CENTRE FOR EDUCATION POLICY
& EQUALISING OPPORTUNITIES

Monetary and time investments in children's education: how do they differ in workless households?

Centre for Education Policy and Equalising Opportunities (CEPEO)

Silvan Häs, Jake Anders and Nikki Shure

Working Paper No. 20-10

April 2020

Highlights

- Children growing up in workless households are of increasing interest for social policy. While the academic literature discusses the “attainment gap” of children growing up in workless households, little research has been done to understand how workless parents invest their resources in their children’s education, as compared to non-workless parents.
- We analyse data from the 2012 cycle of the Programme for International Student Assessment (PISA), matching children from a workless background to children with an otherwise similar background whose parents are not workless.
- Our results indicate that workless parents tend to invest less money in their children’s education. However, we do not find evidence that children from a workless household receive fewer paid-for out-of-school lessons.
- Children growing up in workless households receive greater time investments by their parents in the form of homework help. Our estimates indicate that this association is stronger in single parent households.
- Policymakers should be aware of these differences in monetary and time investments between workless and non-workless households in the design of policy aiming to reduce the attainment gap between these groups.

Why does this matter?

Our results – together with more country-specific research – can help inform policy makers interested in closing the attainment gap of children from workless households.

Monetary and time investments in children's education: how do they differ in workless households?

Silvan Häs s.has@ucl.ac.uk
Jake Anders jake.anders@ucl.ac.uk
Nikki Shure nikki.shure@ucl.ac.uk

UCL Institute of Education*

21st April 2020

Abstract

Around 9% of children in the European Union live in households in which no parent is working. Children living in these workless households are of increasing interest to researchers, policy makers, and the wider public. Workless households not only have lower income on average but are also widely considered to be at risk of social exclusion. In this paper, we study the relationship between parents' employment status and their time and monetary investments in their child's education using data from the Programme for International Student Assessment (PISA). We use matching methods and regression analysis to compare educational investments made in children from a workless background to children with at least one working parent, but otherwise very similar background characteristics. Our analyses indicate that parents' worklessness is associated with lower monetary investments in their children's education. However, we do not find a difference in monetary investments in the form of commercial tutoring. In terms of time investments, we find that workless parents – especially workless single parents – spend more time helping their child doing homework. These findings could help guide future social policy aimed at equalising opportunities for children living in workless households. Conditional on a deeper understanding of the implications of worklessness on country level, measures such as educational vouchers or stipend programmes specifically aimed at socially disadvantaged children could be introduced.

*The authors are part of the European Training Network *OCCAM*. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 765400. For more information, see https://etn-occam.eu.

1. Introduction

Public attention is frequently drawn towards various issues arising from unemployment and worklessness. For instance, in Germany families living off social benefits ("Hartz IV") are subject to regular media coverage, especially highlighting insufficient funds for children and their education (Öchsner 2018; Schäfer 2019). Adults not working and households with no adult household member in employment (i.e. workless households) have been identified as a vulnerable group, both by policy makers and researchers (e.g. McClelland 2000; Mynarska et al. 2015).

Aside from public interest, a broad range of scientific literature connects socio-economic background and education. Links between parents' income and children's educational achievement have been researched in the past decades (e.g. Yeung et al. 2002). Two main mechanisms have been identified linking parental income with their child's school performance: parental stress and parental investment (e.g. Conger et al. 1992; Yeung et al. 2002). On the one hand, low income puts stress on parents and therefore limits their ability for 'good parenting'. On the other hand, with low income come budget constraints, which reduce the potential for parents' monetary educational investments in their children.

Economists have modelled parental investments in a child's education as investments in human capital. Models, such as those introduced by Becker and Tomes (1986) and Aiyagari et al. (2002), suggest that utility maximising parents caring for their children's future utility level choose to invest money in their child's education depending on their own wealth and productivity. Solon (2004) suggests that higher income parents not only have greater possibilities to invest in their children, but also have a greater incentive to do so.

These theoretical findings have been supported by empirical studies, which establish the proposed link between parental income and monetary investments in a child's education. Richer parents are found to spend more money on their child's primary and secondary education compared to less wealthy parents (Mauldin et al. 2001). In the United States, the proportion of income allocated to the education of a child below the age of 24 has been around five percent for most households between the 1970s and 2000s (Kornrich and Furstenberg 2013). Households belonging to the lowest 10% in terms of income, however, spent around 20% of their income on their child's education, imposing greater restrictions on their household's budget compared to richer households. Parental education level appears to be an important factor in explaining future educational success (e.g. Black et al. 2005) and is also found to be strongly associated with the amount of money parents spend on their child's education (Mauldin et al. 2001).

However, monetary investments are not the only investments that parents can make in their child's education; they can also use their own time. In the economic model of Aiyagari et al. (2002), parental time investments are deemed to be efficient only if made by individuals with low productivity levels elsewhere in the economy, while high-skilled parents would rather pay for child care. One result from this model is that highly productive parents do not invest their time in their child's education unless if forced to do so by an imperfect child care market.

Some empirical studies, however, find results in contrast to the theoretical predictions by Aiyagari et al. (2002): as with monetary investments, richer and better educated parents tend to allocate more of their time towards their child's education compared to those less wealthy and well educated (Guryan et al. 2008). Hence, parents who have high 'human capital' to pass on to their children also spend more time doing so.

Another important factor for time investments identified in previous empirical research is family composition. Children growing up with two biological parents receive more time investments than those living with a single parent (Kalil et al. 2014).

Several studies have analysed parental occupational status and time investments. Parents who are not in employment tend to spend more time investing in their child's education (Guryan et al. 2008). Furthermore, for mothers who do not hold employment when their child is younger than five years old, Bernal (2008) find a positive effect on the child's cognitive abilities. While these studies focus on individual parents not working, Parsons et al. (2014) focus on workless households, i.e. households in which no adult households member is in employment. They observe that in the average workless household parents tend to spend less time reading to their young children or taking them to the library (ibid).

Literature that has explored the link between parental occupation status and a child's outcomes has had two main foci: firstly, estimating the link between parents' worklessness and children's school performance, or more generally, their educational trajectories. In the context of the United Kingdom, the UK Department for Work and Pensions (2017) recently analysed children from workless households and their performance in school. This study found that one in eight children lives in a workless household. In their educational career children from a workless background appear to struggle more often and perform worse; the findings led the British government to widen the scope of their *Troubled Families Programme*. Also, parental worklessness is found to be associated with lower parental education levels and single parenthood as well as lower educational attainment of young children (Parsons et al. 2014).

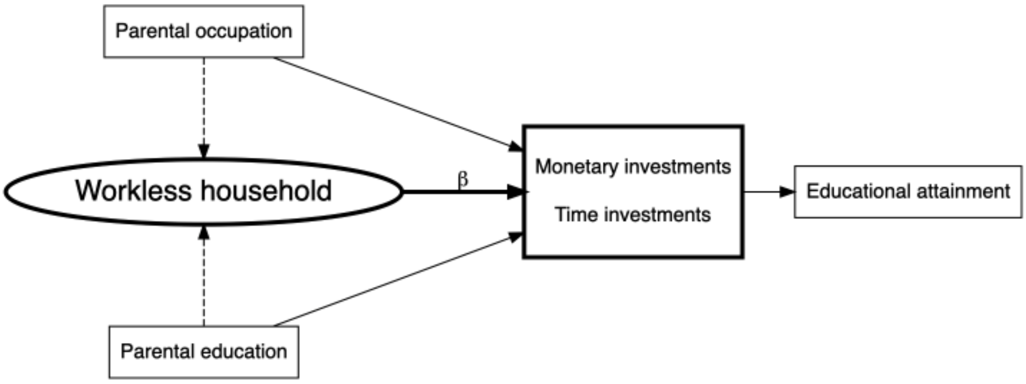
The other major focus of this literature has been on more general economic intergenerational effects of worklessness. Worklessness has been identified as a factor for social mobility. Especially in weak labour markets, adolescence in a workless household increases the probability of becoming workless later on in life (Friedman et al. 2017; Macmillan 2014, findings for the UK). A strong association between worklessness and educational attainment is observed in many European countries (Macmillan et al. 2018). Especially for boys there appears to be a link between growing up with a workless background and being jobless and poor later on in life, especially in countries in which the attainment gap is found to be strong. However, even though worklessness correlates between generations in many countries, many studies do not detect causal effects (Mäder et al. 2015, for intergenerational transmission of worklessness between fathers and sons). Schoon (2014) does not find evidence for a 'culture of worklessness' with a causal link between growing up in a workless household and being unemployed as a young adult. While much of the literature focusses on the relationship between fathers and sons, Berloffa et al. (2017) find that maternal worklessness is associated with a lower risk of youth unemployment in many European countries.

Overall, the current literature on parental worklessness does not focus on its implications on educational investments, and similarly the academic discussion of parental educational investments leaves out worklessness. In this study, we contribute to the existing literature in several ways.

First, worklessness and both monetary and time investments in education have been identified as potential factors affecting educational attainment and intergenerational mobility. However, there remains a gap in understanding how worklessness and educational investments are linked. While studies such as UK Department for Work and Pensions (2017) find differences in educational performance between children from workless and working background families, this study aims to shed light on specific mechanisms which lead to such differences.

Second, building on Macmillan et al. (2018), we use the extensive background data of the *Programme for International Student Assessment* (PISA) to look at parents' worklessness and its implications. Where the existing literature on monetary and time investments in education mainly uses country-specific datasets – with most research conducted for the United States and more recently the United Kingdom – this study includes a wide range of OECD and partner countries covered by the PISA study.

Figure 1: Parental worklessness and investments in their child's education.



Notes: β indicates the strength of the association between a workless household and parental investments in their child's education. Figure based on Jerrim and Macmillan (2015) and Leibowitz (1974)

In short, the literature introduced above suggests that parents with more money spend more on their child's education and parents not in employment spend more time with their child. However, other factors such as parental education levels need to be taken into account. Figure 1 sums up which factors are likely to affect the amount of money and time parents invest in their child's education – which influences a child's educational attainment. While parental characteristics such as education and occupation levels can affect monetary and time investments in a child's education directly (e.g. highly educated parents tend to spend more time with their child), they are also linked to the propensity of a household to become workless. In this study, we are interested in how monetary and time investments are different in workless households, correcting for observable family characteristics (i.e. β in the figure).

As workless parents may have less money but more time to spend, the hypothesis we test in this paper is that workless parents invest less money - but more time - in their child's education. Our findings partly confirm the first part of this hypothesis: while workless households generally spend less money on children's education, we do not find that children in workless households are less likely to receive commercial out-of-school lessons. Overall, workless parents – especially single parents – tend to spend more time helping their child with its homework, confirming the second part of our hypothesis. However, results are not clear-cut and differ across subsets of the PISA data.

The remainder of this paper is structured as follows. In Section 2, the data from the PISA study is introduced. We highlight differences in the characteristics between children with workless and working parents in Section 3. Section 4 introduces the empirical method. This includes the preprocessing of the data using a matching approach, ensuring only children from workless and non-workless households that are similar in observed background characteristics are compared, as well as the subsequent analysis of the matched sample. A detailed description of the results can be found in Section 5 with additional robustness checks in Appendix Section C. Finally, Section 6 discusses the paper's key findings and reflects on the advantages and limitations of researching worklessness and investments in education using the PISA data.

2. Data

In this paper, we use data from the 2012 cycle of the Programme for International Student Assessment (PISA). Since its launch in 2000, PISA has assessed 15-year-old school children's reading, mathematics and science skills. The PISA study is conducted every three years in all OECD countries as well as a growing number of partner countries.

After participating in the PISA tests, all students answer questions about themselves, their family background, learning habits, and more. Starting with PISA 2006, some participating countries also ask parents to fill in a questionnaire containing questions similar to those in the student questionnaire as well as additional information on the family's financial background, educational spending, and time spent with their children. In 2012, participating countries and regions with parent questionnaire data are: Belgium, Chile, Croatia, Germany, Hong Kong, Hungary, Italy, Macau, Mexico, Portugal and South Korea. For more information about all participating countries and jurisdictions, and their sample sizes, see Table A.1. This study makes use of the detailed information provided by the student background questionnaire and – where available – parent questionnaire.

The background questionnaire has varied significantly over the years, such that not all information required for the analysis of parental worklessness and their educational investments is available in each PISA cycle. Parents' occupational status (workless or non-workless) is provided by the student questionnaires of PISA cycles 2000, 2003, 2009 and 2012. Only the PISA 2012 data contains items related to both parental homework help and *commercial* out-of-school lessons from both the student and parent questionnaires, making it best suited research how workless parents invest their time and money in their child's education. In total around 480,000 pupils from 65 countries or jurisdictions were surveyed by PISA in 2012, of which around 100,000 pupils in 11 countries also returned a parent questionnaire.

Key for all analyses conducted throughout this paper is measurement of parental worklessness. Following Macmillan et al. (2018) and in line with the definition used by the UK government (Great Britain. Office for National Statistics 2019), a household is considered workless if none of the parents[1] living in the household hold any kind of employment. Reversely, a household is not workless if at least one parent is in employment. Hence, it is possible to find a parent currently not in employment in a non-workless household. To observe worklessness, it is necessary to know which parents are living with the child as well as details about those parents' employment situation. This information is only available through the student questionnaire, not the parent questionnaire. Thus, a variable indicating a workless background can only be created if the respective questions have been answered by the student.[2]

This definition of worklessness together with the use of PISA data has some shortcomings. For one, worklessness of a household is only measured at one point in time at which the child is 15 years old. This does not reflect the complexity of the occupational biographies of parents but merely reduces it to a one point in time measure. Furthermore, the reasons for worklessness are unknown and therefore the characterisation of the group of children growing up in workless households remains potentially heterogeneous. In this study, it is not possible to differentiate between deliberate worklessness (due to e.g. retirement or wealth) and involuntary worklessness e.g. after losing a job or due to illness.

The PISA dataset provides several background characteristics which we use throughout this paper, first to describe different characteristics of workless and non-workless children (section 3) and then as control variables (section 5). The background characteristics used throughout this study contain students' gender, both parents' occupation and education level, a family's immigration background, and whether it is a single or two parent household. For workless parents, we use their last held employment to determine occupation level. As shown in Section 3, children of workless and non-workless parents differ in various ways. Table 1 lists all background variables we use throughout this study.

Parental occupation is reported by the students (and parents in countries with a parent questionnaire) in form of an open question. If a parent is not working, students are asked to

[1] This includes step-parents and legal guardians.

[2] This excludes Israel from this study as households composition is not assessed. Also, due to a very small amount of workless background children (only 10 observations) we exclude Liechtenstein from further analysis.

Table 1: Description of variables used for matching.

Variable	Type	Source	Formula symbol
Mothers' occupation level	Continuous	SQ / PQ	O_m
Fathers' occupation level	Continuous	SQ / PQ	O_f
Gender	Binary	SQ	G
Immigration status	Binary	SQ	I
Mothers' education	Categorical	SQ / PQ	E_m
Fathers' education	Categorical	SQ / PQ	E_f
Single parent household	Binary	SQ	S
Mother's age	Categorical	PQ	A_m
Father's age	Categorical	PQ	A_f

Notes: SQ *means the variable is measured through the student questionnaire and* PQ *indicates it is observed from the parent questionnaire. Parental education levels are observed in six ISCED categories and recoded into three categories,* low, medium *and* high *to ensure comparability between student and parent questionnaire as well as between countries. Parental age is observed in categories* younger than 36, 36-40, 41-45, 46-50 *and* 51 or older.

name the last held occupation. The occupations are then ultimately coded into a continuous quantitative variable using the International Socio-Economic Index of Occupational Status (ISEI) (OECD (Organization for Economic Cooperation and Development) 2014). Parental education level is assessed in both parent and student questionnaire by using a multiple choice question. Jerrim and Micklewright (2014) show that children's reports of their parents' occupation are generally accurate while for parental education level children and parents show 'moderate' agreement in their responses.

For the analyses of educational investments as observed from the student questionnaire, we only use background information reported by students. In use background information reported by parents only when focussing on dependent variables from the parent questionnaire. This is done to ensure consistent data sources (both across countries and within countries).

Monetary investments can be observed indirectly in the student questionnaire and directly in the parent questionnaire. In the student questionnaire, around two thirds of students are randomly assigned booklets containing the following question:

> *Thinking about all school subjects: on average, how many hours do you spend each week on the following?*
>
> *d) Attend out of school classes organised by a commercial company, and paid for by your parents*
>
> *e) Study with a parent or other family member*

(PISA 2012 student questionnaire)

We use the former part of the question (d) as a proxy variable for monetary investments – given that such classes are usually expensive and represent a sizeable financial commitment by parents in their offspring's education (Kassotakis and Verdis 2013; Dang and Rogers 2008). We use the second item (e) to measure parental time investments. The response rate of students presented with this question is 87% commercial tutoring and 90% for parental homework help.

The parent questionnaire provides a more direct view on parents' monetary investments by asking specifically for the amount of money parents annually spend on their child's education:

> *In the last twelve months, about how much would you have paid to educational providers for services?*
>
> - *Nothing*
> - *More than **0** but less than **W***
> - ***W** or more but less than **X***
> - ***X** or more but less than **Y***
> - ***Y** or more but less than **Z***
> - ***Z** or more*

(PISA 2012 parent questionnaire – monetary investments)

Each country decides on the values for **W**, **X**, **Y**, and **Z**. We account for these difference by recoding the item in three categories – *low*, *medium*, *high* – to ensure comparability across countries.

Furthermore, parents are asked how frequently they help their child with its mathematics homework:

> *How often do you or someone else in your home [help you child with his/her mathematics homework]?*
>
> - *Never or hardly ever*
> - *Once or twice a year*
> - *Once or twice a month*
> - *Once or twice a week*
> - *Every day or almost every day*
>
> *(PISA 2012 parent questionnaire – time investments)*

We use this variable to complement the analysis of the time investment variable from the student questionnaire (option *e* from student questionnaire). However, the item from the parent questionnaire is limited to mathematics homework only. A potential minor limitation of the measures for time investment arises from the phrasing of the question: it is not restricted to parental homework help but allows also for help from other people living in the household, such as grandparents or siblings. The response rate for the questions from the parent questionnaire is 85% and 86%, respectively.

2.1. Missing data

As in many other surveys, the PISA study suffers from missing data. Table 2 shows the proportion of missing data in selected variables from the student questionnaire. Particularly high rates of missing data can be observed for worklessness as well as fathers' – and especially mothers' – occupation levels. As the workless variable can only be constructed if the household composition is known, missing values in household composition cause missing values in the workless variable. Furthermore, if a child is living with e.g. its grandparents or in a foster home, this results in a missing value for worklessness. The return rate of the parent questionnaire is high in most countries (87%), noteworthy exceptions being Belgium and Germany with response rates of 49% and 58%, respectively.

A closer look on missing values in the parental occupation variables from the student

Table 2: Missing data

	Household comp.	Jobless house-hold	Mothers' educa-tion	Fathers' educa-tion	Mothers' occupa-tion	Fathers' occupa-tion	Immigra-tion status
OECD	.09	.13	.05	.08	.20	.13	.03
Partner countries	.11	.20	.03	.05	.33	.17	.03
Total	.10	.15	.04	.07	.25	.14	.03

Notes: Rate of missing observations for selected variables from the student questionnaire.

questionnaire reveals some interesting patterns. First, mothers' occupation is frequently missing in some countries, such as the United Arab Emirates (62%), Tunisia (69%), Jordan (78%) and Turkey (82%), much higher rates than for fathers' occupation (betweeaan 13 and 26%) or parental education (between 2.6 and 5.8%). This indicates that in these countries mothers' occupation may be unknown because it is less common for mothers to be employed.[3] Second, in single parent households information for the remote parent not living with the child is missing more frequently. 32% of children living with a single mother don't report their fathers' occupation compared to only 9% in a two-parent households (similar, but less pronounced for single fathers). As most single parent households are in fact single mother households, this affects missing data in fathers' occupation more.

Analysing only students with complete information could potentially cause bias as observations are not likely to be left out at random. We use *multiple imputation* to impute missing values (Rubin 1987), using the R-package *mice* (Buuren and Groothuis-Oudshoorn 2011). In the following paragraphs we describe how we adjust for potential missing data mechanisms in our imputation algorithm for both data from the student and parent questionnaire.

While the questionnaires are mostly the same in all participating countries, both the extent to which missing data occurs and the mechanisms that cause them may differ between countries. As a result, we perform multiple imputation for all participating countries separately. Furthermore, we do not impute households' occupation status (workless or non-workless), as this is the dependent variable of interest in this study. Missing values in this variable are mainly caused by unknown or unusual (i.e. no parent present) household composition. Last,

[3]According to World Bank data (World Bank 2019a; World Bank 2019b), female labour force participation in countries with high amounts of missing data for mothers' occupation is much lower than for males. The correlation coefficient between the difference of female and male labour force participation and the difference of missing data for mothers and fathers in PISA is 0.87, indicating a strong correlation between the two. In short: in countries where women are employed at lower rates than men, children report mothers' occupation less frequently than fathers'.

by design of PISA 2012, only around 2/3 of all students were assigned questionnaire booklets containing the outcome variables *commercial tutoring* and *parental homework help*. We exclude students who did receive questionnaires not containing the outcome variables from further analysis of dependent variables from the student questionnaire without causing bias as they are missing completely at random. All of the above results in a student population of interest of 270,175 observations from 63 countries, of which 101,555 require some degree of imputation. For the analysis of the parent questionnaire, out of a total of 101,175 observations from 11 countries 46,903 require imputation on at least one of the variables of interest.

As discussed previously, information about the remote parent is often missing in single parent households. To take this into account, we apply different imputation algorithms for children in two-parent households and single parent households.[4]

In countries with a parent questionnaire, both parents and students report on parental occupation level and parental education level. We make use of this additional information when imputing. To account for the proportion of missing data in the dataset at hand, we create 30 imputed datasets which we use for all subsequent steps of analysis. We then reconcile the results by using *Rubin's Rule*.

3. Differences between children in workless and non-workless background families

3.1. Background Variables

The background variables provided by the PISA study provide insights into each participating child's personal background. This section focusses on the circumstances children with workless parents grow up in and how this compares to those from a non-workless background. In doing so, this helps to provide a better understanding of the similarities and differences between these two heterogeneous groups of children that will be used in the analyses conducted later on in this study.

The PISA study covers many countries and jurisdictions with different prevalence of work-

[4]For Japan, Perm (Russia), and Iceland there are too few observations of either single or two-parent households to split the dataset up and we impute the pooled dataset.

lessness. Around seven percent of students in OECD countries and 14% in partner countries report living in a workless household (see Table A.1 for worklessness rates in each participating country). On average, workless background children in OECD countries score 38 points lower on the PISA test in mathematics (36 in reading and 37 in science) than their peers from non-workless background. In partner countries the difference is slightly less pronounced with a difference in mathematics score of around 25 points (25 in reading, 23 in science). PISA scores for children from a workless background are the same as for those from working households only in Macau, Singapore, Thailand and Albania. Scores for workless children are not substantially lower, or sometimes even higher, only in Macau, Singapore, Thailand and Albania.

Next, Table 3 reports summary statistics for parents' occupation level in OECD and partner countries for workless and non-workless household parents.[5] First, note that for both workless and non-workless parents the observed occupation levels range from 11 (i.e. subsistence farmers[6]) to 89 (i.e. judges). This means that at least some parents with the highest occupation levels are workless and some with the lowest are in employment, ensuring common support in this variable. Overall, parents in workless households are overrepresented in lower occupation levels. The median workless household mother has an occupation level of around 28 (e.g. sales assistant), whereas the median for non-workless household mothers is 45 (e.g. secretary). For fathers from a workless household, the median value is 28 (elementary worker), whereas as those from a working household have a median occupation level of 36 (electrician). Thus, on average workless parents' last job was in 'lower' occupations compared to parents in a non-workless household.

Very similar observations can be made for parental education level, where workless parents have lower education levels compared to those in employment. In OECD countries, around 17-18% of parents from a non-workless household have a low education level[7] compared to 34-35% in workless households. A similar but less pronounced pattern can be observed for participating partner countries.

The living and family conditions of children from a workless background differ in various other ways. Around half of all children living in a workless household in an OECD country are raised

[5]Being "not workless" refers to the household not the individual parent. If the mother holds a full time job and the father is workless, this previous occupation level contributes to the "not workless" distribution.

[6]This includes crop and livestock farmers, hunters and gatherers, and trappers.

[7]ISCED 2 or lower, equivalent to 9 years of schooling or less in the United States (see Miller 2007, for US equivalents).

Table 3: Summary statistics for parents' occupation level in OECD countries for workless and non-workless background children.

(a) Mother's occupation level

	min	25%	median	mean	75%	max	N
not workless	11	27	45	46	65	89	203,306
workless	11	23	28	36	50	89	10,077
Total	11	25	44	45	65	89	213,383

(b) Father's occupation level

	min	25%	median	mean	75%	max	N
not workless	11	26	36	44	62	89	219,519
workless	11	21	28	34	44	89	13,700
Total	11	26	35	43	62	89	233,219

Notes: Occupation levels of mothers (upper table) and fathers (lower table) in workless and non-workless households, classified using the ISEI scale. Reported figures are minimum, 25th percentile, median, mean, 75th percentile, maximum and the number of observations (N). The number of observations differs for mothers and fathers because of missing information on the occupation (last) held. No weights applied.

by a single parent as opposed to only around 12% in non-workless households. In general, parents in workless households appear to be older on average.[8] and are more likely to have an immigration background. Furthermore, workless households appear to have a lower annual income.[9]

3.2. Parental investments in education

As described in greater detail in section 2, monetary investments are observed from the student questionnaire using the proxy variable *commercial tutoring*. As shown in Table 4, around 16% of students in OECD countries and more than 34% in partner countries attend at least one hour of commercial tutoring per week. The raw difference between workless and non-workless background children appears to be very small on average in both OECD and partner countries.

The differences between countries in terms of prevalence of commercial out-of-school lessons is shown in Figure 2. While in countries such as Norway, Sweden and Denmark only a very small proportion of around four percent of students report to attend any commercial tutoring,

[8]Information only available for countries with parent questionnaire. Details in appendix.

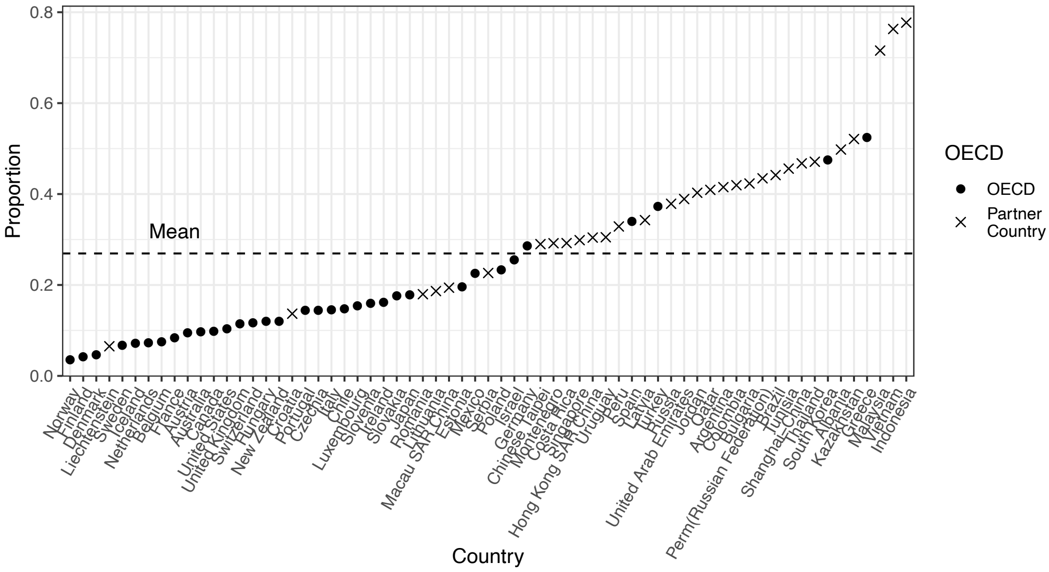
[9]Information available for countries with parent questionnaire except Italy

Table 4: Commercial tutoring for children from a workless and non-workless background.

	not workless	workless
OECD	0.161	0.163
Partner Country	0.371	0.348

Notes: Difference in the prevalence of at least one hour per week of commercial tutoring in workless and non-workless households, not adjusted for any background characteristics. Senate weights applied.

Figure 2: Proportion of children with at least one hour of commercial tutoring per week across countries.



Notes: Countries on horizontal line sorted from smallest to greatest proportion of commercial tutoring. The dashed line shows the average proportion. Student weights applied.

around half of students the OECD countries South Korea and Greece and up to almost 80% of students in Malaysia, Vietnam and Indonesia report attending at least one hour per week in commercial out-of-school lessons. Overall, in most OECD countries commercial tutoring seems to be far less common than in non-OECD countries.

In the parent questionnaire, educational expenses are measured more directly. [10] On average, around 35% of non-workless parents fall into the lowest expense category, whereas more than 45% of workless parents only spend low amounts on their child's education. Non-workless parents are in turn overrepresented in the medium and high expense category. The raw difference between workless and non-workless parents is comparably high in Chile, Hong Kong,

[10]Data from the parent questionnaire on education expenses provided in the PISA dataset is regrouped as described in section B.1 in order to ensure comparability across countries.

Table 5: Parental homework help for children from a workless and non-workless background.

	not workless	workless
OECD	0.436	0.447
Partner Country	0.524	0.542

Notes: Difference in the prevalence of at least one hour per week of parental homework help in workless and non-workless households, not adjusted for any background characteristics. Senate weights applied.

Table 6: Parental mathematics homework help for children from a workless and non-workless background.

	Never or hardly ever	Once or twice a year	Once or twice a month	Once or twice a week	Every day or almost every day	N
not workless	0.457	0.112	0.189	0.178	0.0637	80,723
workless	0.495	0.0919	0.155	0.181	0.0761	7,611

Notes: Frequency with which parents report to help their child with his or her mathematics homework for workless and non-workless households. The last column shows the number of observations N. *Data available for countries with a parent questionnaire. All values are raw values and have not been adjusted for any background characteristics. Senate weights applied.*

Hungary and Portugal. A very small raw difference can be observed in Belgium and Denmark.

We measure parental time investments in their child's education using the reported number of hours per week spent studying with a parent or other family member. In OECD countries, the rate of students reporting that they spend at least one hour per week studying with a parent or family member is around 45% and for partner countries slightly above 50%. Similarly, for countries with a parent questionnaire parents report how regular they assist their child with its *mathematics* homework. Around 50% of parents report to help their child with its mathematics homework at least a few times per year, around 1/4 of parents reporting to help on a weekly basis (see Table 6). Overall, there are only small raw difference between children from a workless and those from a non-workless household.

In this section we give a broad overview of the situation which children from workless and non-workless households grow up and live in. Workless parents tend to be less well educated, have held lower occupation levels, are more likely to have an immigration background, and are often single parents. However, the raw difference in educational investments of money and time is very small. The next sections introduce statistical analysis methods relevant to measure the association between worklessness and parental educational investments more accurately, making sure to not compare apples with oranges.

4. Methods

This section introduces the methods used to understand how monetary and time investments in a child's education differ if parents are workless compared to households in which at least one parent is in employment. Section 3 has shown how workless and non-workless households differ in many background characteristics, while educational investments are roughly the same regardless of parental occupation status, if not adjusting for any of the differences in background characteristics. Now the aim is to take into account all those background characteristics in order to assess the association between worklessness and parental educational investments were those other characteristics to be equalised.

For this, we use a combination of two methods: First, we use a *matching* approach to construct a sample of pupils from workless households that can realistically be compared with the sample of pupils from non-workless households. The basic intuition is that for each child from a workless household we find a sufficiently alike counterpart, in terms of their observable characteristics, in a workless household. All those identified in this way are included in the matched sample. Workless background children who do not have a counterpart growing up in similar circumstances but with non-workless parents (and vice versa) are discarded.

Second, we apply linear and logistic regression modelling to the matched dataset. This helps account for any remaining imbalance in terms of observable characteristics and improves the precision of the measured association between parental occupation status and parental educational investments.

4.1. Matching

As described in section 3, children from a workless background are different on average from those whose parents are in employment in many regards. Household characteristics such as parental occupation and education, single parenthood and immigration status are likely to affect both the probability of worklessness and the outcomes of interest related to educational investments.

In this study,we use matching as a tool to construct a comparison group of pupils from non-workless background with whom we can meaningfully compare pupils from workless backgrounds. Even though growing up in a workless household is not a 'treatment' in the normal sense, matching helps reduce bias and model dependency in the estimation of the association between worklessness and educational investments (Ho et al. 2007). Sometimes, regression modelling alone is used to achieve similar aims. However, when comparing a 'treatment' group with a very dissimilar comparison group, this can suffer from problems with extrapolation, leading to unreliable estimates for the regressors in the model. This is potentially the situation here, given how different workless and non-workless households may be.

In the matching literature it is often recommended to apply different matching specifications and use the best balanced dataset for further analysis. The robustness of results can then be tested across a range of techniques that produce datasets with a similar good balance.[11] However, it is not that clear how to define the best balance (Stuart 2010). Ideally it would be possible to compare the multidimensional distribution of all matching variables between 'treatment' and 'comparison' group and settle for the approach that minimises this difference. As this is not feasible in practice, researchers are left with no incontestable way of comparing different matched datasets.

Because of these shortcomings and the fact that most matching approaches applied to this dataset resulted in very similar balance improvements, we present the matching methods used for the analyses in the main body and the resulting satisfactory balance improvements of this matching approach in this section. We present the results obtained from the subsequent analysis of this matched dataset in section 5. Additionally, in the Appendix we present balance improvements of selected other matching approaches as well as the exact same analyses as robustness checks to all findings.

As with our imputation of missing data, we run the matching algorithm separately for each country as well as for single and two-parent households.[12] For the implementation of matching in this study, we use the R package *MatchIt* (Ho et al. 2011). This package implements many matching methods and allows for detailed specifications.

[11] In the context of PISA data and matching, e.g. Rutkowski et al. (2018) use different matching techniques and report results as robustness checks.

[12] As there is too few observations for Japan, Iceland and Perm (Russia), we do not split the dataset up between single and two-parent households.

4.1.1. Matching Methods Applied

Next, we introduce the matching methods and specifications used for preprocessing the data from both student and parent questionnaire for further analyses.[13] An overview of available matching techniques and advice on the implementation in practice can be found in Stuart (2010), which has set the foundation for the following paragraphs.

For all analyses presented in the main body of this paper we use one-to-one nearest neighbour matching: for each observation in the *workless* group this method finds the closest match from the *non-workless* group, according to a distance measure, here the *propensity score*. The propensity score is the probability that an observation with certain characteristics belongs to the 'treatment' – i.e. *workless* – group. As the true propensity score cannot generally be observed, most commonly the propensity score is estimated using a binary response model where the dependent variable is the treatment status (i.e. workless or non-workless). We use a logistic regression of the following form to estimate the propensity score:

$$\text{logit}(ps) = \gamma_0 + \vec{\gamma}_1 B_{(\cdot)}, \tag{1}$$

where $\gamma_1 B_{SQ}$ includes all relevant background variables when matching data from the student questionnaire: parental education, parental occupation, and immigration status as well as – for Perm (Russia), Iceland, and Japan – a dummy variable indicating a single parent household. Similarly, $\gamma_1 B_{PQ}$ includes all variables from the student questionnaire and additionally mothers' and fathers' age.[14] Rosenbaum and Rubin (1984) introduce the use of the propensity score for matching as it helps overcome the *curse of dimensionality*.[15] A guide on how to best implement propensity score matching in applied research can be found in Caliendo and Kopeinig (2008).

One-to-one nearest neighbour propensity score matching is the most commonly applied matching method and it offers a wide range of additional specifications from which to choose.

[13]For information about matching methods used as robustness checks, see Appendix.

[14]Hence,

$$\vec{\gamma}_1 B_{SQ} = \delta_1 G + \delta_2 I + \delta_3 O_m + \delta_4 O_f + \vec{\delta}_5 E_m + \vec{\delta}_6 E_f (+\delta_7 S),$$

and

$$\vec{\gamma}_1 B_{PQ} = \delta_1 G + \delta_2 I + \delta_3 O_m + \delta_4 O_f + \vec{\delta}_5 E_m + \vec{\delta}_6 E_f + \vec{\delta}_7 A_m + \vec{\delta}_8 A_f (+\delta_9 S).$$

For a overview of all variables and their formula symbols, see Table 1.

[15]The curse of dimensionality is an issue common to the analysis of data with many covariates: depending on the statistical method used, more covariates (i.e. higher dimension) cause the data to be too sparse which causes the method to perform poorly or fail.

The background variables may predict worklessness differently in each country and for different household compositions (single or two-parent): single parenthood was found to be strongly associated with worklessness in previous studies (e.g. Macmillan et al. 2018), as only one parent needs to be workless instead of two. We take this into account by running the matching algorithm separately for each country and household composition[16], including the estimation of the propensity score. Also, for each country and household composition we require the algorithm to discard those workless and non-workless background children from matching, which are outside the common support of the propensity score. Matching is carried out without replacement. This means, that once an observation from the comparison group has been matched, it cannot be matched to another workless child, even if it were the closest match in terms of propensity score. This prevents one non-workless background child to be matched to several workless background children.

4.1.2. Balance improvement

As the main purpose of matching is to create a balanced dataset in which workless and non-workless background children are very similar in their background characteristics, we check the balance improvement due to matching. Different measures can be used to check the balance of a dataset before and after matching. We mainly use the *absolute standardised bias in means*[17] (reported in this section) combined with visually checking the distribution of the propensity score and mothers' and fathers' occupation levels before and after matching (see Appendix).

Table 7 shows the change in standardised bias of all relevant variables of interest for the student questionnaire. A standardised bias smaller than 0.25 is considered to be balanced enough for further analyses (Stuart 2010; Ho et al. 2007). Before matching, most variables in both two-parent and single households are unbalanced. Matching improves balance very well for all variables, reducing the standardised bias well below the 0.25 threshold with all variables having a standardised bias between 0.00 and 0.05. Moreover, the distributions of

[16] Except Perm (Russia), Iceland, and Japan, where we pool single and two-parent household due to small number of observations and require exact matching on household composition.

[17] The absolute standardised bias in means is computed as follows:

$$SB = \frac{|\bar{x}_T - \bar{x}_C|}{\tilde{s}_T}, \tag{2}$$

where $\bar{x}_T$ and $\bar{x}_C$ denote the mean value of variable x for treatment (workless) and comparison (non-workless) group, respectively, and $\tilde{s}_T$ denotes the observed standard error of the treatment group.

Table 7: Absolute standardised bias in means before and after matching – student questionnaire.

	Two-parent household			*Single parent household*		
Variable	Before	After	Improvement	Before	After	Improvement
Gender	0.06	0.01	78.89%	0.03	0.01	75.46%
Immigration status	0.01	0.01	10.91%	0.12	0.01	90.32%
Occupation level father	0.59	0.03	95.68%	0.16	0.01	96.16%
Occupation level mother	0.56	0.01	97.58%	0.35	0.01	95.83%
Education level father - low	0.51	0.03	94.39%	0.23	0.03	87.27%
Education level father - medium	0.10	0.02	81.62%	0.06	0.01	83.78%
Education level father - high	0.53	0.02	97.08%	0.18	0.04	79.01%
Education level mother - low	0.60	0.02	97.03%	0.38	0.05	87.85%
Education level mother - medium	0.13	0.01	92.87%	0.01	0.00	64.40%
Education level mother - high	0.63	0.01	97.93%	0.41	0.05	87.23%

Notes: Matched dataset generated as described in this section; i.e. one-to-one nearest neighbour propensity score; matching algorithm run separately for each country and household composition. All numbers are averaged over all 30 imputations.
Two-parent household*: Workless background children discarded for lack of common support: 89-158. Unmatched workless background children: 0. Total number of observations in matched dataset: 31,652-31,786.*
Single parent household*: Workless background children discarded for lack of common support: 564-746. Workless background children unmatched: 499-560. Total number of observations in matched dataset: 16,428-16,684. Variation in figures due to random differences between the 30 imputed datasets.*

the propensity score and mothers' and fathers' occupation level of the matched dataset are well balanced over their full support (see Appendix): slight imbalances remain only for low occupation levels of fathers.

Table 8 shows the standardised bias in data from the parent questionnaire before and after matching. Before matching, many variables in both two-parent and single parent households are unbalanced, especially mothers' and fathers' occupation level. Despite initial imbalance being stronger in two-parent households, matching succeeds in bringing all variables' standardised bias in means well below the 0.25 threshold with almost no imbalance remaining. However, some notable imbalance remains for single parent households after matching: the standardised bias in means is brought well below 0.25 with variables measuring mothers' occupation and education level retaining an imbalance of above 0.10.

Table 8: Absolute standardised bias in means before and after matching - parent questionnaire.

	Two-parent household			*Single parent household*		
Variable	Before	After	Improvement	Before	After	Improvement
Gender	0.08	0.01	90.27%	0.05	0.03	38.16%
Immigration status	0.02	0.01	49.05%	0.01	0.01	28.77%
Occupation level father	0.68	0.01	97.84%	0.26	0.06	77.28%
Occupation level mother	0.69	0.01	98.53%	0.48	0.11	78.10%
Education level father - low	0.52	0.01	97.39%	0.28	0.07	74.84%
Education level father - medium	0.27	0.01	95.24%	0.10	0.04	59.87%
Education level father - high	0.40	0.01	97.84%	0.23	0.05	78.87%
Education level mother - low	0.60	0.01	98.05%	0.44	0.12	71.80%
Education level mother - medium	0.34	0.01	96.83%	0.20	0.08	58.95%
Education level mother - high	0.46	0.01	97.78%	0.34	0.08	76.36%
Father Age <36	0.02	0.01	38.98%	0.05	0.02	40.27%
Father Age 36-40	0.04	0.01	64.32%	0.01	0.03	-821.85%
Father Age 41-45	0.13	0.01	93.82%	0.12	0.03	72.85%
Father Age 46-50	0.26	0.01	96.11%	0.11	0.03	75.05%
Father Age >51	0.30	0.01	96.87%	0.18	0.04	76.30%
Mother Age <36	0.10	0.01	88.93%	0.04	0.02	58.05%
Mother Age 36-40	0.03	0.01	61.45%	0.01	0.03	-744.65%
Mother Age 41-45	0.20	0.01	94.93%	0.15	0.01	90.11%
Mother Age 46-50	0.11	0.01	89.62%	0.06	0.01	82.83%
Mother Age >51	0.23	0.01	95.92%	0.19	0.05	72.28%

Notes: Matched dataset generated as described in this section; i.e. one-to-one nearest neighbour propensity score matching; algorithm run separately for each country and household composition. All figures in the table are averaged over all 30 imputations.
Two-parent household*: Workless background children discarded for lack of common support: 9-36. Unmatched workless background children: 0. Total number of observations in matched dataset: 10,590-10,644.*
Single parent household*: Workless background children discarded for lack of common support: 78-129. Workless background children unmatched: 247-285. Total number of observations in matched dataset: 6,014-6,090. Variation in figures due to random differences between the 30 imputed datasets.*

4.2. Regression Modelling

The matched datasets can now be analysed using the same methods one would have applied to an unmatched dataset with the advantage of reduced model dependency. This is the recommended approach throughout the matching literature in order to find the best estimates for the association of interest (e.g. Ho et al. 2007; Stuart 2010).

When analysing the student questionnaire, we focus on the probability of receiving commercial (C_{binary}) and parental (P_{binary}) out-of-school lessons and how it differs between workless and non-workless background children. We analyse two-parent and single parent households separately. These models can be represented as follows:

M1 – commercial tutoring

$$\Pr[C_{binary} = 1] = G(\beta_0 + \beta_1 WL + \vec{\gamma}_1 B_{SQ} + \epsilon_{M_{1.1}}) \tag{3}$$

M2 – parental homework help

$$\Pr[P_{binary} = 1] = G(\beta_0 + \beta_1 WL + \vec{\gamma}_1 B_{SQ} + \epsilon_{M_{2.1}}). \tag{4}$$

where $\gamma_1 B_{SQ}$ controls for the observable background characteristics also used for matching, i.e. gender, immigration status, mothers' and fathers' occupation level as well as their education level (see Table 1). The link function $G(\cdot)$ translates the linear core into probabilities. If G is the identity function, a *linear probability model* is estimated using Ordinary Least Squares (OLS). In the main body of this paper we report results from the linear probability model, as they are easy to interpret and can be estimated best given the data structure at hand. WL denotes the households' occupation status (workless or not-workless) and ϵ represents the error term. As an additional robustness check, we analyse the number of hours spent attending out-of-school lessons. For details and results, see Section C.

M3 – monetary investments; and M4 – time investments The dependent variables obtained from the parent questionnaire are ordinal: parents report on their educational expenses and helping their child with its homework in distinct ranked categories. We use two approaches to analyse this data. In the first approach, we use different cut points[18] to recode the ranked categories into a binary variable, which we then analyse using a linear probability model:

$$\Pr[V = 1] = \beta_0 + \beta_1 WL + \vec{\gamma}_1 B_{PQ} + \epsilon_{M_{3/4}}, \tag{5}$$

where the background variables summarised in $\vec{\gamma}_1 B_{PQ}$ are parental education and occupation, mothers' and fathers' age, and students' immigration status and gender. the dependent variable V represents all binary versions of the original variable using different cut points.

However, by transforming an ordinal outcome variable into multiple binary ones, valuable information from the data is lost within each logistic regression. Therefore, we use an *ordered logistic regression* as second approach which aims at avoiding this issue by using the full information of the categorical dependent variable in a single model, instead of scattering this information across several logistic regressions. In both approaches, logistic regression with cut points and ordered logistic regression, we use the same background variables as shown above. We report results from the ordered logistic regression together with the results from different cut points.

With all linear probability models, we apply country fixed effects to compute the standard errors for the estimates. This is done by introducing country dummies as covariates in the regression. When analysing the data from the parent questionnaire with an ordered logistic regression, we include country dummies in the estimation as country fixed effects can not be implemented. All reported standard errors (and resulting p-values and confidence intervals) are computed clustering at the country level.[19]

We apply models using data from the student questionnaire (M1 to M4) to the fully matched samples for two-parent and single parent households as well as to several subsets of this. Most importantly, we separately analyse OECD countries and partner countries. As described at several points in Section 3 OECD and partner countries differ in many regards. For instance it

[18]For clarification, consider a dependent variable consisting of three ranked categories, $A < B < C$. This variable can now be transformed into a binary dependent variable by merging two of the original categories. The first option is to merge A and B to a new category such that the resulting new depending variable γ takes value 1 if category A or B holds true, and 0 in case of category C. The second option would be to merge B and C, such that $\gamma = 1$ if category A holds true, and 0 otherwise.

[19] From the way PISA data is collected, it would be most natural to cluster at school level. Due to the comparably small number of observations in some countries this is not feasible. Therefore, we cluster on country level instead.

is less common in OECD countries to receive any form of out-of-school lessons. Furthermore, we distinguish between countries with high public spending on education (above median according to UNESCO Institute for Statistics (UIS) (2020)) and countries with low education expenditure (below median). Lastly, we state countries with a parent questionnaire separately to allow for comparison with additional results obtained for those countries.

Note that the analyses for each subset do not involve a different matching approach, as all matching is performed within countries.

5. Results

In this section we present our estimates for the association between parental worklessness and money and time parents invest in their child's education. We obtain our estimates by first preprocessing the data with *Matching* techniques and subsequently using regression analyses for estimation (see Section 4). We present the point estimates for the regression coefficients of the worklessness variable (β_1) as well as corresponding standard errors and significant levels. When reporting results from an ordered logistic regression from the analysis of the parent questionnaire, we report the *average marginal effect* rather than the actual model estimates to ensure comparability with the results from the linear probability models. The marginal effects of worklessness show by how much the probability of being in a higher category of the dependent variable differs if parents are workless instead of non-workless, depending on a broad range of background characteristics.

First, we present results around monetary educational investments in Section 5.1. Results for time investments are shown in Section 5.2. Note that robustness checks (heterogeneity analysis, variations in regression, no matching, differently matched sample; see Section C) are in line with the findings presented in this section.

5.1. Monetary Investments

Models M1 and M3 estimate the association between parental worklessness and their monetary investments in their child's education (see section 4.

Table 9 shows the results from a linear probability model M1 applied to *all countries*, *OECD*

Table 9: M1 – Association between worklessness and commercial tutoring from a linear probability model applied to different subsets of the matched PISA data.

	Two-parent household		*Single parent household*	
Data	Estimate	Standard error	Estimate	Standard error
All countries	-0.009	0.008	-0.008	0.008
OECD	-0.009	0.011	-0.001	0.009
Partner countries	-0.010	0.011	-0.016	0.013
PQ	0.004	0.011	-0.015	0.012

$*\ p < 0.1$, $**\ p < 0.05$, $***\ p < 0.01$

Notes: Regression on matched sample run separately for two-parent and single parent households. Adjusted for gender, mothers' and fathers' occupation level, mothers' and fathers' education level, and immigration status. Standard errors clustered at country level. Country fixed effects.

countries, *Partner countries*, and *Countries with a parent questionnaire*. Data from the student questionnaire is used with a binary variable indicating whether or not a child attends commercial tutoring as dependent variable. The point estimates of the association between worklessness and commercial tutoring are mostly negative, close to 0, with none being statistically significant.

Next, we analyse the data from the parent questionnaire using model M3. Table 10 presents the estimates for the association of worklessness with parental educational expenses. As educational expenses are measured in ordered categories, we present both the results of linear probability models with different cut points of the categorical expense variable as well as the results of an ordered logistic regression (see Section 4). The results indicate that workless background children living in both single parent and two-parent households are more likely to be in a lower expense category compared to their non-workless background peers. Our estimates suggest that children from a workless background are 2.4–2.5 percentage points more likely to be in the lowest expense category and around 3.9–4.0 percentage points less likely to be in the highest expense category. There appears to be no difference between children living in a single parent or two-parent household. Robustness checks confirm these results with similar or stronger point estimates (see Section C). It is noteworthy that the analysis of the student questionnaire did not show any difference in the prevalence of commercial tutoring in countries with a parent questionnaire. This suggests that commercial tutoring does not capture overall spending on a child's education: in countries with a parent questionnaire there appears to be robust negative association between parental worklessness and educational expenses as a whole, while this differences cannot be detected when focusing on commercial tutoring only – a specific kind of educational expense.

Table 10: M3 – Association between parental worklessness and monetary investments using data from the parent questionnaire.

	Two-parent household		*Single parent household*	
Regression	Estimate	Standard error	Estimate	Standard error
low | medium, high	-0.039***	0.013	-0.040**	0.017
low, medium | high	-0.025***	0.009	-0.024**	0.011
Ordered logistic regression	-0.036***	0.010	-0.037***	0.013

$* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$

Notes: First two rows – linear probability models with different cut points for the categorical dependent variable: between low income and merged medium and high income (first row) and between merged low and medium income and high income (second row). Standard errors clustered at country level. Country fixed effects.
Third row – ordered logistic regression. For comparability we report the average marginal effect *and the corresponding standard error, which allows the magnitude of the regression coefficients to be compared. Standard errors clustered at country level. Country dummies included (no country fixed effects).*
Adjusted for gender, mothers' and fathers' occupation level, mothers' and fathers' education level, immigration status, and mothers' and fathers' age.

5.2. Time Investments

In this section we present the results from models M2 and M4, estimating the association between parental occupation status and parents' time investments in education.

Table 11 shows the results from analysing the student questionnaire. We find a statistically significant association between parental worklessness and parental homework help in both two-parent and single parent households, while being larger in single parent households. In two-parent households, we find this association only in OECD countries, where children living in a workless household are around two percentage points more likely to being helped by their parents with their homework. The estimate for partner countries and countries with parent questionnaire is small and not significant. In single parent households, we find a significant association in partner countries (5% level) and OECD countries (10% level). Here, children from a workless background are around 3 percentage points more likely to receive parental homework help, compared to children with similar background characteristics who live in a non-workless household. However, we do not find this association in countries with a parent questionnaire.

As Table 12 shows we do find no difference in parental *mathematics* homework help between workless and non-workless background children when analysing data from the parent question-

Table 11: M2 – Association between worklessness and parental homework help from a linear probability model applied to different subsets of the matched PISA data.

	Two-parent household		*Single parent household*	
Data	Estimate	Standard error	Estimate	Standard error
All countries	0.013**	0.007	0.026***	0.009
OECD	0.021**	0.010	0.021*	0.012
Partner countries	0.007	0.009	0.031**	0.015
PQ	0.001	0.012	0.013	0.018

$*\ p < 0.1$, $**\ p < 0.05$, $***\ p < 0.01$

Notes: Regression on matched sample run separately for two-parent and single parent households. Adjusted for gender, mothers' and fathers' occupation level, mothers' and fathers' education level, and immigration status. Standard errors clustered at country level. Country fixed effects.

naire. The dependent variable is categorical and indicates how regularly parents report to help their child its mathematics homework. The point estimates are close to 0 and insignificant for all cut-off points of the linear probability models as well as for the ordered logistic regression, both for two-parent and single parent households.

These results suggest that overall there is a positive association between parental worklessness and time investments in a child's education. Unsurprisingly, this association appears to be stronger in single parent households compared to two-parent households: many two-parent households only have one parent in full-time employment which leaves time for the other parent to help the child with their homework.

However, the subset of countries with a parent questionnaire does not show any association between parental worklessness and parents helping their child doing homework. This indicates that the association between worklessness and parental homework help – while being overall positive – differs between countries, with no association on average in parent questionnaire countries.

6. Conclusion

Children living in workless households are often surrounded by a particular economic and social disadvantage, potentially putting strains on their educational careers. While workless parents may have more time to spend on their child's education, they might face tighter economic constraints. This potentially changes how parents decide to invest in their child's

Table 12: M4 – Association between parental worklessness and parental mathematics homework help using data from the parent questionnaire.

	Two-parent household		*Single parent household*	
Regression	Estimate	Standard error	Estimate	Standard error
A | BCDE	0.005	0.011	-0.008	0.023
AB | CDE	0.003	0.013	-0.015	0.020
ABC | DE	0.009	0.009	0.009	0.016
ABCD | E	0.003	0.006	0.001	0.009
Ordered logistic regression	0.006	0.010	-0.004	0.020

$* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$

Notes: First four rows – linear probability models with different cut points for the categorical dependent variable, indicated by ' | '. Abbreviations: A: 'Never or hardly ever'; B: 'Once or twice a year'; C: 'Once or twice a month'; D: 'Once or twice a week'; E: 'Every day or almost every day'. Standard errors clustered at country level. Country fixed effects.
Last row – ordered logistic regression. For comparability we report the average marginal effect *and the corresponding standard error, which allows the magnitude of the regression coefficients to be compared. Standard errors clustered at country level. Country dummies included (no country fixed effects).*
Adjusted for gender, mothers' and fathers' occupation level, mothers' and fathers' education level, immigration status, and mothers' and fathers' age.

education by spending their financial and time resources.

I studied these hypotheses using PISA 2012 and applying different statistical methods to find robust estimates for the association between parental worklessness and educational investments.

I find that parental worklessness is associated with lower overall spending on education in both two-parent and single parent households. However, offering their child access to commercial tutoring appears not to be a relevant channel in which non-workless parents spend their additional resources. On the other hand, workless parents – especially single parents – appear to spend more time helping their child with its homework compared to non-workless parents with otherwise similar characteristics. While we find this pattern in many subsets of the data analysed, we do not find higher time investments into workless background children in countries with a parent questionnaire, i.e. Belgium, Chile, Croatia, Germany, Hong Kong, Hungary, Italy, Macau, Mexico, Portugal, and South Korea – regardless of whether students or parents report on this. Overall, our results suggest that children from workless households on average receive lower monetary investments in their education, while in some countries children – especially those from single parent households – receive higher time investments.

However, these results come with limitations. First, while the methods we use go beyond an analysis of correlations – comparing children growing up in workless and non-workless households with otherwise very similar background characteristics – it would not be appropriate to interpret the results as *causal*. Second, despite PISA offering a unique international perspective on the ramifications of worklessness, country level interpretation of our results is limited and often not possible. Third, we are lacking a time dimension in the data to see and analyse different patterns of worklessness. Similarly, we don't have detailed information about potential reasons for worklessness, such as age, illness, wealth, or unemployment. Last, the proportion workless households from the PISA study is a coarse estimate. For the UK, PISA data suggests that around 7% of children grow up in a workless household, compared to around 12% according to UK Department for Work and Pensions (2017) and 17% according to Eurostat (2019).

Next, further research on country level could study the difference in educational investments between workless and non-workless households. This could help address the limitations stated above. Country-specific characteristics of the education and welfare system could be taken into account, potentially allowing for causal claims. Also, from the PISA study parental employment status is only observed at one point in time, not taking into account the complexities of individual employment history, including the reasons for worklessness. Further research could fill that gap by using longitudinal data which includes information on households' employment status across several points in time.

Results from our study as well as potential country level analyses could help guide how policy makers approach the education of children growing up in workless households. For instance, in countries in which monetary investments by workless parents are lower compared to non-workless households, unemployment benefits could help account for that by introducing a voucher system redeemable for educational expenses.

References

Aiyagari, S.Rao, Jeremy Greenwood and Ananth Seshadri (2002). 'Efficient Investment in Children'. In: *Journal of Economic Theory* 102.2, pp. 290–321. ISSN: 00220531. DOI: 10.1006/jeth.2001.2852.

Becker, Gary S. and Nigel Tomes (1986). 'Human Capital and the Rise and Fall of Families'. In: *Journal of Labor Economics* 4 (3, Part 2), S1–S39. ISSN: 0734-306X, 1537-5307. DOI: 10.1086/298118.

Berloffa, Gabriella, Eleonora Matteazzi and Paola Villa (2017). 'The Influence of Parental Employment Status on Children's Labor Outcomes. Does the Gender of Parents and Children Matter?' In: *Journal of Research in Gender Studies* 7.2, pp. 134–164. ISSN: 2164-0262. DOI: 10.22381/JRGS7220178.

Bernal, Raquel (2008). 'The Effect of Maternal Employment and Child Care on Children's Cognitive Development'. In: *International Economic Review* 49.4, pp. 1173–1209. ISSN: 00206598. DOI: 10.1111/j.1468-2354.2008.00510.x.

Black, Sandra E, Paul J Devereux and Kjell G Salvanes (2005). 'Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital'. In: *The American Economic Review* 95.1, p. 13.

Buuren, Stef van and Karin Groothuis-Oudshoorn (2011). 'Mice: Multivariate Imputation by Chained Equations in R'. In: *Journal of Statistical Software* 45.3. ISSN: 1548-7660. DOI: 10.18637/jss.v045.i03.

Caliendo, Marco and Sabine Kopeinig (2008). 'Some Practical Guidance for the Implementation of Propensity Score Matching'. In: *Journal of Economic Surveys* 22.1, pp. 31–72. ISSN: 0950-0804, 1467-6419. DOI: 10.1111/j.1467-6419.2007.00527.x.

Conger, Rand D., Katherine J. Conger, Glen H. Elder, Frederick O. Lorenz, Ronald L. Simons and Les B. Whitbeck (1992). 'A Family Process Model of Economic Hardship and Adjustment of Early Adolescent Boys'. In: *Child Development* 63.3, pp. 526–541. ISSN: 0009-3920. DOI: 10.2307/1131344.

Dang, H.-A. and F. H. Rogers (2008). 'The Growing Phenomenon of Private Tutoring: Does It Deepen Human Capital, Widen Inequalities, or Waste Resources?' In: *The World Bank Research Observer* 23.2, pp. 161–200. ISSN: 0257-3032, 1564-6971. DOI: 10.1093/wbro/lkn004.

Eurostat (2019). *Jobless Households - Children*. Share of persons aged 0-17 who are living in households where no-one works. URL: https://data.europa.eu/euodp/en/data/dataset/SnnopD7PYlHQQ6Tgf2tuaA (visited on 05/12/2019).

Friedman, Sam, Daniel Laurison and Lindsey Macmillan (2017). 'Social Mobility, the Class Pay Gap and Intergenerational Worklessness: New Insights from The Labour Force Survey'. In: p. 56.

Great Britain. Office for National Statistics (2019). 'Workless Households for Regions across the UK'. In: p. 8. URL: https://www.ons.gov.uk/employmentandlabourmarket/peoplenotinwork/unemployment/bulletins/worklesshouseholdsforregionsacrosstheuk/2018#workless-households.

Guryan, Jonathan, Erik Hurst and Melissa Kearney (2008). 'Parental Education and Parental Time with Children'. In: *Journal of Economic Perspectives* 22.3, pp. 23–46. DOI: 10.1257/jep.22.3.23.

Ho, Daniel E., Kosuke Imai, Gary King and Elizabeth A. Stuart (2007). 'Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference'. In: *Political Analysis* 15.03, pp. 199–236. ISSN: 1047-1987, 1476-4989. DOI: 10.1093/pan/mpl013.

– (2011). 'MatchIt: Nonparametric Preprocessing for Parametric Causal Inference'. In: *Journal of Statistical Software* 42.8. ISSN: 1548-7660. DOI: 10.18637/jss.v042.i08.

Jerrim, John and Lindsey Macmillan (2015). 'Income Inequality, Intergenerational Mobility, and the Great Gatsby Curve: Is Education the Key?' In: *Social Forces* 94.2, pp. 505–533. ISSN: 0037-7732, 1534-7605. DOI: 10.1093/sf/sov075.

Jerrim, John and John Micklewright (2014). 'Socio-Economic Gradients in Children's Cognitive Skills: Are Cross-Country Comparisons Robust to Who Reports Family Background?' In: *European Sociological Review* 30.6, pp. 766–781. ISSN: 0266-7215, 1468-2672. DOI: 10.1093/esr/jcu072.

Kalil, Ariel, Rebecca Ryan and Elise Chor (2014). 'Time Investments in Children across Family Structures'. In: *The ANNALS of the American Academy of Political and Social Science* 654.1. Ed. by Marcia J. Carlson and Daniel R. Meyer, pp. 150–168. ISSN: 0002-7162, 1552-3349. DOI: 10.1177/0002716214528276.

Kassotakis, Michael and Athanasios Verdis (2013). 'Shadow Education in Greece'. In: *Private Tutoring Across the Mediterranean*. Ed. by Mark Bray, André E. Mazawi and Ronald G. Sultana. Rotterdam: SensePublishers, pp. 93–113. ISBN: 978-94-6209-237-2. DOI: 10.1007/978-94-6209-237-2_6.

Kornrich, Sabino and Frank Furstenberg (2013). 'Investing in Children: Changes in Parental Spending on Children, 1972–2007'. In: *Demography* 50.1, pp. 1–23. ISSN: 0070-3370, 1533-7790. DOI: 10.1007/s13524-012-0146-4.

Leibowitz, Arleen (1974). 'Home Investments in Children'. In: *Journal of Political Economy* 82 (2, Part 2), S111–S131. ISSN: 0022-3808, 1537-534X. DOI: 10.1086/260295.

Macmillan, Lindsey (2014). 'Intergenerational Worklessness in the UK and the Role of Local Labour Markets'. In: *Oxford Economic Papers* 66.3, pp. 871–889. ISSN: 1464-3812, 0030-7653. DOI: 10.1093/oep/gpt046.

Macmillan, Lindsey, Paul Gregg, John Jerrim and Nikki Shure (2018). 'Children in Jobless Households across Europe: Evidence on the Association with Medium- and Long-Term

Outcomes'. In: *Journal of Poverty and Social Justice*. ISSN: 1759-8273. DOI: 10.1332/175982718X15265611809224.

Mäder, Miriam, Regina T Riphahn, Caroline Schwientek and Steffen Müller (2015). 'Intergenerational Transmission of Unemployment – Evidence for German Sons'. In: *Journal of Economics and Statistics* 235 (4+5), p. 21.

Mauldin, Teresa, Yoko Mimura and Mark Lino (2001). 'Parental Expenditures on Children's Education'. In: *Journal of Family and Economic Issues* 22.3, p. 21.

McClelland, Alison (2000). 'Effects of Unemployment on the Family'. In: *The Economic and Labour Relations Review* 11.2, pp. 198–2012. DOI: https://doi.org/10.1177/103530460001100204.

Miller, David C (2007). 'Comparative Indicators of Education in the United States and Other G-8 Countries: 2006'. In: p. 22.

Mynarska, Monika, Bernhard Riederer, Ina Jaschinski, Desiree Krivanek, Gerda Neyer and Livia Oláh (2015). 'Vulnerability of Families with Children: Major Risks, Future Challenges and Policy Recommendations'. In: *"Families and Societies" Project, Working Paper* 49, p. 107.

Öchsner, Thomas (2018). '1,06 Euro für Bildung'. In: *sueddeutsche.de. politik*. ISSN: 0174-4917. URL: https://www.sueddeutsche.de/politik/armut-1-06-euro-fuer-bildung-1.3906460 (visited on 30/05/2019).

OECD (Organization for Economic Cooperation and Development) (2014). *PISA 2012 Technical Report*, p. 472.

Parsons, Samantha, Ingrid Schoon and Anna Vignoles (2014). 'Parental Worklessness and Children's Early School Achievement and Progress'. In: *Longitudinal and Life Course Studies* 5.1, pp. 19–41. ISSN: 17579597. DOI: 10.14301/llcs.v5i1.230.

Rosenbaum, Paul R. and Donald B. Rubin (1984). 'Reducing Bias in Observational Studies Using Subclassification on the Propensity Score'. In: *Journal of the American Statistical Association*, p. 9.

'Multiple Imputation for Nonresponse in Surveys' (1987). In: *Wiley Series in Probability and Statistics*. Ed. by Donald B. Rubin. Hoboken, NJ, USA: John Wiley & Sons, Inc., pp. i–xxix. ISBN: 978-0-470-31669-6 978-0-471-08705-2. DOI: 10.1002/9780470316696.fmatter.

Rutkowski, David, Leslie Rutkowski, Justin Wild and Nathan Burroughs (2018). 'Poverty and Educational Achievement in the US: A Less-Biased Estimate Using PISA 2012 Data'. In: *Journal of Children and Poverty* 24.1, pp. 47–67. ISSN: 1079-6126, 1469-9389. DOI: 10.1080/10796126.2017.1401898.

Schäfer, Christoph (2019). 'Hartz-IV-Kinder haben 54 Cent für Bildung'. In: *FAZ.NET*. ISSN: 0174-4909. URL: `https://www.faz.net/aktuell/wirtschaft/arm-und-reich/ueber-hartz-iv-und-warum-kinder-nur-54-cent-fuer-bildung-erhalten-16213781.html` (visited on 19/08/2019).

Schoon, Ingrid (2014). 'Parental Worklessness and the Experience of NEET among Their Offspring. Evidence from the Longitudinal Study of Young People in England (LSYPE)'. In: *Longitudinal and Life Course Studies* 5.2. ISSN: 17579597. DOI: `10.14301/llcs.v5i2.279`.

Solon, Gary (2004). 'A Model of Intergenerational Mobility Variation over Time and Place'. In: *Generational Income Mobility in North America and Europe*. Ed. by Miles Corak. Cambridge: Cambridge University Press, pp. 38–47. ISBN: 978-0-511-49254-9 978-0-521-82760-7. DOI: `10.1017/CBO9780511492549.003`.

Stuart, Elizabeth A. (2010). 'Matching Methods for Causal Inference: A Review and a Look Forward'. In: *Statistical Science* 25.1, pp. 1–21. ISSN: 0883-4237. DOI: `10.1214/09-STS313`.

UK Department for Work and Pensions (2017). *Improving Lives: Helping Workless Families.* OCLC: 1018037637. ISBN: 978-1-78425-923-5.

UNESCO Institute for Statistics (UIS) (2020). *Education: Government Expenditure on Education as a Percentage of GDP*. http://data.uis.unesco.org/index.aspx?queryid=3373. URL: `http://data.uis.unesco.org/Index.aspx?DataSetCode=EDULIT_DS#` (visited on 10/01/2020).

World Bank (2020). *GDP per Capita, PPP (Current International $)*. URL: `https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD` (visited on 13/01/2020).

– (2019a). *Labor Force Participation Rate, Female (% of Female Population Ages 15+) (Modeled ILO Estimate)*. URL: `https://data.worldbank.org/indicator/SL.TLF.CACT.FE.ZS` (visited on 30/12/2019).

– (2019b). *Labor Force Participation Rate, Male (% of Female Population Ages 15+) (Modeled ILO Estimate)*. URL: `https://data.worldbank.org/indicator/SL.TLF.CACT.MA.ZS` (visited on 30/12/2019).

Yeung, W. Jean, Miriam R. Linver and Jeanne Brooks–Gunn (2002). 'How Money Matters for Young Children's Development: Parental Investment and Family Processes'. In: *Child Development* 73.6, pp. 1861–1879. ISSN: 0009-3920, 1467-8624. DOI: `10.1111/1467-8624.t01-1-00511`.

Appendices

A. Additional tables

Table A.1: Overview of all countries and jurisdictions participating in the PISA 2012 cycle.

Country	N_0	P^{workless}	N_{SQ}	N_{SQ}^{workless}	N_{SQ}^{matched}	N_{PQ}	N_{PQ}^{workless}	N_{PQ}^{matched}
Albania	4,743	0.12	2,394	386	768			
Argentina	5,908	0.09	3,040	344	604			
Australia	14,481	0.06	8,558	597	1,180			
Austria	4,755	0.03	2,871	97	188			
Belgium	8,597	0.06	5,119	318	628	7,777	518	1,022
Brazil	19,204	0.13	9,117	1,657	3,314			
Bulgaria	5,282	0.05	2,844	173	338			
Canada	21,544	0.04	12,860	548	1,084			
Chile	6,856	0.08	3,964	369	720	5,934	560	1,096
Colombia	9,073	0.10	4,115	611	1,220			
Costa Rica	4,602	0.12	2,403	366	580			
Croatia	5,008	0.15	3,171	527	984	4,746	750	1,488
Czechia	5,327	0.03	3,335	105	206			
Denmark	7,481	0.09	4,488	474	938			
Estonia	4,779	0.04	2,844	119	228			
Finland	8,829	0.05	5,262	315	600			
France	4,613	0.05	2,729	148	294			
Germany	5,001	0.03	2,500	92	174	3,857	149	282
Greece	5,125	0.11	3,147	373	726			
Hong Kong SAR China	4,670	0.07	2,923	220	430	4,362	333	646
Hungary	4,810	0.07	2,930	242	462	4,406	357	714
Iceland	3,508	0.03	2,117	76	134			
Indonesia	5,622	0.11	2,373	409	742			
Ireland	5,016	0.10	3,036	323	634			
Israel	5,055	excluded						
Italy	31,073	0.05	19,402	958	1,894	29,185	1,460	2,912
Japan	6,351	0.02	3,969	66	132			
Jordan	7,038	0.18	3,361	860	1,503			
Kazakhstan	5,808	0.12	3,441	459	854			
Latvia	4,306	0.04	2,385	121	222			
Liechtenstein	293	excluded						
Lithuania	4,618	0.08	2,736	251	494			
Luxembourg	5,258	0.04	3,170	152	294			
Macau SAR China	5,335	0.05	3,287	151	276	4,928	242	470
Malaysia	5,197	0.12	2,913	427	679			
Mexico	33,806	0.11	17,626	2,514	4,386	26,430	3,743	6,898
Montenegro	4,744	0.16	2,769	480	926			
Netherlands	4,460	0.05	2,674	121	220			
New Zealand	4,291	0.07	2,640	194	362			
Norway	4,686	0.03	2,845	87	160			
Perm Russia	1,761	0.04	1,014	44	88			
Peru	6,035	0.13	3,258	562	1,114			

Country	N_0	P^{workless}	N_{SQ}	N_{SQ}^{workless}	N_{SQ}^{matched}	N_{PQ}	N_{PQ}^{workless}	N_{PQ}^{matched}
Poland	4,607	0.08	2,837	219	432			
Portugal	5,722	0.06	3,342	223	410	5,022	341	672
Qatar	10,966	0.09	5,193	622	1,070			
Romania	5,074	0.15	2,849	529	1,040			
Russia	5,231	0.06	2,950	201	374			
Serbia	4,684	0.12	2,750	381	758			
Shanghai China	5,177	0.08	3,180	275	544			
Singapore	5,546	0.04	3,381	146	272			
Slovakia	4,678	0.06	2,760	205	394			
Slovenia	5,911	0.05	3,588	168	332			
South Korea	5,033	0.05	3,012	172	324	4,528	270	520
Spain	25,313	0.07	15,490	1,084	2,138			
Sweden	4,736	0.03	2,801	97	188			
Switzerland	11,229	0.03	6,888	250	496			
Taipei China	6,046	0.06	3,755	221	438			
Thailand	6,606	0.10	3,098	428	856			
Tunisia	4,407	0.12	2,155	330	566			
Turkey	4,848	0.21	2,598	690	1,291			
United Arab Emirates	11,500	0.15	6,095	1,135	1,900			
United Kingdom	12,659	0.07	7,408	561	1,054			
United States	4,978	0.05	2,862	195	370			
Uruguay	5,315	0.06	2,746	211	402			
Vietnam	4,959	0.37	2,807	1,227	2,352			

Notes: N_0: total number of participating students in PISA 2012. $NA^{workless}$: proportion workless observations. N_{SQ}: number of observations suitable for analysis of dependent variables from the student questionnaire. $N_{SQ}^{workless}$: number of workless *observations for analysis of student questionnaire. $N_{SQ}^{matched}$: number of observations after propensity score matching. N_{PQ}: number of suitable observations from parent questionnaire in PISA 2012. $N_{PQ}^{workless}$: number of suitable workless observations from parent questionnaire. $N_{SQ}^{matched}$: number of observations after propensity score matching of parent questionnaire data.*

B. Details about methods

B.1. Education expense variable from parent questionnaire

This section describes how the categorical data from the monetary expense variable obtained from the parent questionnaire is transformed to ensure better comparability across countries. In table B.1 the original education expense categories as reported in the PISA 2012 data are shown. The actual monetary values defining the categories are created for each country separately. Even though the PISA technical report specifies how those categories should be constructed, it remains rather non-transparent in practice. In the course of this paper, relative

Table A.2: Monetary and time investments for all countries with a parent questionnaire.

	Monetary investments		*Time investments*	
Data	Estimate	Standard error	Estimate	Standard error
Belgium	0.008	0.035	0.043	0.039
Chile	-0.065**	0.028	-0.045	0.035
Croatia	-0.026	0.025	-0.011	0.028
Germany	-0.134*	0.078	0.108*	0.061
Hong Kong SAR China	-0.071*	0.040	0.013	0.041
Hungary	-0.078**	0.040	0.026	0.032
Italy	-0.027	0.017	0.017	0.022
Macau SAR China	0.013	0.048	-0.003	0.050
Mexico	-0.028**	0.013	-0.011	0.012
Portugal	-0.125***	0.045	0.082*	0.042
South Korea	-0.074*	0.044	-0.017	0.044

$* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$

Notes: Results from an ordered logistic regression pooling single and two-parent households.

expenses within each country are used. For this, each category should include a comparable amount of observations across countries. As table B.1 highlights, the amount of observations in each category varies tremendously across countries. For instance, the lowest expense category contains around 1.6% of observations in Belgium and about 20.6% in Germany. In countries such as Hungary and Mexico the highest three categories contain only very few observations, whereas in Italy the highest category is the second largest.

For better comparability those six categories are regrouped to become three new categories aiming at balancing category sizes across countries. Table B.2 shows the sizes of those new categories. Even though balancing is not perfect, the categories help ensure better comparability across countries. Note that categories are merged for each country separately as each country has specific characteristics that need to be taken into account.

C. Robustness Checks

C.1. Heterogeneity Analysis

We obtained the results shown earlier in this section by pooling a large set of countries which are different in many aspects such as education and welfare system, overall economic power as well as the prevalence of commercial tutoring and parental homework help. In this section

Table B.1: Education expenses for all countries with parent questionnaire.

	0	0–W	W–X	X–Y	Y–Z	>Z
Belgium	0.016	0.057	0.234	0.321	0.167	0.205
Chile	0.113	0.283	0.433	0.084	0.024	0.063
Croatia	0.201	0.151	0.255	0.148	0.086	0.159
Germany	0.206	0.122	0.295	0.152	0.133	0.092
Hong Kong SAR China	0.101	0.281	0.152	0.361	0.077	0.027
Hungary	0.212	0.325	0.293	0.094	0.035	0.042
Italy	0.041	0.174	0.270	0.155	0.101	0.259
Macau SAR China	0.063	0.471	0.221	0.125	0.054	0.065
Mexico	0.148	0.423	0.380	0.024	0.010	0.015
Portugal	0.142	0.228	0.571	0.047	0.010	0.003
South Korea	0.067	0.321	0.267	0.170	0.102	0.073

Notes: Data from the parent questionnaire of PISA 2012. Displayed are the six spending categories as in the PISA dataset without recoding. The absolute values for W, X, Y, and Z differ in each country and are not reported in the PISA technical report.

we show our estimates when pooling over different subsets of countries.

First, parents might adjust their behaviour according to the education system they raise their children in: the need for parental money and time to be invested in children's education could be lower in countries with a highly prioritised education system. We use data on public spending on education as proportion of GDP (UNESCO Institute for Statistics (UIS) 2020) and split the PISA countries at their median spending into low and high spending countries.

Second, there might be a difference between wealthy and less wealthy countries. Therefore, we split up the dataset into high and low GDP per capita (purchasing power equivalents) countries according to World Bank (2020).

Third, the prevalence of commercial tutoring and parental homework help differs substantially between countries. This might be reflected in the magnitude of our estimates. To account for this potential difference, we analyse countries separately depending on the proportion of commercial and parental tutoring, respectively, again splitting up the dataset into below and above median.

Table C.1 shows the estimates for the association between parental worklessness and commercial tutoring. We find no significant results for any of the subsets. The point estimates differ depending on which subset is analysed. Point estimates are higher in countries with more common commercial tutoring. Children growing up in single may have a stronger disadvantage in countries with low public spending on education and countries with low overall

Table B.2: Education expenses for all countries with parent questionnaire. Three new categories created from the original six categories with increased balance are shown.

	low	medium	high
Belgium	0.307	0.321	0.372
Chile	0.396	0.433	0.171
Croatia	0.352	0.403	0.245
Germany	0.328	0.448	0.225
Hong Kong SAR China	0.383	0.514	0.104
Hungary	0.212	0.618	0.170
Italy	0.216	0.425	0.360
Macau SAR China	0.535	0.221	0.244
Mexico	0.571		0.429
Portugal	0.370	0.571	0.060
South Korea	0.387	0.267	0.345

Notes: Recoded educational expense categories based on data from the PISA 2012 parent questionnaire. No medium *category built for Mexico.*

GDP. However, these differences should be taken with pinch of salt as standard errors for all point estimates are large leaving wide confidence intervals for all estimates.

When focussing on the association between worklessness and parental homework help, the heterogeneity analysis in inconclusive with estimates changing in the opposite directions for two-parent and single parent households. As before, all differences in estimates are possibly random effects and are well within the respective confidence intervals.

Next, we present country specific ordered logistic regression estimates for the eleven countries with a parent questionnaire. Recall that overall we find a statistically significant association between workless parents and educational expenses and no association for parental mathematics homework help. We compute the country specific estimates not differentiating between two-parent and single-parent households to maximise sample sizes for each regression. In nine out of eleven countries, we find a negative association between worklessness and monetary investments. In Germany, Hong Kong, and South Korea these estimates are statistically significant on the 10% level; in Chile, Hungary, and Mexico estimates are statistically significant on the 5% level; and in Portugal we find an association statistically significant on the 1% level. For mathematics homework help, we find a positive association in six out of eleven countries, with only Germany and Portugal having estimates statistically significant at the 10% level.

Table C.1: M1 – Heterogeneity analysis for the association between parental worklessness and commercial tutoring.

	Two-parent household		*Single parent household*	
Data	Estimate	Standard error	Estimate	Standard error
Public spending on education as % of GDP				
High	-0.013	0.012	-0.003	0.010
Low	-0.012	0.011	-0.017	0.012
GDP per capita (PPP)				
High	-0.008	0.015	-0.003	0.011
Low	-0.010	0.009	-0.010	0.011
Prevalence of commercial tutoring				
High	-0.014	0.012	-0.018	0.013
Low	-0.002	0.008	0.003	0.009

$* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$

Notes: Linear probability model on matched sample run separately for two-parent and single parent households. Adjusted for gender, mothers' and fathers' occupation level, mothers' and fathers' education level, and immigration status. Standard errors clustered at country level. Country fixed effects.

Table C.2: M2 – Heterogeneity analysis for the association between worklessness and parental homework help.

	Two-parent household		*Single parent household*	
Data	Estimate	Standard error	Estimate	Standard error
Public spending on education as % of GDP				
High	0.015*	0.009	0.023*	0.013
Low	0.009	0.011	0.028*	0.015
GDP per capita (PPP)				
High	0.021*	0.011	0.022	0.014
Low	0.008	0.008	0.031**	0.014
Prevalence of parental homework help				
High	0.009	0.008	0.029**	0.014
Low	0.018	0.012	0.025*	0.013

$* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$

Notes: Linear probability model on matched sample run separately for two-parent and single parent households. Adjusted for gender, mothers' and fathers' occupation level, mothers' and fathers' education level, and immigration status. Standard errors clustered at country level. Country fixed effects.

Table C.3: M3 & M4 – Country specific estimates.

	Educational expenses		*Mathematics homework help*	
Country	Estimate	Standard error	Estimate	Standard error
Belgium	0.008	0.035	0.043	0.039
Chile	-0.065**	0.028	-0.045	0.035
Croatia	-0.026	0.025	-0.011	0.028
Germany	-0.134*	0.078	0.108*	0.061
Hong Kong SAR China	-0.071*	0.040	0.013	0.041
Hungary	-0.078**	0.040	0.026	0.032
Italy	-0.027	0.017	0.017	0.022
Macau SAR China	0.013	0.048	-0.003	0.050
Mexico	-0.028**	0.013	-0.011	0.012
Portugal	-0.125***	0.045	0.082*	0.042
South Korea	-0.074*	0.044	-0.017	0.044

$* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$

Notes: Estimates computed by using an ordered logistic regression with two-parent and single parent households pooled together for larger sample sizes. Average marginal effects presented.

C.2. Continuous dependent variable

The dependent variables from the student questionnaire, commercial tutoring and parental homework help, are both continuous variables with values between 0 and 30. For the purpose of the main analyses, we recode them to become a binary variable indicating whether or not a student receives any commercial tutoring or parental homework help on a weekly basis. Tables C.4 and C.5 show the results when applying a linear model to the original continuous variable. Overall, we find similar results: no association between worklessness and commercial tutoring and a positive association between worklessness and parental homework help. However, standard errors are larger and some of the initially significant estimates for OECD and partner countries, respectively – especially for single parent households – are now statistically not significant. This could partly be explained by a larger influence of unrealistic outliers (30 hours per week of parental homework help).

C.3. No matching

All results presented in the main body of this paper preprocess the PISA data using matching techniques to improve balance in many background characteristics between workless and non-workless households. In the following, we present estimates resulting from regression analyses without prior matching. As before, we find no association between worklessness

Table C.4: Association between worklessness and commercial tutoring from a linear model.

	Two-parent household		*Single parent household*	
Data	Estimate	Standard error	Estimate	Standard error
All countries	-0.054	0.038	0.008	0.042
OECD	-0.028	0.046	0.022	0.037
Partner countries	-0.071	0.056	-0.003	0.081
PQ	0.001	0.051	-0.031	0.056

$*\ p < 0.1$, $**\ p < 0.05$, $***\ p < 0.01$

Notes: Regression on matched sample run separately for two-parent and single parent households. Adjusted for gender, mothers' and fathers' occupation level, mothers' and fathers' education level, and immigration status. Standard errors clustered at country level. Country fixed effects.

Table C.5: Association between worklessness and parental homework help from a linear model.

	Two-parent household		*Single parent household*	
Data	Estimate	Standard error	Estimate	Standard error
All countries	0.062*	0.037	0.089**	0.045
OECD	0.101**	0.048	0.065	0.058
Partner countries	0.033	0.055	0.120	0.073
PQ	0.070	0.069	0.101	0.087

$*\ p < 0.1$, $**\ p < 0.05$, $***\ p < 0.01$

Notes: Regression on matched sample run separately for two-parent and single parent households. Adjusted for gender, mothers' and fathers' occupation level, mothers' and fathers' education level, and immigration status. Standard errors clustered at country level. Country fixed effects.

and commercial tutoring and a significant association between worklessness and educational expenses. We find a notable difference to the results presented in the main body of this paper in the association between worklessness and homework help in two-parent households: before, we found no association in countries with a parent questionnaire, whereas in the unmatched sample we find an association in data from the student questionnaire, and – to a lesser extent – in data from the parent questionnaire.

C.3.1. Monetary investments

Table C.6: Association between worklessness and commercial tutoring from a linear probability model applied to different subsets of the matched PISA data – no matching.

	Two-parent household		*Single parent household*	
Data	Estimate	Standard error	Estimate	Standard error
All countries	0.008	0.010	0.003	0.006
OECD	-0.001	0.012	0.006	0.007
Partner countries	0.011	0.012	-0.002	0.011
PQ	0.016**	0.007	0.002	0.008

$*\ p < 0.1$, $**\ p < 0.05$, $***\ p < 0.01$

Notes: Regression on unmatched sample run separately for two-parent and single parent households. Adjusted for gender, mothers' and fathers' occupation level, mothers' and fathers' education level, and immigration status. Standard errors clustered at country level. Country fixed effects.

C.3.2. Time investments

Table C.7: Association between parental worklessness and monetary investments using data from the parent questionnaire – no matching.

	Two-parent household		*Single parent household*	
Regression	Estimate	Standard error	Estimate	Standard error
low | medium, high	-0.050***	0.015	-0.043***	0.014
low, medium | high	-0.030*	0.018	-0.025**	0.010
Ordered logistic regression	-0.047***	0.017	-0.038***	0.010

$*\ p < 0.1$, $**\ p < 0.05$, $***\ p < 0.01$

Notes: First two rows – linear probability models with different cut points for the categorical dependent variable: between low income and merged medium and high income (first row) and between merged low and medium income and high income (second row). Standard errors clustered at country level. Country fixed effects.
Third row – ordered logistic regression. For comparability we report the average marginal effect *and the corresponding standard error, which allows the magnitude of the regression coefficients to be compared. Standard errors clustered at country level. Country dummies included (no country fixed effects).*
Adjusted for gender, mothers' and fathers' occupation level, mothers' and fathers' education level, immigration status, and mothers' and fathers' age.

Table C.8: Association between worklessness and parental homework help from a linear probability model applied to different subsets of the matched PISA data – no matching.

	Two-parent household		*Single parent household*	
Data	Estimate	Standard error	Estimate	Standard error
All countries	0.035***	0.006	0.031***	0.008
OECD	0.038***	0.008	0.022***	0.009
Partner countries	0.023***	0.007	0.040***	0.014
PQ	0.019***	0.007	0.017	0.015

$*\ p < 0.1$, $**\ p < 0.05$, $***\ p < 0.01$

Notes: Regression on unmatched sample run separately for two-parent and single parent households. Adjusted for gender, mothers' and fathers' occupation level, mothers' and fathers' education level, and immigration status. Standard errors clustered at country level. Country fixed effects.

Table C.9: Association between parental worklessness and parental mathematics homework help using data from the parent questionnaire – no matching.

	Two-parent household		*Single parent household*	
Regression	Estimate	Standard error	Estimate	Standard error
A | BCDE	0.017*	0.008	0.001	0.020
AB | CDE	0.018	0.013	-0.003	0.015
ABC | DE	0.024*	0.010	0.015	0.012
ABCD | E	0.010	0.007	0.002	0.008
Ordered logistic regression	0.022*	0.011	0.005	0.017

$* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$

Notes: First four rows – linear probability models with different cut points for the categorical dependent variable, indicated by ' | '. Abbreviations: A: 'Never or hardly ever'; B: 'Once or twice a year'; C: 'Once or twice a month'; D: 'Once or twice a week'; E: 'Every day or almost every day'. Standard errors clustered at country level. Country fixed effects.
Last row – ordered logistic regression. For comparability we report the average marginal effect *and the corresponding standard error, which allows the magnitude of the regression coefficients to be compared. Standard errors clustered at country level. Country dummies included (no country fixed effects).*
Adjusted for gender, mothers' and fathers' occupation level, mothers' and fathers' education level, immigration status, and mothers' and fathers' age.

C.4. Coarsened exact matching

Coarsened exact matching (CEM) requires exact matches on all variables included in the algorithm. While exact matching is impractical for continuous variables, CEM allows to set up categories, within which exact matches are required. We use this as a robustness check. As exact matches are increasingly difficult to find the more variables are included in the algorithm, we perform matching using only parental occupation and education levels (four variables). For countries with high proportions of missing data of mothers' occupation, we additionally require exact matching on a missing data dummy for mothers' occupation. We divide parents' occupation variables into 10 equally sized categories to make exact matching possible. As with propensity score matching, we run the algorithm separately for single and two-parent households.

Since CEM is a very restrictive matching approach as all categories must be matched exactly, more observations remain unmatched: more than 3,000 out of 9,859 children living in single parent households remain unmatched. This results in a sample less representative of workless background children than for propensity score matching in which the amount of unmatched observations remains comparably low. On the other hand, more non-workless background children are included in the matched dataset when matching with CEM, making estimates more robust.

Tables C.10 and C.11 show balance improvements in all variables of interest when applying CEM to student and parent questionnaire data, respectively. As guaranteed by CEM, parental education levels are now exactly matched. Occupation levels improve really well as they are required to fall into a very narrow range. However, as no matching is performed on gender and immigration status, balance in these variables does not improve. For the parent questionnaire data, balance is strongly reduced for most variables indicating parental age. Furthermore, the amount of workless background children in the final matched dataset is lower than under propensity score matching. However, as more non-workless background children are included as a control (making the application of weights necessary for all analyses), estimates can be expected to be more precise.

Overall, CEM comes with the advantage of having close to perfect matches in variables chosen for matching. This comes with a simple trade off: the more variables are included for CEM, the fewer workless background children can be matched to similar enough non-workless background peers. Hence, we choose to match only on the most important parental

Table C.10: Absolute standardised bias in means before and after matching – student questionnaire.

	Two-parent household			*Single parent household*		
Variable	Before	After	Improvement (in percent)	Before	After	Improvement (in percent)
Gender	0.06	0.03	54.51	0.03	0.06	-68.38
Immigration status	0.01	0.03	-175.29	0.12	0.09	21.34
Occupation level father	0.59	0.01	98.18	0.16	0.00	98.12
Occupation level mother	0.56	0.00	99.87	0.35	0.00	99.49
Education level father - low	0.51	0.00	100.00	0.23	0.00	100.00
Education level father - medium	0.10	0.00	100.00	0.06	0.00	100.00
Education level father - high	0.53	0.00	100.00	0.18	0.00	100.00
Education level mother - low	0.60	0.00	100.00	0.38	0.00	100.00
Education level mother - medium	0.13	0.00	100.00	0.01	0.00	100.00
Education level mother - high	0.63	0.00	100.00	0.41	0.00	100.00

Notes: Matched dataset generated as described in this section; i.e. coarsened exact matching. All numbers are averaged over all 30 imputations.
Two-parent household*: Workless background children discarded for lack of common support: 0. Unmatched workless background children: 743-866. Total number of observations in matched dataset: 121,525-123,870.*
Single parent household*: Workless background children discarded for lack of common support: 0-38. Workless background children unmatched: 3056-3195. Total number of observations in matched dataset: 19,119-19,649. Variation in figures due to random differences between the 30 imputed datasets.*

Table C.11: Absolute standardised bias in means before and after matching – parent questionnaire.

	Two-parent household			*Single parent household*		
Variable	Before	After	Improvement (in percent)	Before	After	Improvement (in percent)
Gender	0.08	0.06	25.21	0.05	0.07	-34.33
Immigration status	0.02	0.06	-269.32	0.01	0.04	-152.81
Occupation level father	0.68	0.02	97.69	0.26	0.00	98.45
Occupation level mother	0.69	0.00	99.48	0.48	0.01	98.81
Education level father - low	0.52	0.00	100.00	0.28	0.00	100.00
Education level father - medium	0.27	0.00	100.00	0.10	0.00	100.00
Education level father - high	0.40	0.00	100.00	0.23	0.00	100.00
Education level mother - low	0.60	0.00	100.00	0.44	0.00	100.00
Education level mother - medium	0.34	0.00	100.00	0.20	0.00	100.00
Education level mother - high	0.46	0.00	100.00	0.34	0.00	100.00
Father Age $<$36	0.02	0.10	-445.66	0.05	0.05	-18.46
Father Age 36-40	0.04	0.14	-308.55	0.01	0.08	-2072.11
Father Age 41-45	0.13	0.15	-19.79	0.12	0.09	22.71
Father Age 46-50	0.26	0.11	58.28	0.11	0.04	68.62
Father Age $>$51	0.30	0.37	-21.86	0.18	0.21	-12.00
Mother Age $<$36	0.10	0.08	20.82	0.04	0.10	-161.36
Mother Age 36-40	0.03	0.13	-328.42	0.01	0.06	-1933.03
Mother Age 41-45	0.20	0.11	48.41	0.15	0.05	62.91
Mother Age 46-50	0.11	0.05	51.66	0.06	0.01	74.02
Mother Age $>$51	0.23	0.27	-14.25	0.19	0.20	-1.78

Notes: Matched dataset generated as described in this section; i.e. coarsened exact matching. All numbers are averaged over all 30 imputations.
Two-parent household*: Workless background children discarded for lack of common support: 0. Unmatched workless background children: 57-80. Total number of observations in matched dataset: 61,943-63,303.*
Single parent household*: Workless background children discarded for lack of common support: 0-7. Workless background children unmatched: 414-485. Total number of observations in matched dataset: 9,302-9,598. Variation in figures due to random differences between the 30 imputed datasets.*

background variables, i.e. mothers' and fathers' education and occupation levels.

C.4.1. Monetary investments

The results from a CEM-matched sample show – as for the results from the main body of this paper – that no significant association between worklessness and commercial tutoring can be found.

Table C.12: Association between worklessness and commercial tutoring from a linear probability model applied to different subsets of the matched PISA data – coarsened exact matching.

	Two-parent household		*Single parent household*	
Data	Estimate	Standard error	Estimate	Standard error
All countries	-0.001	0.008	-0.005	0.008
OECD	-0.003	0.010	0.002	0.010
Partner countries	-0.005	0.010	-0.014	0.014
PQ	0.008	0.007	-0.011	0.013

$* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$

Notes: Regression on CEM-matched sample run separately for two-parent and single parent households. Adjusted for gender, mothers' and fathers' occupation level, mothers' and fathers' education level, and immigration status. Standard errors clustered at country level. Country fixed effects.

Table C.13: Association between parental worklessness and monetary investments using data from the parent questionnaire – CEM.

	Two-parent household		*Single parent household*	
Regression	Estimate	Standard error	Estimate	Standard error
low | medium, high	-0.049***	0.013	-0.046***	0.015
low, medium | high	-0.031**	0.014	-0.024**	0.010
Ordered logistic regression	-0.030***	0.008	-0.036**	0.015

$* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$

Notes: Linear probability models with different cut points for the categorical dependent variable: between low income and merged medium and high income (first row) and between merged low and medium income and high income (second row). For comparability we report the average marginal effect *and the corresponding standard error for the ordered logistic regression, which allows the magnitude of the regression coefficients to be compared. Standard errors clustered at country level. Country fixed effects (linear probability model) and country dummies included (ordered logistic regression).*
Adjusted for gender, mothers' and fathers' occupation level, mothers' and fathers' education level, immigration status, and mothers' and fathers' age.

C.4.2. Time investments

As in the main body of this paper, we find a significant association between worklessness and parental homework help.

Table C.14: Association between worklessness and parental homework help from a linear probability model applied to different subsets of the matched PISA data – coarsened exact matching.

	Two-parent household		*Single parent household*	
Data	Estimate	Standard error	Estimate	Standard error
All countries	0.016**	0.006	0.024**	0.012
OECD	0.029***	0.009	0.012	0.013
Partner countries	0.004	0.008	0.040**	0.018
PQ	0.009	0.007	0.008	0.022

$* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$

Notes: Regression on CEM-matched sample run separately for two-parent and single parent households. Adjusted for gender, mothers' and fathers' occupation level, mothers' and fathers' education level, and immigration status. Standard errors clustered at country level. Country fixed effects.

Table C.15: Association between parental worklessness and parental mathematics homework help using data from the parent questionnaire – CEM.

	Two-parent household		*Single parent household*	
Regression	Estimate	Standard error	Estimate	Standard error
A | BCDE	0.009	0.009	-0.006	0.024
AB | CDE	0.009	0.014	-0.012	0.020
ABC | DE	0.016	0.010	0.006	0.016
ABCD | E	0.006	0.006	0.003	0.008
Ordered logistic regression	0.012	0.011	-0.004	0.021

$* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$

Notes: Linear probability models with different cut points for the categorical dependent variable, indicated by ' | '. Abbreviations: A: 'Never or hardly ever'; B: 'Once or twice a year'; C: 'Once or twice a month'; D: 'Once or twice a week'; E: 'Every day or almost every day'. For comparability we report the average marginal effect *and the corresponding standard error for the ordered logistic regression, which allows the magnitude of the regression coefficients to be compared. Standard errors clustered at country level. Country fixed effects (linear probability model) and country dummies (ordered logistic regression).*
Adjusted for gender, mothers' and fathers' occupation level, mothers' and fathers' education level, immigration status, and mothers' and fathers' age.

@cepeo_ucl

ucl.ac.uk/ioe/cepeo