Essays on child development in the UK

Richa Saun

A thesis presented for the degree of
Doctor of Philosophy

Department of Economics
School of Politics, Economics and International Relations

University of Reading
12 January 2024

The Essays on child development in the UK

Richa Saun

Abstract

This thesis presents three essays, which address the child development from early age to youth in the UK. The first essay addresses the diverging and converging impact of bilingual household environment on the child's vocabulary and behavioural development across ages 3 to 7 years compared to the monolingual counterparts. The second essay considers the importance of parental mobility descisions taken at age 5 on the vocabulary and behavioural development scores across ages 7 to 14 years and highlights the impact of moving to more or less deprived area of current residence on the child outcomes in terms of change in the vocabulary and behavioural scores. The final essay explains the impact of 'COVID' shock on the youth preferences to continue education. To capture the impact of parental and household inputs on the youth's likelihood to continue education, the final essay utilises the change in parental work status and socioeconomic condition during the pandemic.

The Impact of Individual Parental Bilingualism on Children's Vocabulary and Behavioural Development in the UK

The current chapter examines the impact of individual parental bilingualism on children's vocabulary and behavioural development in the UK. It uses the Millennium Cohort Study (MCS) for children aged 3 to 7 years to look at the age-specific and cumulative effects of individual parental bilingualism. The study finds that the bilingual child fares worse in vocabulary development at age 3 but catches up with their monolingual counterparts by age 7. Mother being bilingual has a significant negative impact on the vocabulary scores for children aged 3 to 5 years. There is a significant negative impact of bilingual father on a child's behavioural development at age 5. The negative impact on behavioural outcomes continues till age 7, in case of mother bilingual. As a cumulative process, mother-only bilingual declines vocabulary development and father-only bilingual have a significant adverse effect on a child's behavioural development compared to their English-only counterparts.

The Impact of Parental Residential Mobility Decisions on Child Development

This chapter examines the age-specific impact of changing residence on a child's vocabulary and behavioural development. It uses the MCS dataset for children aged

3 to 14 years in England. The study aims to analyse the impact of parental decisions to change residence captured by the changes in the movement between Indexes of Multiple Deprivation (IMD). IMD deciles are used to account for the changes in child development for the families moving to higher or lower IMD. The child development at each age group is captured as the change between the test scores obtained at ages 3 and over. The current study uses household movements with combined and individual deciles of the IMDs. The overall age-specific analysis shows that there lies a significant negative impact of the changing residence at age 5 on the vocabulary outcome. In case of internalsing scores, the movements along the deciles of the overall IMD movements, show a positive impact on the internalising scores as one moves down the lower decile of the IMD. The positive impact on internalising scores highlights an increase in the behavioural issues. Among the movements along the deciles of the individual IMDs, households moving to higher or lower decile of the health IMD show a significant change in vocabulary and behavioural scores.

The Impact of COVID on the youth's preferences to continue education

This study investigates the interrelationship between parental employment and socioeconomic dynamics during the COVID-19 pandemic and the subsequent educational aspirations of the youth. Using the pre-COVID and COVID UK MCS data, the study finds approximately 4 percentage point increase in the likelihood of youth preferring to pursue education when fathers transition to unemployment, contrasted with those with employed fathers. Conversely, when mothers transition to unemployment during the pandemic, there was a significant 3 percentage point decrease in the propensity of youth towards further education. These trends prevail across "single and two-parent households" for mother and father and for father only in case of "two-parent only households" analysis. Notably, there was an overall positive inclination to continue education in the case that fathers transitioned jobs for both household types, and a negative inclination in the case of mothers compared to parents remain in unemployment. Specific economic shifts within the family, such as a notable financial improvement for mothers/fathers, decreases/increases positive inclination, and a minor financial decline for mothers/fathers increases/decreases the positive inclination. Hence, this research offers valuable insights into how household economic transitions during global crises can shape youth educational ambitions.

Declaration

I, Richa Saun, confirm that the work presented in this thesis is my own. Where the information has been derived from other sources, I confirm that this has been indicated in the thesis. This work has not been submitted for any other degree or professional qualification.

Signed:

Date: 12/1/2024

Acknowledgements

I would like to express my sincere gratitude to all those who have contributed to the completion of this thesis. First and foremost, I would like to thank my supervisors- Dr Carl Singleton, Dr Neha Hui and Dr Samantha Rawlings for their constant support, guidance, encouragement and patience throughout my PhD. Their insightful feedback and constructive criticism have been invaluable in shaping the direction and quality of this research. I am also thankful to my assessors Professor Sarah Jewell, and Dr Vivien Burrows for their comments, and suggestions on my chapters, which helped to polish my research. I would like to express my sincere gratitude to the staff members of the Economics Department for their constructive feedback during the PhD seminars. In particular, I wish to take this opportunity to thank Professor Uma Kambhampati, Professor Simonetta Longhi and Dr Stefania Lovo for their enlightening feedback during the PhD seminars. My gratitude extends to the comments and suggestions by my internal examiner Dr Andy Chung and my external examiner Professor Gurleen Popli (University of Sheffield). I am immensely thankful for their comments provided on my Thesis chapters that led to a great improvement in my Thesis.

I would like to express my gratitude and thanks to my parents, my sister and brother-in-law for their unending words of pray and financial support. Without their emotional and financial support, this thesis may not have been possible. I would like to express my affection to my nieces who have been a constant source of strength and motivation throughout my PhD journey. I am also thankful to my other relatives for their words of support and encouragement throughout the PhD process.

I am also grateful to my friends and colleagues in the Economics Department for their supportive and encouraging words. Their friendship, advice, and feedback have been invaluable. I am grateful for their support and encouragement and I look forward to achieving my future endeavours.

Finally, I would like to thank God for all the blessings that helped me throughout the PhD journey.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

"The highest rate of return in early childhood development comes from investing as early as possible, from birth through age five, in disadvantaged families. Starting at age three or four is too little too late, as it fails to recognise that skills beget skills in a complementary and dynamic way. Efforts should focus on the first years for the greatest efficiency and effectiveness. The best investment is in quality early childhood development from birth to five for disadvantaged children and their families."-James J. Heckman

In the past two decades, economists have shifted their attention from child poverty to child development. The shift in the focus away from child poverty indicates that child development is not a direct impact of the parental income, although income can not be denied as one of the important determinants of child poverty, but it does not mark as an appropriate measure of the resources available to the child (Heckman, 2011). Before delving deeper into the inequality of resource allocation between families, it is important to assess the early stages of child development and to recognise that children are born into families/households over which they have no control. Cunha & Heckman (2007) has recognised the fact that childbirth occurs by chance in a family and is an important consideration that a child is born with their parents' genes, education, economic status, and surrounding environment, all of which shapes a child's growth. Hence, focusing on family impact on the early child outcomes such as cognitive and non-cognitive (behavioural-emotional) development plays an important role. Focusing on family dynamics, intact families show a greater investment of resources in their children compared to less stable families and single parent households (Heckman, 2011). These gaps are more pronounced in the early ages of child development. There have been increasing numbers of family environment disparities on child development in recent years in terms of parental inputs (Doepke & Zilibotti, 2022). These disparities continue throughout childhood until adulthood, the outcomes of which can be seen in the form of disparities in school test scores, continuing education to university or college, and employment status.

Apart from the household input, the timing and the effectiveness of the UK government policies are important to address the growing inequality in the early stage development of children, from the birth of the child. There are challenges in the effectiveness of human development policies, for example, the outcomes of Education For All declaration in 1990, have led to significant advancements in school access and levels of literacy and numeracy that raised children's potential. But there lacks equity in the outcomes of such measures. The urban population still receives the majority share of funds and the poor rely on government spending on primary education in many countries, resulting in much less secondary-level spending and none on tertiary education. The educational inefficiencies equally exist in developed countries and policy inefficiency has been highlighted by Cunha & Heckman (2007). The timing of these policies plays a pertinent role in the adult outcomes. There are stages in the child's development which are called "sensitive periods" for the acquisition of cognitive and non-cognitive skills. Hence, economists Cunha & Heckman (2007) & Heckman (2011) emphasise the early stages of child development as a "critical period" for skill development. The effectiveness of the policies also depends upon the fact that skills are skill development compounds itself, building on previous development in order to continue developing. Hence, a higher stock of cognitive skills in one period raises the stock of cognitive skills in the next period. The dynamic complementarity here illustrates that skills acquired in early years support the formation of later life skills acquired when children attend school or university. It also implies that early investments should be followed by later investments for the early investment to be productive. This is important in the case of programs aimed at disadvantaged adolescents. Cunha & Heckman (2007) has highlighted that disadvantaged adolescents' educational returns show that educational investments made in a previous time period ranging from early childhood to adolescent is more productive and yields higher return at later stages of the child's development path for more able children. Also, the dynamic complementarity is seen in children who have greater early cognitive and non-cognitive skills, meaning they are more efficient in later learning of both types of skills. As pointed out by Heckman (2011), the dynamic complementarity is evident from the early intervention by the Abecedarian, and Perry Preschool programs which promote greater efficiency in learning in school and help reduce behavioural difficulties.

The current thesis uses The ONS Millennium Cohort Study (MCS) dataset to find the impact of parental and household inputs into cognitive and non-cognitive outcomes of child development, specifically disadvantaged children. The MCS is a survey focused on the child development milestones starting from birth through to teenage years. The survey is conducted every three years with the first one starting in 2000-02 (England and Wales: 2000-01, Scotland and Northern Ireland: 2000-02) where the average age of the child is 9 months. The survey is conducted in all four UK countries (England, Wales, Scotland and Northern Ireland). The dataset focuses on the parental figures and asks questions about early years development milestones like grossmotor and finemotor development which are essential to child development. From age 3 onwards, the child interviews are conducted to record their age-specific cognitive child development such as vocabulary scores and non-cognitive outcomes such as total difficulties and proscocial skills. The chapters in the thesis utilise cognitive and behavioural development questions asked to the child throughout the sweeps, which are important in addressing the growing inequality in educational performance among advantaged and disadvantaged households. The set of questions asked centres on their socioeconomic well-being, employability, and test scores from child development assessments which help government agencies focus on different aspects of the economy by designing ward-specific policies that act as an key input to children's early years educational development.

To give broader relevant context from the UK, the UNICEF (2018) report using the Eurostat data for 2018, shows that the percentage of children enrolled in pre-school between the age of 3 and the first years of school in the UK are between 60%-80% lower than the majority of other European countries and between 20%-30% of childcare under 3 years. The UNICEF (2018) 'Progress in International Reading Literacy Study' (PIRLS) 2016 aims to evaluate the reading comprehension performance of 4,000 grade 4 students (average age of 10 years old) from 150-200 schools per country based on the difference between the performance of the upper 90th percentile and the lowest 10th percentile in the Reading Literacy tests. Within the UK, in England, there is a performance gap of 200 points and 202 points in Northern Ireland. The overall ranking of the UK is between 25 and 24 among the 31 countries included in the study, therefore ranking among the lowest European countries and slightly higher in ranking than Australia, Chile, Israel and New Zealand.

The current research focuses on the household heterogeneity in shaping the educational choices of the children from early age to youth. There are studies (Almond & Currie, 2011; Anders, 2017; Bono et al., 2016; Cobb-Clark et al., 2018; Clifton-Sprigg, 2016) focusing on the parental years of schooling, employment, socioeconomic status impact on the child outcomes in terms of years of schooling and class performances, conducted on various parts of the world. Li et al. (2022) has highlighted that there are two parental pathways which have a strong impact on the academic performance of the child, namely the parents competing for high-quality education for their children and parenting behaviour and educational support developing children's learning habits. There are regional variations in student performance where urban students' academic performance is heavily affected by their socioeconomic status (SES) compared with more rural areas. The higher the family's SES, the greater the children's academic achievement. Findings by Chevalier et al. (2010) using British cross-section data, highlight the intergenerational transmission of education and finds that leaving school early is related to parental background, using the least squares estimation there still remains a significant impact of maternal education after including household income on the schooling of the children. The analysis by Chevalier et al. (2010) also highlights the significant impact of maternal education on the number of years of schooling for daughters which becomes stronger as daughters age. Current research is motivated by the findings by Francesconi & Heckman (2016) on the importance of attributes shaped in early life on the child's later life outcomes, which shows the significant impact of socioeconomic inequality faced by the child in their early years, also the early emergence of skill gap as a result of children born to high-achieving parents resulting in children born to smarter parents achieving more as compared to their peers, highlighting the importance of heritability in child achievements later in life. Also, Attanasio et al. (2022) focuses on children's experiences in early life being important in determining their cognitive and socio-emotional development later in life, and looks into parental ability to invest in a child's development, termed as a financial constraint and informational constraint, where parents are shown to have an impact their child's level of human development. Hence, to analyse the process of child development it is crucial to recognise the growing heterogeneity in parental input.

The present research focuses on child development, starting from age 3 to late teens. To find the impact of parental and household input into child development, current research uses the MCS to analyse data on the household heterogeneity in terms of parental role throughout the child's development process. By using the child production function introduced by Cunha et al. (2005), also used by Bono et al. (2016), Cobb-Clark et al. (2018) and Dickerson & Popli (2016), the current thesis analyses child outcomes in terms of the cognitive and non-cognitive development right from 3 to late teens. The young-age child outcome is analysed in terms of the youth preferences to continue in education after taking into account the endogeneity of child's previous age achievements.

To begin with, the first chapter of this thesis focuses on the role of the maternal and paternal language spoken to the child aged 3-7 years and its impact on the child's cognitive and behavioural outcomes. Under the child production function framework, this chapter uses the MCS panel data from the UK to find the impact on cognitive and behavioural development. The data consists of cognitive and behavioural measures such as the British Ability Scales (BAS) and BAS Word Reading test for vocabulary development; Pattern Construction and National Federation for Educational Research (NFER) number skills for numerical ability; and Strengths and Difficulties (SDQ) scores for measuring internalising and externalising (also known as prosocial) behaviours of the child.[1] The panel nature and the presence of the essential indicators of the cognitive and behavioural measures help to highlight the outcomes of child development throughout these age groups. The UK is a multi-ethnic and multicultural country, hence, addressing the child outcome gaps as a result of household heterogeneity such as the language spoken (as compared to the English language spoken at home) by the parents plays a significant role in the child's cognitive development. Previous studies by Cobb-Clark et al. (2018); Duursma et al. (2007); Sorenson Duncan & Paradis (2020); Hakuta & Diaz (1985) have highlighted the importance of the maternal and paternal roles in the language spoken at home. The longitudinal nature of the MCS data helps to analyse the diverging (between ages 3-5 years) and converging (between ages 5-7 years) gaps in cognitive development by using the child production function approach. The main findings of this chapter show that both bilingual parents have a consistent negative impact on vocabulary scores but a positive impact on behavioural scores throughout ages 3 to 7 years. The maternal negative impact on vocabulary scores of being bilingual is until age 5 with no considerable impact on behavioural development. But, having both bilingual parents at age 3 and a mother only bilingual at age 7 leads to more behavioural issues in the child. The paternal bilingualism leads to a significant increase in age 3 behavioural issues accompanied by a marginally significant positive impact of 0.1 standard deviation of lagged maternal bilingualism. But by age 7, there is an emergence of significant maternal bilingualism leading to an increase in behavioural issues by 0.26 standard deviations and a significant decline of 0.1 s.d. of paternal bilingualism. But, there is a significant decline of age 5 maternal bilingualism on the behavioural development. The current findings corroborate with the previous findings by Cobb-Clark et al. (2018) on the overall impact of bilingual households on the child vocabulary and behavioural scores. In contrast to the previous studies by Cobb-Cobb-Clark et al. (2018) & Clifton-Sprigg (2016), the age-specific findings in the first chapter show positive significant impact of individual parental bilingualism

[1]The vocabulary and numerical abilities correspond to cognitive abilities.

on the behavioural issues. Therefore, the findings on the parental bilingualism impact on the child's vocabulary and behavioural outcomes in the first chapter resemble the previous findings by Hammer et al. (2009) in the US, & Sorenson Duncan & Paradis (2020) in Canada & De Houwer (2007) & Mattock et al. (2010) in the European countries. Hence, the first chapter has contributed to the literature on parental input into child development, focusing on paternal, maternal impacts in the UK.

The second chapter focuses on the parental mobility decisions on the child's cognitive and behavioural development. This chapter studies the impact of moving to higher or lower deciles of the IMD on the cognitive and behavioural outcomes of the child. This is achieved by focusing on the child development outcomes as a result of the households' movements to areas having different deciles of the current residence IMD, between ages 3 and 7, 3 and 11 and 3 and 14 years. The analysis highlights the impact of the parental decisions taken at age 5 of the child, to whether to move to a new residence (having different decile than the current residence IMD decile) compared to those who did not move to a residence with a different IMD or did not change their residence at all. This chapter also utilises the panel nature of the MCS data in the UK to find the change in cognitive and behavioural development for age groups 7 to 11 by focusing on the parental mobility decisions made at age 3. The findings show a significant decline in the vocabulary scores of those moving to lower deciles of the IMD at age 7. Previous studies (Beck et al., 2016; Fiorini & Keane, 2014; Flouri et al., 2013; Gambaro & Joshi, 2016; Gambaro et al., 2017) have focused on the impact of change in residence on vocabulary and behavioural scores and have highlighted the negative impact of household dynamics such as moving house due to changes in parental marital status, parents facing frequent transitions in employment status, maternal health issues, and economic instability on the cognitive and behavioural development of the child. This chapter focuses on the marginal impacts of moves to the lower/higher decile of the IMD and includes the current decile of the IMD as control by recognising the importance of previous period child inputs in the child production function framework. The findings present a significant decline in vocabulary development but has an insignificant impact on behavioural scores. This allows the conclusion that movements to area having lower deciles of the IMD at an early age leads to a significant decline in vocabulary development. But, on average, the impact on behavioural outcome is insignificant when moving from higher to lower deciles of the IMD.[2]

The final chapter in the thesis is an analysis of the impact of COVID-19 on the educational preferences of 17-year-olds. This chapter combines the MCS COVID-19

[2]The MCS data lacks the neighbourhood characteristics that could explain it in more detail.

survey data with the MCS data collected pre-COVID up to age 17 in the UK. The MCS COVID-19 data is collected for three COVID-19 waves (April 2020, September 2020, and February 2021) and consists of information on parental employment during the pre-COVID period. The chapter focuses on information about parental Standard Occupational Classification (SOC) and measures the impact of the SOC on child preferences after COVID. The preferences of the young people are classified as "whether they want to continue with/start the course or drop it". These preferences help to capture the impact of COVID on young people's preferences to join the labour force in the coming years. Previous studies (Anders et al., 2021; Hupkau et al., 2020; Major et al., 2021; Werner & Woessmann, 2021) have highlighted the scarring effect of the COVID-19 pandemic on the global issues of child development and the labour market in the UK. The chapter highlights the emergence of young people not in education, employment, or training (NEET), increasing the pool of discouraged workers. The chapter uses the parents' employment status, and the change in the financial situation showing that the COVID-19 pandemic hindered educational mobility in the UK. The individual parental analysis shows that the heterogeneity in parental job transition leads to negative maternal impacts on the likelihood of continuing education and positive impacts of paternal employment on the likelihood of staying in education. Hence, this chapter has highlighted the importance of the rising issue of the NEET 17-year-old cohorts as they comprise the group with the potential to enter the labour force after completing their education. Evidently, they face declining involvement in education, which negatively impacts the economy's Human Development Index (HDI). Therefore, the third chapter highlights the responsiveness of the youth to the changes in their education and career paths and puts forward the need to address the severity of the educational shock due to COVID-19.

The findings in this thesis have important implications for policymakers. For instance, policymakers should focus on parental household heterogeneity after considering the socioeconomic factors relevant to child development at an early age. By realising the importance of early-age household, neighbourhood and school input, policies aimed at children's cognitive development should provide targeted support to schools, by helping to conduct regular activities that focus on cognitive development. The child's cognitive performance should be reported to the parents and the schools should provide the necessary training required to parents after considering the family heterogeneity. Hence, by following this approach, the government can effectively implement policies which will lead to the closing of gaps in the child's cognitive development. For children's behavioural development, the government should adopt area-specific policies that play a major role in the determinants of child behaviour later in life. There is also a need for the government to focus on the inputs determining youth's career prospects. Although there lies an importance of early life child intervention, there also exists a need to design policies that help address the youth moving away from education by supporting the selected youth financially who have sound educational achievements but whose further academic achievements have been hindered by shocks such as COVID-19.

Chapter 2

The Impact of Individual Parental Bilingualism on Children's Vocabulary and Behavioural Development in the UK

2.1 Introduction

Language has a pivotal role in the early ages of child development (Bialystok et al., 2012; Byers-Heinlein & Lew-Williams, 2013; Taylor-Leech, 2013). A potential challenge for some parents appears when a child is exposed to more than one language at home. Previous research on child development has found that children have greater ability to grasp more than one language at an early age (more recently Brotherson, 2005; Huttenlocher et al., 1991; Schiller, 2010 and previously Asher & García (1969) and Snow & Hoefnagel-Höhle (1978)) However, De Houwer (2015) stressed the importance of harmonious bilingual development on a child's well-being. The term 'harmonious' refers to the exposure of the bilingual child to two or more languages at home, leading to children's largely equal proficiency in the languages spoken at home. The degree of proficiency depends upon the inclination towards bilingualism in early years. In bilingual families, children face conflict related to the interpersonal relations between themselves and older family members regarding the languages spoken, raising concern over the harmony or well-being of these bilingual families. Early life well-being affects the later life of the children and the individuals surrounding the child (relatives, peers) as they grow. Hence, harmony between the two or more languages would mean bilingual parents/family individuals having neutral views regarding the language spoken to and by the child. This complexity in parents' decisions is faced by most bilingual families worldwide but varies depending on the community in which they reside. Moreover, the bilingual parental use of the resident country's language relies on the degree of attachment to the mother tongue language, the expectation from their child to learn the mother tongue, and the willingness to assimilate into the community by learning the resident country language. Common questions faced by bilingual/multilingual households are whether their child would be confused by speaking more than one language, whether speaking more than one language could make them fall behind in learning the resident country language, and whether the child should initially gain a good knowledge of their mother tongue and subsequently learn other languages.

In support of the above questions, findings by Hammer et al. (2009) suggest that maternal usage of Spanish did not affect children's English vocabulary, hence, mothers should not be afraid of speaking in their first language in the fear of child losing interest in English. Rather, it is found that maternal input was necessary for the children to learn Spanish. In the same vein, Byers-Heinlein & Lew-Williams (2013) found that bilingualism will not confuse; instead, it could improve a child's knowledge of vocabulary. One of the earlier studies by Hakuta & Diaz (1985) highlights the importance of additive bilingualism. Additive bilingualism refers to bilingual children learning the second language in addition to their mother tongue, without any fear of losing their ethnicity or linguistic identity. They point out that introducing bilingualism only in transitional programs in the US could make children adapt to subtractive bilingualism. Hence, additive bilingualism helps bilingual children to adapt a new language into their set of skills and improve cognitive skills. In contrast, a bilingual child in a subtractive bilingualism environment faces a replacement of its first language by the second one in the form of usage of the second language as an exclusive form of instruction in schools.[1]

This paper explores the early ages of children in the UK before entering school at age 3 and at the first year (Key Stage 1; age 5-7) of school in a two-parent family, so, spanning ages between 3-7 years old. The study uses the longitudinal nature of the MCS to estimate a child's age-specific and cumulative development of cognitive and behavioural patterns. Strand et al. (2015) in their study of English as an Additional Language (EAL), studies show that overall, EAL students catch up with the First Language English (FLE) students by age 16 in vocabulary assessment. However, Strand et al. (2015) also reports that within the EAL group, the performance gap is more than that between EAL and FLE students. Factors that influence the EAL performance gap are ethnicity, age of arrival of foreign-born children, change of school, neighbourhood deprivation, region of residence, age in months and gender. Cobb-Clark et al. (2018) have used the above factors in their analysis of the impact of household bilingualism on the vocabulary of children aged 3-14, and behavioural outcomes. The current study focuses on the impact of parental identity in bilingual (English and other language) households on children's vocabulary and behavioural development. Hence, it is important to understand the role of parental identity in shaping the child's outcomes. In the context of present research, the simplest way to understand the parental identity is to focus on the significant parental role the child is exposed to given the family environment and the individual maternal and paternal roles such as the degree of maternal and paternal proximity of the child and the parental time investment influencing the vocabulary and behavioural development of the child. Cobb-Clark et al. (2018) also using the MCS dataset, find

[1] The National Curriculum Programme in England is in line with Hakuta & Diaz's (1985) contribution. The programme aims to enable children to communicate their ideas and thoughts through speech and writing by using their knowledge of phonology and grammatical structure. Therefore, the child learns to compare their use of English grammar and spelling to foreign language (Department for Education, 2013).

that bilingual children's vocabulary development is not significantly different from that of monolingual children. Similar studies on bilingualism in the US by Hammer et al. (2009) on 3-year-old children in Head Start and kindergarten and Duursma et al. (2007) on fifth-grade English learners find that the parental usage of English at home does not significantly impact the vocabulary test scores. The US studies discussed above considered proficiency as a measure of the vocabulary outcome, while in the UK, Cobb-Clark et al. (2018) use the MCS data and Clifton-Sprigg (2016) uses Scottish Government data find the relative performance of bilingual children in vocabulary as well as in behavioural outcomes.

This research will contribute to a deeper understanding of bilingualism in the UK and, to some extent, in the US and Canada.[2] In the UK, the study delves into the role of the parental identities speaking other than English to the child, highlighting the importance of household heterogeneity in child development outcomes. By exploring the early ages of children in the UK before entering school (age 3) and at the first and second years (Key Stage 1; age 5 to 7 years) of school in a two-parent family. This study aims to achieve this by using the longitudinal nature of the MCS to observe the heterogeneity of bilingual families and exploring parental level data, such as whether the mother, or father speak non-English languages compared to their monolingual counterparts from ages 3 to 7.

By using the production function introduced by Cunha et al. (2005), Cunha et al. (2010) & Todd & Wolpin (2003), also used by Bono et al. (2016), Cobb-Clark et al. (2018) & Dickerson & Popli (2016), the current study analyses child outcomes as a function of individual parental bilingualism. In line with Cobb-Clark et al. (2018), the lagged child test scores are controlled to avoid biased estimates due to achievement at earlier ages. The research design is a cumulative value-added model; this approach is not

[2]Using the English vocabulary test scores and behavioural development test scores together for children aged 3 to 7 years, the current study aims to fill the gap in the reflections in the US and Canada, consisting of separate vocabulary and behavioural analysis in the US and only vocabulary analysis in Canada. Moreover, the controls used in the present study are similar to those used in child development literature for Canada and the US.

new to this data type, and Cobb-Clark et al. (2018) have also used this approach. Following Cobb-Clark et al.'s (2018) approach to the lagged bilingual status of parents, the current study uses the contemporaneous and lagged parental bilingual status. The main findings of the present study show that contemporaneous maternal bilingualism has a significant negative impact on vocabulary development throughout ages 3 to 5 which becomes positive and significant at age 7. Hence, the considerable negative effect of a mother speaking non-English on vocabulary scores endures up to age 5. Also, there is no significant bilingual penalty of age 3 parental bilingualism on age 5 and 7 vocabulary development. Over time, maternal bilingual penalty inflicts negative vocabulary scores and lagged vocabulary scores diminish their positive impact throughout ages 5 to 7, also, lagged behavioural issues has a declining impact on vocabulary development. At the same time, individual parental bilingualism has varied impact on behavioural development. At age 3, there is marginal significant positive impact of maternal and significant positive impact of paternal bilingualism which is transferred to highly significant positive impact of maternal bilingualism and highly significant negative impact of paternal bilingualism on age 7 behavioural issues, with a significant negative impact of age 5 maternal bilingualism on age 7 behavoural development. Over time there remains no significant impact of parental bilingualism on behavioural issues. Also, there is a prominent increase in behavioural issues due to increasing lagged behavioural outcomes but a persistence decline of the impact of vocabulary scores.

Hence, by using the value-added models across ages 3 to 7, the current findings coroborates with Cobb-Clark et al. (2018) on the overall impact of bilingual households on vocabulary scores. The current study finds that a significant negative impact of individual maternal bilingualism exists on vocabulary scores which remains negative till age 5, but turns significant and positive at age 7. The age-specific analysis shows that behavioural issues increase significantly with an increase in the lagged SDQ scores. Over time, no significant impact of parental bilingualism exists on behavioural issues, but the behavioural issues increase significantly for the mother speaking in non-English at age 7. Hence, the current study contributes to the child development literature on individual parental bilingualism's impact on child development scores in the early years by using the value-added models under the framework of the child production concept.

The remainder of the chapter is organised as follows: section 2.2 discusses some related studies; section 2.3 describes the data and variables used for analysis; section 2.4 shows the econometric framework used in the analysis; section 2.5 presents the results; section 2.6 discusses the main findings and compares them with previous studies; and section 2.7 concludes.

2.2 Literature Review

To date, several studies (Bialystok, 2001; Cobb-Clark et al., 2018; Goldberg et al., 2008; Hammer et al., 2009; De Houwer, 2007; Kirk et al., 2015; Sorenson Duncan & Paradis, 2020) have confirmed the positive relationship between bilingualism and cognitive development by using factors such as maternal language use at home, parental education, parental migrant status, country of birth, SES, and child output to the language spoken at home as inputs in their longitudinal studies.

Byers-Heinlein & Lew-Williams (2013) focuses on the bilingual parents' concern of speaking a native language at home as it can confuse their child. They find that bilingual infants are sensitive to the information that distinguishes the languages spoken at home and find no evidence that bilingualism could lead to confusion and intellectual fatigue; instead, it acts as an input to a child's cognitive development in later life. But, it also suggests that bilingualism is not essential input for child development. Instead, it focuses on the importance of the strategies adopted by bilingual families in choosing the language to communicate to children where a one-person-one-language approach is not necessary. Pearson (2008) describes the confusion among bilingual children as code-mixing: mixing words of the language spoken at home and other different languages outside the home, which, among bilingual children can lead to an increase in their vocabulary size (Lanza, 1997). Moreover, Bedore et al. (2012) points to the fact that bilingual parents have better knowledge of language development and can provide a good measure for the current vocabulary skills of the child. Bedore et al. (2012) highlights a significant impact of different language measures (semantics and morphosyntax) on proficiency. Their study on 1,029 Spanish-English bilingual pre-kindergartners (aged 4-5) provided evidence of the interaction of language characteristics, such as the frequency of constructions and the user's phonetic salience. They stressed that the age of first exposure matters less than the current use, highlighting a strong correlation of 0.95 between the input and output of the language. It is evident from their findings that the age of exposure provides 35% of variance in language dominance, whereas, the current use of language explains 60% of variance.

De Houwer (2007) uses a survey of 1,899 bilingual families with children aged between 6-10 in Flanders, Belgium, consisting of one of the parents speaking a language other than the majority (Dutch) language. Her study involves the individual role of mother and father speaking a minority language, and suggests that parental use of minority languages has no impact on a child's use of a minority language. Also, findings from her research reveal that teaching children to speak two languages highly depends on parental input patterns, and parents should decide on language patterns at an early stage of child development. In the same vein, Mattock et al. (2010) observed the word learning performance of 17-month-old infants speaking English only, French only and bilingual English-French children using Switch task and found that bilingual children can use their varying phonetic skills to identify words, and bilingual parents should use both languages to help their child gain knowledge of both home country and host country languages.

Bialystok (2001), Collins et al. (2011) & Yang et al. (2011) suggest that language competencies have significant impact on children's well-being and those who have dual language competencies feel more confident and secured. Collins et al. (2011) finds that Spanish and English competencies account for 13% of variance in interpersonal strengths and 25% of total variance of intrapersonal strengths. However, some studies have found bilingual children have more behavioural issues than monolinguals and found that the children of immigrants exposed to more than one language exhibit shyness, low social skills, and assertive behaviour compared to native children (Cobb-Clark et al., 2018; Özerk et al., 2011; Spomer & Cowen, 2001; Sun et al., 2021).

2.2.1 Household Impact

This subsection highlights the household-only specific dynamics shaping the child's cognitive and non-cognitive development from age 3 onwards.

Carneiro et al. (2003) and Cunha et al. (2010) analyse the importance of cognitive and non-cognitive skills in the early years, shaping socioeconomic success in the later years of life. Cunha et al. (2010) uses Peabody Picture Vocabulary (PPVT) at ages 3-4 and at ages 5-6 as well as measures of maternal cognitive skills focusing on language development such as Word Knowledge and Paragraph Composition. Hence, by giving importance to the child's cognitive tests, the findings by Carneiro et al. (2003) & Cunha et al. (2010) demonstrate the essential role of families and schools in shaping the human capital accumulation of the child in terms of cognitive development from early age. Delving into the household inputs, economists such as Chiswick & Gindelsky (2014) in the US & De Houwer (2007) in Belgium found that the prevalence of bilingualism in the home environment depends upon the country of origin of the parents, whether they are foreign-born, sharing similar linguistic ancestry and whether a child is living in a neighbourhood where the parents' language is spoken. Also, Cummins (2005) in the US suggests that not all bilingual parents in the host country make their children use their home country language regularly. This is explained by the findings of Chiswick & Gindelsky (2014) in the US that having at least one parent proficient in English reduces the probability of a child speaking in other languages by 19%, at the same time, having a foreign-born parent increases the probability by 39%. Asian and Hispanic children are 19% and 20% more likely to speak a non-English language at home compared with their non-Hispanic and non-Asian counterparts after controlling for other factors. These household bilingualism studies highlight the fact that a significant relationship exists between the degree of parental attachment to their home country's language and the language spoken at home with the child.

Studies by Byers-Heinlein & Lew-Williams (2013), & De Houwer (2015) suggest bilingualism is a key to cognitive learning. However, parents in bilingual households limiting a child to learn only the host country's language delays a child's cognitive development. This is validated by the recent study on parental fluency by Sorenson Duncan & Paradis (2020), highlighting the impact of maternal linguistic input and education on Language 1 (L1) (first language learners i.e., monolingual), Language 2 (L2) (second language learners i.e., bilingual) on the L1 scores on the child's syntactic ability. The Sorenson Duncan & Paradis (2020) study suggests that children who have more years of exposure to L2 in school and interact more in L2 with their mother had more knowledge of word types in L2. Also, the degree of cognitive development not solely depends upon the parental inputs but also on the child's abilities as shown by Métraux (1965) using the historical cross-sectional survey, consisting of a small sample of 25 bilingual families (French and English). The Métraux (1965) findings show that children who spoke early and could not stop talking, extrovert and adaptable to new situations quickly learned both languages. In contrast, introverted, resistant children show difficulty adapting to new conditions and the children with less physical energy were late second language learners, hence, highlighting the role of child attributes as determinants of language development.

Therefore, in a bilingual household, a child could quickly learn both languages. As Clifton-Sprigg (2016) demonstrates in her study on the impact of bilingualism in young children under six years of age, where bilingual children's cognitive and non-cognitive skills are compared with monolingual children. In her research, the bilingual child has one parent foreign-born and the other UK born with monolingual children where both parents are natives. The findings of her study show an average low score in an English vocabulary naming exercise from the bilingual children compared to their monolingual counterparts at the age of 3. However, this gap in vocabulary test scores converges by age 5. She finds no significant positive impact of bilingualism on emotional development. In the same vein, Cobb-Clark et al. (2018) using the MCS data finds the bilingual gap in vocabulary to start converging at age 7 in general but he does not address the individual parental role in vocabulary development. At the same time, they find a positive relationship between bilingualism and the child's emotional development, mainly in the case of boys.

This paper focuses on the importance of individual parental bilingualism on child outcomes. The study is motivated by the previous findings of Goldberg et al. (2008) in Canada & Hammer et al. (2009) in the US, highlighting the importance of maternal bilingualism in child outcomes. Goldberg et al. (2008) points to the maternal use of English at home as one of the input factors in English as a second language (ESL) children. Their findings suggest no significant impact on the average use of English at home or in the PPVT outcome measures for 19 minority L1 children. Successively, Hammer et al. (2009) in the US, suggests that mothers with a poor command of English can cause their child to develop less well-formed sentences in English. In contrast, the mother who speaks in the home country language produces more well-formed sentences and rich English vocabulary. Hence, in line with these findings, the current research highlights the age-specific maternal role in influencing child outcomes and finds a significant positive impact of maternal bilingualism on the child's outcomes through ages 3 to 7 with a significant positive impact as the child reaches age 7.

2.2.2 Review of measurement approach and research designs used in previous studies

As the literature on bilingualism develops, authors have switched to different vocabulary and non-cognitive development measures to analyse the relationship between language and child development. Hence, the following sections highlight some measurement and research design considerations of the impact of bilingualism used by multiple authors.

Measures used

Bedore et al. (2012) & Reyes & Hernández (2006) study English and Spanish language development in bilingual children. Bedore et al. (2012)'s study estimated semantic and morphosyntactic development using the Bilingual English Spanish Assessment, consisting of sentence completion and sentence repetition tests designed to identify speech and language learning impairments of children aged 4 to 6, whereas Reyes & Hernández (2006) used the Boston Naming Test and the PPVT in both English and Spanish. The interviewer administered the tests to the children to obtain an index of the children's ability in both languages, specifically naming and comprehension vocabulary. Similarly, Kim et al. (2020), in Canada, used Education Quality and Accountability Office reading measures such as multiple-choice questions, open response reading, and writing comprehension questions to measure English literacy for grades 3, 6, and 10 students. In the UK, studies on bilingualism (Cobb-Clark et al. (2018) and Clifton-Sprigg (2016)) also focus on children's bilingualism at home and exposure to reading activities. Cobb-Clark et al. (2018) & Clifton-Sprigg (2016) measure vocabulary development through age-specific British Ability Scale Test (BAS) on naming vocabulary at ages 3 and 5 and word reading from age 7 to 14. In the BAS naming vocabulary, the child uses their knowledge of objects to identify the picture shown in the flashcards and is assessed on their spoken vocabulary.

Clifton-Sprigg (2016) also used the Picture Similarity measure, which consists of a booklet of images where the child matches the card with a picture in the booklet.

Research designs

The most common research designs used in bilingualism literature are experimental survey designs, cross-sections (in case studies, primarily found in earlier studies) and causal designs. The survey design studies are characterised by the homogeneous selection of the bilingual children, i.e., going to the same kindergarten or prekindergarten, high degree of neighbourhood proximity, having the same mother tongue and the majority of them having the same language dominance. However, the survey designs have paucity as they cannot accumulate the changes in the bilingual status of the household and mobility of the family. Hence, most of the survey questions asked to the bilingual households about the language spoken in the home are population-specific and thus lack comprehensive coverage of the impact of bilingualism. Another type of research design used in the literature on bilingualism is cross-sectional case studies—the case study by Sorenson Duncan & Paradis (2020) of immigrant children in Canada with different L1. Their analysis finds that the frequency of L2 language spoken by mother at home and higher education of mother positively depends upon the language (L1 or L2) which the mother acquired in education. The longitudinal survey data where the findings are based on age-specific measures of cognitive and non-cognitive abilities are used by Cobb-Clark et al. (2018) & Clifton-Sprigg (2016). The current study indicates that bilingual children fare worse in the early ages (3 to 5) and catch up with their peers later, with a significant positive impact on vocabulary scores at age 7.

Methodological approaches

In the UK Bono et al. (2016), Clifton-Sprigg (2016), Cobb-Clark et al. (2018) & Del Boca et al. (2012) have used the child production function approach of Todd & Wolpin (2003) to study the child vocabulary and behavioural outcomes. Bono et al. (2016) focuses on maternal time investment on later child development by using the Principal Components measure. In the above UK studies, the parental language inputs to child development was only analysed by Clifton-Sprigg (2016) & Cobb-Clark et al. (2018). However, all these studies use lagged-period outcomes as a standard instrumental variable to account for the dependence of child ability and current period inputs. The impact of other family factors, like siblings, or area of residence, also accounted for unobserved school effects.

The above studies use time-fixed effects to account for the endogeneity caused due to unobserved school effects, the child's unobserved genetic endowment, household characteristics and bilingual parental decisions to speak a second language at home. Whereas, in the US there are different methodological approaches used in the study of the impact of bilingualism on child development. Some of these are multilevel modelling used by Hoff et al. (2018) to analyse the impact of maternal education on their child's English vocabulary outcomes by conducting a survey study of 92 US-born bilingually developing children. Bohman et al. (2010) used a different approach to logistic growth modelling. Their study finds that the language inputs and outputs have different impacts when a child achieves language semantics and syntax in both types of language. Therefore, the methods used in the above studies are appropriate to ensure unbiased estimates given the type of data. Overall, the studies accounting for the fixed effects provide fair estimates to a large extent.

Limitations of measurement approaches and research designs

The above measurement tests used are not perfect. An earlier study by Bowles & Salthouse (2008) found that the vocabulary test consists of various test formats such as multiple choice synonyms, multiple choice antonyms, producing definition and picture identification. These formats differed in their relationship with the age of the participant. Hence, the early adult participant has a higher score in picture identification and the lowest in identifying definition. At the same time, the late adult participants have the lowest score in the picture identification and the highest in the MC synonyms. Therefore, their findings suggest that using only a single indicator of vocabulary score provides incomplete and biased results about vocabulary knowledge. The current study uses the age-adjusted (accounting for a three-month age gap) vocabulary scores to account for this.

The viability of parental questionnaires was studied by Patterson (1999), who finds that parental questionnaires used to evaluate the home language use of the child are based on the languages the children speak to the reporting parent. Hence, the mean vocabulary size outcome does not provide valid results for bilingual children, whereas the current study using the MCS dataset consists of questions such as language spoken at home with the child, which are asked directly to the parents. Conclusively, it is hard to recognise the best measure of the child's vocabulary. The vocabulary tests differ in their formats and language, where bilingual children can attach their understanding of both mother tongue and the host country language to respond to the questions in the test. Regarding the study design, the cross-sectional studies do not provide the validity of findings to a broader population. Therefore, compared to longitudinal studies that use the time, and household characteristic fixed effects, the cross-sectional studies limit the conclusions to a specific population and lack individuals comprising important characteristics particular to the specific research.

2.3 Data

The analysis uses the MCS longitudinal data, consisting of household inputs (language spoken by the parents, parental time investment, SES, etc.) into child development. The MCS is a collection of approximately 19,000 children born between 1 September 2000, and 31 August 2001, in England and Wales and between 24 November 2000, and 11 January 2002, in Scotland and Northern Ireland.[3] The MCS data set consists of seven sweeps to date. The ages of the children range from 9 months (sweep

[3]The tables A.1 to A.12 in the appendix give details about the age of the child and the years of each MCS sweep along with the variable of interest and controls used in the current research.

1) to 17 years (sweep 7). This study utilises the first four sweeps of the MCS consisting of children aged 9 months to 7 years old, marking the early stages of their development. The information on the household SES, parental heterogeneity, and child development components (vocabulary and behaviour) is collected from the child interview questionnaires. The surveys are mainly conducted with parents/parent-like figures.[4] The survey focuses on the children (cohort members), collecting information on parents, household characteristics, siblings, teachers and neighbours. At the household level, the data set consists of information on socioeconomic conditions, ethnicity and the heterogeneity in the parental relationships.[5]

The current study uses parental and household characteristics to analyse the dynamics of the effect of the person (mother, father, and both) being bilingual in the child's development (vocabulary and behaviour). Tables A.8 and A.9 describe vocabulary and behavioural development measures. In the current study, a household is characterised as bilingual if another 'non-English' language is spoken. English-only speaking households are classified as monolingual.[6] The table 2.1 shows the descriptives of the monolingual and bilingual households for all three age groups (3, 5 and 7 years). In the case of bilingual households, there is a decline in the percentage of parental bilingualism throughout ages 3 to 7. Table 2.2 shows the descriptives of vocabulary and behavioural scores that consist of Z standardised child outcomes with mean 0 and standard deviation 1.[7] In table 2.2 there is a constant decrease in the

[4] In the MCS, questions on a child's cognitive and non-cognitive development at age 9 months are asked to the Main respondent only. 90% of the Main respondents are natural mothers in sweep 1 (9 months age).

[5] The SES condition variable used for the present study is OECD weighted quintile: Lowest, Second, Third, Fourth and Highest. Cobb-Clark et al. (2018) use below 60% of the median UK income indicator. Following Cobb-Clark et al. (2018), the ethnicity of the child is also classified into two groups: 0 "White" 1 "Non-white", but coded as assuming "White" as the base category.

[6] The bilingual family is derived as per the question asked from age 3 to age 7: Is English the language usually spoken at home? Yes-"English only", Yes-"English and other language(s)", No-"Other languages(s) only". Bedore et al. (2012), Duursma et al. (2007) & Kim et al. (2020) also used this type of method. Hence, the households where parents speak more than one language is assumed to be bilingual. Questions regarding the frequency of other than non-English language spoken were not asked from bilingual households. The non-English speaking group is small to consider separately. Moreover, there were no significant changes in the coefficients after removing the non-English-speaking group from the sample.

[7] Table A.1 shows the non-standardised child outcomes for the average ages of the child.

mean vocabulary scores from age 3 to 5, but shows an increase of 0.03 standard deviation at age 7 , whereas there is a decrease in the mean index of Strengths and Difficulties Questionnaire (SDQ) between ages 3 to 5 but an increase of 0.07 standard deviation at age 7, showing on average, by age 7, there is a decline in a child's behavioural problems. The SDQ scores consist of internalising (consisting of emotional, conduct, hyperactivity and peer problems) and externalising (prosocial behaviour) scores ranging from 0 to 10 each. The higher the internalising scores, the more the behavioural issues, whereas the higher the externalising score, the lower the behavioural issues in the child.

2.3.1 Sample Representation

The MCS survey is disproportionately stratified clustered.[8] The stratification was done based on the four countries of the UK (England, Wales, Scotland and Northern Ireland). To adequately represent the ethnic minority and disadvantaged groups, the population of electoral wards was classified as disadvantaged, advantaged and ethnic minority.[9] The variable of interest is the person speaking in non-English at home. To provide precision of the impact of the person speaking non-English on the outcome variable, the current study uses the weights at the country level provided, since bilingual households are more concentrated in the disadvantaged and ethnic minority stratum of a country as compared to the advantaged stratum. In the MCS, a total of 19,495 families (including twins and triplets) were selected for the interview. [10] The present study uses information on sweep 1 and new families entering at sweep 2 of the MCS. For analysis, I include the first cohort member in the twins family to

[8]The sample was selected systematically within each stratum and country. The sampling interval was determined by the number of wards required in the sample to the number of wards present in the population. Finally, the selection was combined with the ratio of the English regions in England and Scotland, and only by ward size in Wales and Northern Ireland.

[9]The MCS technical report by Plewis et al. (2007) characterises disadvantged ward where the population lies on the upper quartile of the 1998 Child Poverty Index (CPI) of England and Wales. These families were characterised as receiving either income support, jobseekers allowance, family credit and disability working allowance. The advantaged stratum consists of the population of the ward lying in the lower quartile of the 1998 CPI. Ethnic minority stratum was only present in England where 30% of the total population was "Black" or "Asian".

[10]Of these, 692 families entered the MCS at sweep 2 and were termed as new families

account for statistical complexity while clustering at the family level using the overall UK non-response weights.[11] Moreover, the households with triplets are not included in the study because data availability is secured access, unlike single and twin families whose data is available under an end-user licence.

Attrition

The Main, Partner or Proxy Partner responded to the MCS interview.[12] The sample consists of attrition weights where: the respondents did not answer all questions and were coded as "don't know" and, "refusal" which led to non-response attrition. Secondly, there were sample losses while conducting MCS interviews from Child Benefit records to the families present and issued to the field because some families moved outside the sample ward. Hence, there was under-representation of the new-movers to the ward, to account for this, weights are assigned to each of the four countries of the UK.[13] Also, the attrition weights allow over-representation of winter births in Scotland and Northern Ireland. I have used the MCS technical report (Plewis et al. (2007)) information to accounting for non-response weights. For analysis, the overall longitudinal UK population weights estimate outcome variables of corresponding sweeps in regression models.[14] The overall weights are used because they are adjusted for non-response and have a broader range than the sample weights,

[11]The twins were concentrated more on Disadvantaged areas and hence, formed an important sample for the analysis. I have clustered the sample using country and area-level variables because of the lack of family-level indicators used in clustering. Moreover, there is statistical complexity in the segregation of population weights among twins in the household which may lead to biased estimates that could account for the same child-level clustering. Also, there are a maximum of 250 families (keeping in mind the varying numbers as per the sweep) which have twins.

[12]The Main respondent is asked all the child, parental and household-specific questions. But selective questions are asked to the Partner respondents, such as, questions related to the involvement with the babies, grandparents and friends of the partner, parent's health, employment and education, interests and time with babies that are also asked to the Main respondent. The Main respondent also answers proxy Partner questions in case the partner is not available at the time of interview.

[13]Some families targeted for interview moved from the address registered with the National Pupil Database (NPD).

[14]The current study focuses on the child development outcome for the UK as a whole consisting of input information from households of each country (England, Wales, Scotland and Northern Ireland).

as the former combines with sample weights and non-response weights for the UK as a whole.[15]

2.3.2 Sample selection

The MCS data consists of families belonging to different ethnicities and cultures. The study considers two-parent households. The definition of parents/parent-like comprises all those who responded to the MCS survey and take care of the child.[16] The study follows the findings from the previous bilingual literature in the UK by Cobb-Clark et al. (2018) & Dickerson & Popli (2016) also using MCS data and Hammer et al. (2009) in the US & Sorenson Duncan & Paradis (2020) in Canada to hypothesise a relation between person being bilingual and child development. The current study contributes to the literature by using the information on who speaks non-English language/s at home. For analysis, the base language is English, and the households where people speak non-English are bilingual.[17]

The questions asked to the main respondent in the household are "who speaks the other language at home?", and the answers are: "mother", "father" and "both" (In the present study I have created mother and father bilingual category by combining both bilingual under mother and father variables) as shown in table 2.1 that represents a list of variables of interest.[18] There is a constant decrease in father speaking in non-English

[15]The attrition is accounted for by the overall UK weights standardised with mean 1 and multiplied together to give an overall unit non-response weight. The non-response is caused by a change in the respondent from the previous sweep.

[16]The sample selection ensures uniformity in the type of parent/parent-like figures present between monolingual and bilingual households. If a certain parent/parent-like figure is not available for bilingual households but it is for monolingual then those cases are not considered for that sweep (age). Similarly, if any of the parent/parent-like figures are absent in monolingual households that observation is not included in analysis. In the current analysis, the monolingual households consists of grandparents, brother and sister interviewed acting as parent/parent-like figure.

[17]In the MCS data, the percentage of households "not speaking in English at all" are small throughout sweeps 1 to 4. Hence, they are also assumed to be bilingual. Moreover, the focus of the study is on the person speaking non-English language/s and a child is assumed to be bilingual if one of the parents speak more than one language at home.

[18]The focus of the study is on the person speaking non-English, table A.2 in appendix gives mean outcome and controls for each parental language input on the bilingual children performance as compared to monolingual counterparts.

from age 3 to 7.[19] The percentage of mother speaking non-English decreases by 1.18 ppt. at age 5 from age 3, at age 7 it declines slightly by 0.34 ppt. from age 5. By age 7, the percentage of mother and father speaking non-English declines and increases from ages 3 and 5, respectively. Previous studies on parental language by Duursma et al. (2007), Hammer et al. (2009) & De Houwer (2007) have used the household parental language interviews to find the impact on child outcomes.[20] The final sample of the study in table 2.1 comprises 7,765 cohort members (children) at age 3 years, 7,243 at age 5 and 6,784 at age 7. Following the approach of Cobb-Clark et al. (2018), I have considered those observations where the information on the measures of child development components (word reading, total difficulty scores) and controls are non-missing throughout.[21]

TABLE 2.1: Descriptive statistics of explanatory variable of interest

	Age 3		Age 5		Age 7	
	Mother (1)	Father (2)	Mother (3)	Father (4)	Mother (5)	Father (6)
No.[%] of monoling.						
Eng. only is spoken	6,992 [90.05%]	7,069 [91.04%]	6,608 [91.23%]	6,689 [92.35%]	6,212 [91.57%]	6,259 [92.26%]
No.[%] of bilingual						
Speaking non-Eng.	773 [9.95%]	696 [8.96%]	635 [8.77%]	554 [7.65%]	572 [8.43%]	525 [7.74%]
Unweighted Total hhld.[%]	7,765 [100.00]	7,765 [100.00]	7,243 [100.00]	7,243 [100.00]	6,784 [100.00]	6,784 [100.00]

Note.-The number of households corresponds to the sample size of monolingual (English only) and bilingual (English and other) households. The sample size is calculated from the pooled data for ages 3 to 7. The number of households is the same as the number of cohort members in the sample. The statistics derived in this table 2.1 are unweighted. Overall UK weights are used in the estimation of regression coefficients only.

The analysis has two steps; the first consists of age-specific analysis for each non-missing observation on controls and variables of interest. The second step is to check for the robustness of the age-specific coefficients obtained. For the second

[19]There are households where some parents move and hence not respond to the questions. Some questions are only asked to those respondents who are present in the previous sweep/s.

[20]The variable "Who speaks in other than non-English at home?" in the MCS dataset consists of the mother, father, cohort member, and any household member speaking non-English at home. The current study focuses on mother, father, and both speaking non-English.

[21]The question on who speaks in non-English at home was asked from age 3 years onward (sweep 2).

step, the panel data is pooled to form an unbalanced panel of 13,260 observations. Further analysis of the pooled unbalanced data was done for each child in the particular sweep and had non-missing observations throughout. More detailed discussion on the specification is done in the Analysis section 2.4.

2.3.3 Components of child development

My study has two major components of child development: cognitive (measured through vocabulary scores) and non-cognitive development (measured through Total Difficulty score; a component of SDQ index). These components are studied frequently in psychology and sociology (e.g. Bialystok et al., 2012; Bowles & Gintis, 2002; Farkas, 2003; Huitt & Hummel, 2003; Keatley, 1992; Singer & Revenson, 1997) and have also been increasingly used in the field of economics. The MCS dataset is a collection of these components. The cognitive abilities at the subsequent sweeps 2 (age 3), 3 (age 5) and 4 (age 7) are observed using the age-varying tests, including the BAS at ages 3 and 5 years and BAS Word Reading test at age seven years (Elliott et al., 1996; Hansen et al., 2010). The BAS Naming vocabulary test at ages 3 and 5 years measures the knowledge of nouns of the cohort member.[22] At age 7, the BAS Word Reading test assesses the cohort member on expressive spoken vocabulary and consists of a set of coloured pictures of objects shown to the child one by one, which they are asked to name. Success depends on the cohort member's previous knowledge of a vocabulary of nouns. The BAS age-standardised t-scores are used for analysis. As a part of common practice in the literature, the t-scores are further standardised to z scores with mean 0 and standard deviation 1 to facilitate interpretation. Recent studies on bilingualism by Cobb-Clark et al. (2018) & Clifton-Sprigg (2016) have also used BAS vocabulary and SDQ index in the UK. Sun (2019) used receptive vocabulary and receptive grammar and SDQ index for behavioural development in Singapore. Hammer et al. (2009) used PPVT-III and Word Reading Ability-2 in the US and Sorenson Duncan & Paradis (2020) have used the Storytelling task in Canada.

[22]The score ranges from 0-80 for BAS naming vocabulary and BAS word reading ranges from 55-145.

Other cognitive test scores (Bracken School Readiness, BAS Picture similarity, Pattern construction and NFER Number skills) are not considered for analysis as these scores do not provide consistent measures throughout ages 3 to 7 of the child. [23] The second component of child development present in the MCS dataset is the non-cognitive development of the child. At sweep 2 (age 3), 3 (age 5) and 4 (age 7), the SDQ index is used to measure the behavioural outcomes of the child.[24]

The SDQ measurement consists of five scales: emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems and prosocial behaviour. Each scale consists of five items. The parents answer to each item based on the three components: 1 'not true', 2 'somewhat true' and 3 'certainly true'. The components range from 0 to 2. 'Somewhat true' is always scored as 1, but 'not true' and 'certainly true' vary with the item. The total score is the sum of the highest score corresponding to each item within a given scale. As there are five items for each scale, for each of the five scales, the score can range from 0 to 10 if all items are completed. The table A.9 provides information on the items asked within each of the SDQ index scale. Further, to obtain the total difficulties score I have calculated the sum of the scores of only the first four scales, excluding the prosocial score. The total difficulties score ranges from 0 to 40, ideally when all the scales are present with complete scores. As per the item description under each scale, I conclude that a higher total difficulty score for a child tends to correspond to more behavioural issues (Goodman, 1997; Johnson et al., 2015). The current study uses total difficulty z scores with mean 0 and standard deviation 1 to facilitate interpretation.

[23] The other cognitive test scores could provide more detailed insights on child development at each age. I will incorporate them for further improvements.

[24] At sweep 1 (9 months old), most questions on non-cognitive outcomes are asked to the mother, as 90% of main respondents are natural mothers in sweep 1. At sweep 1, the Carey Infant Temperament Scale is used to measure behavioural outcomes of the child (Carey, 1972). The behavioural outcomes are measured in the sub-scales of Mood (5 items), Adaptability (2 items), Approach-withdrawal (3 items) and Regularity (4 items). Table A.3 in the appendix provides a summary statistic of 9 months of cognitive and non-cognitive development. They are age three lagged cognitive and non-cognitive measures used in age-specific models.

TABLE 2.2: Descriptive statistics of Z standardised outcome variables

	Z Vocabulary Score			Z SDQ Score		
	Age 3 years (1)	Age 5 years (2)	Age 7 years (3)	Age 3 years (4)	Age 5 years (5)	Age 7 years (6)
Mean	-0.003	0.012	0.030	-0.031	-0.031	-0.016
Median	0.000	0.000	0.000	0.000	0.000	0.000
St. dev.	1.016	1.060	1.039	1.023	1.024	1.068
Min.	-3.000	-3.000	-3.000	-2.000	-1.000	-1.000
Max.	3.000	2.000	2.000	5.000	6.000	5.000
N	7,765	7,243	6,784	8,152	7,282	6,808

Notes.-The author uses the sample of pooled data to calculate descriptive statistics for a child's average ages of 3, 5 and 7. The average age corresponds to sweeps 2, 3 and 4, respectively. The non-standardised test scores variables (a combination of all three sweeps' test scores) are converted to standardised z scores with mean 0 and standard deviation 1 for each age group in the pooled data. High mean vocabulary and total difficulty scores (standardised and non-standardised) correspond to more vocabulary knowledge and behavioural issues respectively. The statistics derived in this table 2.2 are unweighted. Overall UK weights are used in the estimation of regression coefficients only. The non-standardised test scores are shown in appendix table A.1.

2.3.4 Control Variables

The MCS is an extensive collection of child-level data. From this large data set, the current study utilises few variables to act as controls. The selected variables are a collection of the household environment, parental inputs and child-specific variables.

Child-specific controls include child ethnicity, birth weight in kgs (taken for respondents asked at sweep 1 and new respondents at sweep 2) and gender of the child. Parental level information used in the analysis consists of the mother's age at the birth of the child, current age of mother and father, parental education at the time of birth of the child, the year of arrival to the UK and whether the parents were born in the UK.[25] Previous studies by Bono et al. (2016), Cobb-Clark et al. (2018) & Dickerson & Popli (2016) on child development have used the MCS data and considered the above child level and parental level controls. A brief overview of the reasons for including child and parental specific controls are described below:

[25] The table A.12 shows the detailed information about the formation of the young age arrival variable which utilises information from sweeps 1,2,3 and 4 of the MCS dataset.

SES and literacy environment at home are related. English language learner students of fifth grade in the US showed higher SES and a more literacy environment (correlation of 0.37 and 0.28) and had higher English vocabulary scores. Also, higher SES has a strong correlation of 0.29 with the household literacy environment, such as highly educated mothers, having more books, parents frequently reading (Duursma et al., 2007). As the current study focuses on the person speaking non-English, parental level information on the year of arrival corresponds to the familiarity with the host country's language.[26] Staying proficient in English does not require parental inputs (Duursma et al., 2007). The parent's birthplace is a good indicator of familiarity with the language less dependent on family inputs received in bilingual households. Caring for the child is a parental investment that consists of interaction where the child becomes familiar with the language spoken by the carer (Bornstein et al., 1999; Lanza, 1992). The households consist of different age groups of parents; hence taking the ages of mother and father serve as a significant demographic control in the analysis.

Parental inputs include father resident, frequency of taking the child to the library, reading to the child, playing music with the child, and helping in painting from age 3 to 7.[27] These controls have a significant impact on the cognitive and non-cognitive development of the child. At the cohort's age of nine months, parental investment questions were asked, such as the importance of cuddling, stimulation, talking and regularity of contact with the parents (National Centre for Social Research, 2003). Following Cobb-Clark et al. (2018), 9-month questions were used as controls for the analysis of outcome at age 3.[28] Bono et al. (2016) uses a wide range of parental time inputs to monitor child development, but the current study uses only specific parental

[26] I have derived Young migrant indicator by combining the year of arrival to the UK and the year of birth. Following Cortina & Taran (2014), between 15-24 age group is a young migrant.

[27] I have transformed the frequency of parental time investment variables. The higher frequency is coded 1, and the remaining categories are combined to form a base. The categories are combined depending on the relative percentages; those having closer ratios are combined.

[28] Parental time investment questions were asked to both main and partner respondents, but only to the main respondent at age 3 years. In case where both main and partner respond, the category corresponding to higher frequency is taken as parental time investment. Table A.3 in the appendix shows the descriptives of 9 months child development and parental inputs used as lags in the age-specific model for Age 3 outcome.

controls that have a significant impact on the verbal and non-cognitive scores. Mother's age at the birth of child addresses the difference in the development outcomes as a result of children born to younger mothers. Parental education significantly impacts the frequency of speaking in English with the child.[29] But, this effect is mediated by the language of education. Parental higher education in English is associated with more English input. Whereas parents with a higher level of L2 education correspond to more L2 input to their child (Duursma et al., 2007).

Household-level characteristics, such as the number of household members present and the correspondence of geographical location, such as government office regions, are also considered in the analysis to serve as additional controls.

2.4 Method

The current analysis uses the child production function approach introduced by Todd & Wolpin (2003, 2007).[30] The child production function is widely used in educational research. The function evaluates a child's current performance based on the lagged period parental and child inputs, which provides a child's performance trajectory across the school years.

In the context of the current longitudinal study, the child has an innate ability that is assumed to be constant for each age group. For the initial analysis of the impact of inputs on the child outcome, the present study uses an OLS estimator. The following equation gives an estimate of these inputs:

$$Y_{it} = \alpha + \beta P_{it} + \eta H_{it} + \phi_i + \mu_{i,t}. \tag{2.1}$$

[29] The parental academic education is classified as a higher and lower degree. A higher degree corresponds to A/AS level and above. The GCSE and below it are classified as a low degree.

[30] Todd & Wolpin (2007) has found a significant impact of previous period child endowments, parental inputs, household characteristics on the cognitive performance at the school-going age.

Equation (2.1), Y_{it}, corresponds to the child 'i' vocabulary and behavioural outcomes at age t. The component P_{it} indicates the person speaking in non-English. P_{it} is a categorical variable having categories such as "mother speaking non-English" and "father speaking non-English", having a base category of "English only" households (coded as 0).[31] H_{it} is a collection of child-specific inputs (child's gender, ethnicity, birth weight) and current household and parental inputs: OECD weighted income, the number of household members, a parent caring for the child, mother's and father's education, parents born outside the UK, year of arrival, frequency of taking the child to the library, reading to the child, helping in music and painting and addressing the geographical location impact by using the government office regions (GOR) variable. Out of the parental inputs, some of these are time-variant such as frequency of taking child to library, reading to child, playing music and painting with child. ϕ_i captures the fixed effects addressing the changing parental identity speaking non-English through ages 3 to 7. The tables A.4 & A.5 show the household shifts from speaking non-English to English from ages 3 to 7. As the analysis is on the child outcomes, and equation (2.1) consists of unobserved (innate) child endowments such as self awareness, willingness to do things and conscientiousness as well as the presence of some parental inputs such as the decision of bilingual parents whether to speak in English at home, attachment to home country language, proficiency of English, providing environment for child so as to adapt to English at an early age. These factors have confounding effects on the child outcomes which affect the present outcomes of the child.[32] The MCS data is rich in child-level variables accounting for child development for each age group in our analysis. But in the analysis of child outcomes where children differ in ethnicity, parental heterogeneity, parental inputs (comprising of bilingual parental decision to speak in seconnd language), household inputs, information on child ability and attitude have a constant impact on the child outcome at each age group. These measures are not available in the MCS data. Hence, the coefficients of an OLS estimator in

[31]The table A.10 shows the detailed description of the variable "Who speaks non-English at home."
[32]The MCS data lacks information on these factors.

equation (2.1) give biased estimates due to the omitted variable bias caused by not accounting for the above-unobserved factors of child development. ϕ_i is included in the error term μ_{it} and the error term is not independent of these fixed effects caused by the unobserved child characteristics. This gives rise to endogeneity where the following inequality fits.

$$\text{Cov}((P_{it},H_{it}),\phi_i \neq 0)$$

Therefore, partial effects of the person speaking non-English and household inputs cannot have an unbiased impact and the error term μ_{it} is termed idiosyncratic.[33] Equation (2.1) cannot be estimated until the following condition holds true.

$$\text{E}((P_{it},H_{it}) \mid \mu_{it} + \phi_i) \neq 0.$$

Where the combination of P_{it} and H_{it} denotes the household bilingual environment, parental and child-specific inputs independent of the child's inherent skills. The above 'expectations' inequality condition is valid because parental bilingual decisions are influenced by parental level controls and the child's inherited abilities are absent in the current data. To account for time-variant parental and child input factors violating the strict exogeneity assumption, the fixed effects at the child level could give valid unbiased estimates.

Therefore, to check the validity of the fixed effects estimate, the current analysis uses the concept by Wooldridge (2010), also adopted by Cobb-Clark et al. (2018), where the "leads" of the variable of interest are added with the current variable to check for strict exogeneity. The test suggests there are time-variant factors that violate the strict

[33] The measurement error is less likely to have an impact as the variable asked through the survey consisted of categories of responses (Mother/Mother like Father/Father like, cohort member-only, other household members). The survey is a detailed collection of information about households which the interviewer asked. Moreover, the sample means were approximately the same throughout each sweep. This gives an insight that measurement error is less likely to have a significant impact on child outcomes.

exogeneity assumption.[34] Therefore, I utilise within-group estimators and use fixed effects to find a significant impact for both the vocabulary and behavioural scores.

Hence, the current study uses the value-added model for the estimation of child outcomes.[35] This approach is also adopted by Bono et al. (2016), Cobb-Clark et al. (2018) & Fiorini & Keane (2014). The analysis uses the approach of Bono et al. (2016) to include lagged parental and household inputs to measure the current period child ability. The value-added model uses the time invariant child and parental specific inputs as well as the combination of time-variant contemporaneous and lagged parental and household inputs. The combination of contemporaneous and the lagged inputs (parental, household and child specific) help to remove the endogeneity. The model assumes that individual parental investments decline at the same rate as other child-level and household inputs. Also, following Harris et al. (2010), it is expected that the contemporaneous parental investment will be some constant time the previous period input and will disappear once accounted for current time child outcome. Therefore, it is viable to include lagged parental and household inputs, given the child endowment and parental input decline at a geometric rate.

The age-specific value added specification observes the impact of contemporaneous and lagged child and parental inputs into current development scores. This considers the previous age inputs impact on the current age development outcomes. For example, a one standard deviation increase in the child's lag development scores on the current age outcomes. The fixed effects analysis addresses the impact of variables of interest by addressing the unobserved child endowments, and household and parental inputs.In the current analysis, the FE estimates are not consistent due to the failure

[34]The strict exogeneity test is conducted to find out whether the variables used to account fixed effects are independent of the explanatory and variables of interest, I use F statistic test. The test outcome for vocabulary scores and total difficulties score are: F=0.43 with p=0.66 and F=9.23 with p=0.008. The F statistic obtained is greater than the F critical value in the case of behavioural scores; hence the strict exogeneity hypothesis is rejected. Therefore, the Fixed effects estimates cannot be unbiased, so the strict exogeneity null hypothesis is rejected.

[35]The term value added is widely used in macroeconomics and refers to the contribution of the factors of production to increase the value of the product. The model resembles this by observing the current child performance based on current and previous inputs and outputs to child performance including lagged child outcome measures.

of the exogeneity assumption. Hence, the study relies on the cumulative value-added approach to estimate unbiased estimates of child outcomes.[36] The CVA estimates show the average impact of the coefficients by using one period lag of the variables. The study adopts the CVA analysis because of the biased partial effects of the individual parental bilingualism which arise due to their strong joint significance. Therefore, to make the current study precise and minimise bias, I adopt the CVA model. Hence, the equation (2.1) is transformed to the following equation as per the model specification:

$$Y_{it} = \alpha + \beta P_{it} + \psi P_{it-1} + \eta H_{it} + \gamma H_{it-1} + \lambda Y_{it-1} + \varepsilon_{it}. \tag{2.2}$$

where P_{it-1} denotes lag period parental bilingualism. λ is a indicator of persistence in the development scores across the time periods. Y_{it-1} addresses time constant child specific heterogeneity arising due to family background, child endowments, etc. on current period child's development. Also, H_{it-1} is a collection of child level (including lag vocabulary and behaviour scores), parental and household-level characteristics for the last time (sweep) 't-1' of child 'i'. The error term ε_{it} captures the rate of development of a child's vocabulary and behaviour. For equation (2.2) to give unbiased estimates, the rate of acquiring current period vocabulary and behavioural score should be independent of the previous period outcome score as well as the confounding parental inputs. The study aims to estimate the contemporaneous and lagged effect of individual parents' language use on child outcomes. In addition, to ensure precision, I have included only those respondents who cared for the child most of the time and also

[36]The GMM (Gemometric Methods of Moments) estimation to account for the unbiased estimates due to omitted variable bias was also used. Since, the T=3 (where 'T' corresponds to the number of sweeps) in the current study, therefore, the Arellano Bond test of serial correlation yield no outcome. Hence, the validity of GMM estimates could not be known. I also used simple GMM-IV estimates, but the test is not valid since there should be minimum of four waves for the results to be valid. But, in the current GMM-IV validity test, the second-period difference gave no observations. Hence, the GMM can't be used for checking for unbiased estimates in the analysis.

responded to the survey. Finally, for uniformity, I have restricted the parental relations present for both monolingual and bilingual children for each age.[37]

2.5 Results

2.5.1 Estimating parental impact on Vocabulary and Behavioural Development by child age.

The analysis begins with estimating child outcomes for each age group (3 to 7); this comprises lagged household and parental inputs acting as controls to the current period child outcome. Tables 2.3 and 2.4 show the age-specific estimates of child vocabulary and behavioural outcomes.

Vocabulary scores

Age-specific child vocabulary scores are shown in columns 1, 2 and 3 of table 2.3. Columns(1) to (3) of Table 2.3 show the effects of age-specific lagged parental bilingual status on vocabulary score.

Other things being equal, in table 2.3 at age 3 years, the mother speaking in non-English has 0.375 standard deviations marginal significant impact of bilingual penalty on the vocabulary score and father speaking in non-English has insignificant bilingual penalty of 0.13 standard deviations. This is interesting to know because the literature (De Houwer, 2007, 2015) suggests that father has a significant positive impact on child's vocabulary in two-parent and single-parent households at child's average age of 3. The significant decline in the vocabulary score at age 3 due to maternal bilingualism is a result of the majority of time devoted by the mother talking to the child in non-English involving correct usage of words and pronunciation of the

[37]The sample does not necessarily include the same parents (identified through person number and relation to the child in the MCS data) throughout age 3 to 7 years. But, those observations are included where the parents (not necessarily the same) responded to all the questions selected for the analysis. Also, at each sweep, some parental relations are present for monolinguals but not for bilinguals. Therefore, the parental relations which are present for both monolinguals and bilinguals are considered and the rest are ignored.

TABLE 2.3: The Effect of a person speaking in non-English on Vocabulary scores.

	Value added models			Indv.FE	Cum.val.add (CVA)
	Age 3	Age 5	Age 7		
	(1)	(2)	(3)	(4)	(5)
Z Vocabulary Score					
Person Speaking (n) (Base=English only)					
Mother	-0.375*	-0.193***	0.205**	-0.094	-0.029
	(0.136)	(0.029)	(0.085)	(0.128)	(0.073)
Father	-0.130	0.056	-0.032	0.011	0.003
	(0.103)	(0.098)	(0.081)	(0.127)	(0.082)
Person Speaking (n-1) (Base=English only)					
Mother		0.011	0.058		-0.091
		(0.045)	(0.113)		(0.073)
Father		-0.160	-0.006		-0.127
		(0.104)	(0.116)		(0.081)
Z Vocabulary score (n-1)		0.419***	0.266***		0.283***
		(0.011)	(0.020)		(0.009)
Z SDQ score (n-1)		-0.035***	-0.129***		-0.079***
		(0.010)	(0.010)		(0.008)
Child's age in months	0.016***	-0.032***	-0.026***	-0.013***	-0.021***
	(0.002)	(0.004)	(0.005)	(0.005)	(0.003)
Nonwhite (Base=White)	-0.083**	0.017	0.123*	-	0.146***
	(0.033)	(0.068)	(0.058)	-	(0.041)
Constant	-1.015	1.127***	1.784***	0.026	1.308***
	(0.622)	(0.147)	(0.234)	(1.136)	(0.274)
R-squared	0.173	0.288	0.211	0.005	0.136
N	7,765	7,243	6,784		
Observations				13,260	13,260

Notes.- Robust s.e in parenthesis are clustered at family level (stratum within each countries of the UK) for value added and at individual level for cumulative value added (CVA). The value added and CVA regressions are weighted at overall population level to avoid a non-response rate. Outcome is standardised Z vocabulary score with mean 0 and standard deviation 1 for age specific and CVA models. The analysis controls for child's age in months, ethnicity, gender, birth-weight, young migrant, natural father resident, weighted household income, mother's age at the birth of child, mother's and father's highest educational qualification, mother/father born in the UK, government office regions, parent caring most of the time, parental time investments in the form of frequency of reading to child, taking child to library, playing music and painting with child at each age group. The individual fixed effects (FE) model is estimated using unbalanced panel of child wave observations. The CVA model includes sweep dummies. Lags are added for all time variant controls in case of VA model and of selected time variant controls for CVA. Individual FE also includes lead of main variable of interest to check for exogeneity. "-" shows omitted because of collinearity.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

respective language resulting in a 3-year-old child finding this confusing. In a particular setting of bilingual households, the usage of two languages often leads to a state of confusion between the words used (Byers-Heinlein & Lew-Williams, 2013; Guiberson, 2013), this is captured by the vocabulary test scores for the word recognition exercises to test the knowledge of the English vocabulary in the MCS dataset. To further explain the bilingual penalty, it is important to highlight that the marginally significant maternal bilingual penalty is due to the fact that the children at age 3 are capable of keeping up with the languages spoken to them, therefore, in a two parent household even if the mother speaks in the second language, the children can still fare well in vocabulary scores.

The child develops language maturity by age 5, by having experience of speaking and hearing only one or more than one language from the birth as shown in column(2) of Table 2.3, at age 5, after controlling for the observable characteristics, the penalty of mother being bilingual is reduced to 0.19 of a standard deviation but has become more significantly different from zero. Also, the age 3 mother-only bilingual penalty inflicts on age 5 vocabulary score but with insignificant impact of 0.01 of a standard deviation. The age 5 maternal bilingual penalty could be explained due to marked perfection of monolingual children in English by age 5, and also due to the presence of some degree of mother tongue attachment in the bilingual parents that want their child to speak more in non-English before entering year 3 at age 7. Whereas, there is an insignificant age 5 paternal bilingual penalty of 0.06 standard deviations, and also there exists an insignificant age 3 paternal bilingual penalty of 0.16 of a standard deviation on the age 5 vocabulary score. This could be explained by considering that by age 5, the lagged paternal bilingual penalty does not inflicts a significant impact because of the increase in the time spent reading to the child in English than at age 3, declining the bilingual penalty. This is validated by the insignificant negative impact of paternal bilingualism when reading to the child is not included in the analysis(the results are provided on request). Hence, at age 5, the increase in the time spent by fathers reading to the child in English can promote vocabulary development. On average, there is a decrease in the bilingual penalty from age 3, this is accounted by the fact that the child has developed a maturity of the two languages which are constantly being spoken from the birth, to the child, and can easily differentiate the wordings of the two languages. This trend may arise because children at age five years started going to foundation schools, preschools, kindergartens and started gaining exposure to English (Hammer et al., 2009).[38] This decline in vocabulary scores between ages 3 and 5 years is consistent with the findings of Cobb-Clark et al. (2018) using MCS data & Clifton-Sprigg (2016) using the Scottish Government data. Following Cobb-Clark et al. (2018),Cunha & Heckman (2007) and Dickerson & Popli (2016), lagged vocabulary development has a significant positive impact on later age vocabulary development. At age 5, a one standard deviation increase in age 3 vocabulary score increases the age 5 vocabulary score by 0.42 of a standard deviation. But, the age 5 a one standard deviation increase in age 5 vocabulary score increases age 7 vocabulary score by 0.27 s.d. Also there is a decrease in the vocabulary score by 0.04 standard deviation for age 5 and 0.13 s.d. for age 7, which corresponds to higher behavioural problems.

The age 5 maternal bilingual penalty turns out to be an advantage over the monolingual peers on age 7 vocabulary scores with an insignificant impact of 0.06 standard deviation as shown in column(3) of table 2.3. Also, there is a decrease in the age 5 paternal bilingual penalty of 0.01 standard deviations on the age 7 vocabulary scores. The age 7 parental bilingual status inflicts an insignificant advantage on the vocabulary score with 0.21 standard deviations in case of maternal only and insignificant and negative impact of 0.03 standard deviations in case of paternal bilingual over the monolingual peers. The positive trend in the age 7 vocabulary scores are explained by the fact that by age 7 in a mother-only bilingual household, the child catches up with the monolingual score by increasing their reading time, or gradually becomes mature in vocabulary to differ between the words used in the first and the English language. Hence, in a two-parent family where mother speaks non-English, the impact of language confusion by the child is minimised by the increased maturity of the two languages when the child starts going to school and starts interacting with their monolingual peers. There remains a persistent significant positive impact of lagged vocabulary development on the current age 7 vocabulary score, but with a diminished

[38]In the UK, children start going to foundation school at age 3 years, by age 5 years they have developed a prior knowledge of English. Moreover, the ages in months are control for in the analysis. The average age is considered for the analysis. Therefore, the current findings in the UK correspond to the findings in the US by Hammer et al. (2009).

impact compared to age 5, whereas, there is an increasing negative impact of lagged behavioural development on the age 7 vocabulary score.

The age-specific findings are summarised in columns(4) and (5) of table 2.3 for vocabulary scores. [39] The FE estimates in the current analysis attempt to show the individual parental bilingualism after addressing the possible unobserved factors and observing the average impact on the vocabulary scores. The fixed effects model shown in column(4) is not independent of exogeneity with some controls omitted due to collinearity and I assumed them to be inconsistent. After controlling for the year FE, on average, the individual fixed effects coefficients reveal that maternal and paternal bilingualism has no significant impact on vocabulary scores when holding other factors constant. The cumulative model reported in column(5) shows the average impact of age-specific coefficients presented in columns(1) to (3). The one-time period-only lag is included since the other lagged coefficients were not consistent. There exists a biased individual parental bilingualism impact due to the strong joint significance of the bilingual status, hence I adopt the cumulative approach by using one-time-period lag. The maternal bilingual coefficients are not statistically different from zero and does not inflict an overall maternal bilingual penalty, the same is true for the father being bilingual. Whereas, there exists no impact of current and lag paternal and lag maternal bilingual penalty. Lag vocabulary and SDQ scores play a significant impact on the vocabulary scores. A one standard deviation increase in the lagged vocabulary score increases the current vocabulary score by 0.28 standard deviations and a one standard deviation increase in the lagged SDQ score reduce the vocabulary score by 8%-a one standard deviation increase in the lagged behavioural score leads to 0.08 of a standard deviations reduction in the vocabulary score.

[39]The FE estimation is done on the pooled data, both contemporaneous and lagged inputs (parent and household specific) are used in the analysis. In the FE estimation the time dummies for sweeps 3 and 4 are omitted because of collinearity as there are few observations of the inputs in sweeps 3 and 4.

Non-Cognitive scores

Age-specific child behavioural scores are shown in columns 1, 2 and 3 of table 2.4. Columns(1) to (3) of Table 2.4 show the effects of age-specific lagged parental bilingual status on behavioural score.

Other things being equal, in table 2.4 at age 3 years, there is a positive marginal significant impact of 0.05 standard deviation of maternal bilingualism and significant impact of 0.13 standard deviation of paternal bilingualism on the behavioural problems. The increase in the behavioural issues in bilingual households compared to monolingual counterparts at age 3, is explained by the existence of dual language environment in the former household. In bilingual households, apart from the resident country language the other languages are also spoken to the child. In some cases, bilingual parents expect their children to speak in resident country language, but the relatives of the children such as grandparents who lack knowledge of the resident country language expect their grandchildren to speak in the native language to them. This can create confusion in the child's knowledge of language and hence, stress the child when they are unable to communicate in the native language with their relatives, potentially giving rise to behavioural issues in bilingual children.

At age 5, the age 3 paternal bilingualism leads to an insignificant decline in behavioural conduct of the child. Whereas, there is no impact of lagged and current maternal bilingualism on the child's behavioural conduct. Also, the age 5 paternal bilingualism inflicts an insignificant increase of 0.15 of a standard deviation on behavioural issues. The lagged behavioural scores have a positive significant impact of 0.49 standard deviation on the current age scores showing that age 3 behavioural problems increases age 5 behavioural issues but there exists a significant negative impact of 0.06 standard deviation of lagged vocabulary score on the behavioural problems. Hence, higher age 3 vocabulary scores will lead to lower age 5 behavioural issues.

TABLE 2.4: The Effect of a person speaking in non-English on SDQ scores.

	Value added models			Indv.FE	Cum.val.add (CVA)
	Age 3 (1)	Age 5 (2)	Age 7 (3)	(4)	(5)
Z SDQ Score					
Person Speaking (n) (Base=English only)					
Mother	0.050*	-0.011	0.263***	0.144	0.092
	(0.026)	(0.139)	(0.073)	(0.103)	(0.062)
Father	0.129**	0.152	-0.111***	0.074	0.046
	(0.049)	(0.163)	(0.024)	(0.104)	(0.069)
Person Speaking (n-1) (Base=English only)					
Mother		-0.079	-0.207***		-0.028
		(0.099)	(0.049)		(0.059)
Father		-0.072	-0.030		-0.082
		(0.097)	(0.062)		(0.067)
Z Vocabulary score (n-1)		-0.065***	-0.025***		-0.039***
		(0.005)	(0.004)		(0.008)
Z SDQ score (n-1)		0.494***	0.645***		0.401***
		(0.008)	(0.020)		(0.010)
Child's age in months	-0.011	-0.020**	-0.005*	-0.013***	-0.013***
	(0.006)	(0.006)	(0.002)	(0.004)	(0.003)
Nonwhite (Base=White)	0.021	0.068	-0.016	-	0.013
	(0.037)	(0.050)	(0.034)	-	(0.041)
Constant	1.522**	1.991***	0.887***	1.364	2.101***
	(0.551)	(0.395)	(0.260)	(1.350)	(0.287)
R-squared	0.126	0.326	0.440	0.010	0.123
N	7,765	7,243	6,784		
Observations				13,260	13,260

Notes.- Robust s.e in parenthesis are clustered at family level (stratum within each countries of the UK) for value added and individual level for CVA. The VA and CVA regressions are weighted at overall population level to avoid non-response rate. Outcome is standardised Z SDQ score with mean 0 and standard deviation 1 for age specific and CVA models. The analysis controls for a child's age in months, ethnicity, gender, birth-weight, young migrant, natural father resident, weighted household income, mother's age at the birth of child, mother's and father's highest educational qualification, mother/father born in the UK, government office regions, parent caring most of the time, parental time investments in the form of frequency of reading to child, taking child to library, playing music and painting with child at each age group. The individual FE model is estimated using unbalanced panel of child wave observations. The CVA model includes sweep dummies. Lags are added for all time variant controls in case of VA model and of selected time variant controls for CVA. Individual FE also includes lead of main variable of interest to check for exogeneity. "-" shows omitted because of collinearity.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

In contrast to age 5, paternal bilingualism on the behavioural issues is observed in age 7. But, there exists a significant negative impact of lagged maternal bilingualism on the behavioural problems at age 7. The sudden increase in behavioural scores due to maternal bilingualism is accounted for by the increase in the interaction of the mother with the child where the child having gained the maturity of both languages becomes more susceptible to confusion due to increased maternal bilingualism. Also, there is a decrease in the impact of lagged vocabulary scores and an increase in the impact of lagged behavioural scores on age 7 behavioural problems. Hence, similar to the lagged development impact seen in table2.3, higher behavioural problems at age 5 lead to higher age 7 behavioural issues but higher age 5 vocabulary development lead to smaller decrease in behavioural issues at age 7. The current study findings on behavioural development correspond to that of Cobb-Clark et al. (2018) overall in terms of explaining the transitions in the behavioural issues as a result of individual bilingual maternal and paternal roles in inculcating adaptability and peer contact in the child by creating harmonious language environment at home by introducing age-specific individual parental impact.

Similar to Table 2.3, the age-specific findings are summarised in columns(4) and (5) of table 2.4 for behavioural scores.[40] The FE estimates in the current analysis attempt to show that the individual parental bilingualism after addressing the possible unobserved factors and observing the average impact on the vocabulary scores. The fixed effects model shown in column(4) is not independent of the exogeneity with some controls omitted due to collinearity and I assumed them to be inconsistent. After controlling for the year FE, on average, the individual fixed effects coefficients reveal that maternal and paternal bilingualism has no significant impact on behavioural scores when holding other factors constant. The cumulative model reported in column(5) shows the average impact of age-specific coefficients presented in columns(1) to (3). The one-time period-only lag is included since the other lagged coefficients were not consistent. There exists a biased individual parental bilingualism impact due to the strong joint significance of the bilingual status, hence I adopt the cumulative approach by using a one-time-period lag. Similar to the cumulative impact on vocabulary score in table 2.3 there is no impact of paternal and maternal bilingualism, the paternal bilingual coefficients are not statistically different from zero and do not inflict

[40]The FE estimation is done on the pooled data, both contemporaneous and lagged inputs (parent and household-specific) are used in the analysis. In the FE estimation, the time dummies for sweeps 3 and 4 are omitted because of collinearity as there are few observations of the inputs in sweeps 3 and 4.

an overall maternal and paternal bilingual penalty on the behavioural scores. Lag vocabulary and SDQ score play a significant impact on the vocabulary scores. A one standard deviation increase in the lagged vocabulary score decreases the current behavioural score by 0.04 standard deviations and a one standard deviation increase in the lagged SDQ score increases the behavioural score by approximately 40%-a one standard deviation increase in the lagged behavioural score leads to 0.40 of a standard deviations increase in the current behavioural score(more behavioural issues).[41]

2.5.2 Estimating parental impact on Vocabulary and Behavioural Development by child age without ethnicity.

In the analysis, there are approximately 10% of ethnic minority families. Among these ethnic minority families, there are approximately 67% of nonwhite households that are bilingual. To analyse the impact of ethnicity and the bilingualism on the development scores, I evaluate the vocabulary and behavioural scores without ethnicity in tables 2.5 and 2.6 and compare it with the analysis found in tables **??** and 2.4 including the ethnicity variable.

Vocabulary scores

Table 2.5 shows the impact of parental bilingualism without controlling for ethnicity on the z vocabulary development throughout ages 3-7 years. The analysis differs from table 2.3 by excluding ethnicity variable. At age 3, the mother being bilingual penalty is -0.388 standard deviations which accounts for an increase in the negative impact by 0.07 s.d. of vocabulary score when excluding ethnicity. Similarly, there is an increase of father being bilingual penalty by 0.03 s.d. for vocabulary score. Hence, ethnicity plays an important role in determining the vocabulary development at age 3 years.

[41] A set of full controls along with the variables of interest is shown in tables B.1 & B.2 for vocabulary score and tables B.3 & B.4 for the behavioural scores of the same model.

TABLE 2.5: The Effect of a person speaking in non-English on Vocabulary scores without ethnicity.

	Value added models			Indv.FE	Cum.val.add (CVA)
	Age 3 (1)	Age 5 (2)	Age 7 (3)	(4)	(5)
Z Vocabulary Score					
Person Speaking (n) (Base=English only)					
Mother	-0.388**	-0.191***	0.226**	-0.094	-0.016
	(0.133)	(0.029)	(0.087)	(0.128)	(0.074)
Father	-0.162	0.057	-0.009	0.011	0.015
	(0.107)	(0.092)	(0.077)	(0.127)	(0.083)
Person Speaking (n-1) (Base=English only)					
Mother		0.012	0.059		-0.082
		(0.040)	(0.116)		(0.072)
Father		-0.154	0.022		-0.094
		(0.119)	(0.127)		(0.080)
Z Vocabulary score (n-1)		0.420***	0.265***		0.281***
		(0.011)	(0.020)		(0.009)
Z SDQ score (n-1)		-0.035***	-0.129***		-0.078***
		(0.009)	(0.010)		(0.008)
Child's age in months	0.016***	-0.032***	-0.026***	-0.013***	-0.021***
	(0.002)	(0.004)	(0.005)	(0.005)	(0.003)
Constant	-1.020	1.112***	1.754***	0.026	1.268***
	(0.620)	(0.147)	(0.241)	(1.136)	(0.276)
R-squared	0.172	0.288	0.211	0.005	0.134
N	7,765	7,243	6,784		
Observations				13,260	13,260

Notes.- Robust s.e in parenthesis are clustered at family level (stratum within each countries of the UK) for value added and at individual level for cumulative value added (CVA). The value added and CVA regressions are weighted at overall population level to avoid a non-response rate. Outcome is standardised Z vocabulary score with mean 0 and standard deviation 1 for age specific and CVA models. The analysis controls for child's age in months, gender, birth-weight, young migrant, natural father resident, weighted household income, mother's age at the birth of child, mother's age at the birth of child, mother's and father's highest educational qualification, mother/father born in the UK, government office regions, parent caring most of the time, parental time investments in the form of frequency of reading to child, taking child to library, playing music and painting with child at each age group. The individual fixed effects (FE) model is estimated using unbalanced panel of child wave observations. The CVA model includes sweep dummies. Lags are added for all time variant controls in case of value added model and of selected time variant controls for CVA. Individual FE also includes lead of main variable of interest to check for exogeneity. "-" shows omitted because of collinearity.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

In column(2) of table 2.3 and 2.5, there is no difference in the significant penalty at age 5 for a mother speaking non-English with and without ethnicity. Also, there is an insignificant increase in the age 3 paternal being bilingual penalty of 0.15 standard deviation without controlling for ethnicity than 0.16 standard deviation bilingual penalty when controlling ethnicity. Therefore, ethnicity has no significant role in estimating the age 5 vocabulary score due to age 3 bilingual father.

There exists no impact of contempporaneous paternal bilingualism on the age 7 vocabulary scores with and without controlling for ethnicity. There is a similar individual parental bilingual impact on the age 7 vocabulary scores. Therefore, by age 7, the vocabulary development is solely due to maternal bilingualism and its impact remains significant and not much influenced by ethnicity.

By comparing column(5) of Tables 2.3 and 2.5 there is no difference of the individual parental bilingualism on the vocabulary scores when controlling and not controlling for ethnicity. Hence, there are no differences in the z vocabulary score due to ethnic coefficients on the bilingual penalties in the cumulative analysis as shown in column(5) of table 2.3 and table 2.5. Therefore, the decrease in the vocabulary scores is solely due to maternal and paternal bilingualism, but remains insignificant.

Non-Cognitive scores

Table 2.6 shows the impact of parental bilingualism on the z SDQ development throughout ages 3-7 years by excluding ethnicity variable. In table 2.6, at age 3, the mother being bilingual penalty is 0.053 standard deviations increase in the behavioural problems. As compared to maternal penalty in table 2.4, there is a decrease in the behavioural impact by 0.003 s.d. when excluding ethnicity. Similarly, there is an increase in father being bilingual penalty by 0.008 s.d. in behavioural score. Therefore, ethnicity plays an important role in the behavioural issues arising at age 3.

In column(2) of table 2.4, the decrease in behavioural problems at age 5 for a mother speaking non-English at age 3 years is 0.08 standard deviation when controlling ethnicity which is slightly higher than 0.07 standard deviation when excluding ethnicity, remaining insignificant. Similarly, there is a marginal decrease of behavioural problems for father speaking non-English at age 3 years by approximately 0.07 standard deviations when controlling for ethnicity and an insignificant decrease of 0.05 s.d. when not controlling for ethnicity. Also, there remains an insignificant positive impact of age 5 paternal bilingualism on the behavioural scores accounting for 0.152 standard deviation when including ethnicity and 0.160 s.d. when excluding ethnicity. Therefore, ethnicity has an insignificant impact on the age 5 behavioural issues which can be observed by a decline in paternal bilingual penalty after controlling for ethnicity.

TABLE 2.6: The Effect of a person speaking in non-English on SDQ scores without ethnicity.

	Value added models			Indv.FE	Cum.val.add (CVA)
	Age 3 (1)	Age 5 (2)	Age 7 (3)	(4)	(5)
Z SDQ Score					
Person Speaking (n) (Base=English only)					
Mother	0.053*	-0.011	0.254**	0.144	0.093
	(0.028)	(0.141)	(0.077)	(0.103)	(0.061)
Father	0.137**	0.160	-0.121***	0.0739	0.048
	(0.043)	(0.161)	(0.0227)	(0.104)	(0.069)
Person Speaking (n-1) (Base=English only)					
Mother		-0.068	-0.207***		-0.027
		(0.092)	(0.0494)		(0.060)
Father		-0.050	-0.042		-0.079
		(0.111)	(0.066)		(0.066)
Z Vocabulary score (n-1)		-0.066***	-0.025***		-0.039***
		(0.005)	(0.004)		(0.008)
Z SDQ score (n-1)		0.494***	0.646***		0.401***
		(0.008)	(0.020)		(0.010)
Child's age in months	-0.011	-0.020**	-0.005*	-0.013***	-0.0132***
	(0.006)	(0.007)	(0.002)	(0.005)	(0.003)
Constant	1.523**	1.978***	0.900***	1.364	2.098***
	(0.550)	(0.409)	(0.265)	(1.350)	(0.287)
R-squared	0.126	0.326	0.440	0.010	0.123
N	7,765	7,243	6,784		
Observations				13,260	13,260

Notes.- Robust s.e in parenthesis are clustered at family level (stratum within each countries of the UK) for value added and individual level for CVA. The value added and CVA regressions are weighted at overall population level to avoid non-response rate. Outcome is standardised Z SDQ score with mean 0 and standard deviation 1 for age specific and CVA models. The analysis controls for a child's age in months, gender, birth-weight, young migrant, natural father resident, weighted household income, mother's age at the birth of child, mother's and father's highest educational qualification, mother/father born in the UK, government office regions, parent caring most of the time, parental time investments in the form of frequency of reading to child, taking child to library, playing music and painting with child at each age group. The individual FE model is estimated using unbalanced panel of child wave observations. The CVA model includes sweep dummies. Lags are added for all time variant controls in case of VA model and of selected time variant controls for CVA. Individual FE also includes lead of main variable of interest to check for exogeneity. "-" shows omitted because of collinearity.
*** Significant at the 1 percent level.** Significant at the 5 percent level.* Significant at the 10 percent level.

There is an increase in the significance of maternal bilingualism when controlling ethnicity and a decrease in the magnitude of paternal bilingualism negative impact on the behavioural score when controlling for ethnicity. Hence, ethnicity has an important role in determining the impact of maternal bilingualism impact on behavioural score, but vice versa for the case of paternal bilingualism. Also, the overall impact as shown by the cumulative impact in column(5) of tables 2.4 and 2.6 shows no difference in either the maternal-only or paternal-only bilingualism on the child's behavioural issues. Hence, there is no role of ethnicity in determining the behavioural outcomes which is consistent with age 7 bilingual impact on the behavioural scores.

Heterogeneity

There are significant differences in the composition of child and parent-specific inputs. Table 2.7 shows the heterogeneous estimates by the gender of the child and parental education (mother and father).[42] In column 1, of table 2.7 for the current period, there is no impact of parental contemporaneous bilingualism on the vocabulary scores om sons and daughters. But, there is a marginal significant negative impact of 0.20 sd of lagged maternal bilingualism on daughter's behavioural scores. This marginal negative effect is different to findings by Cobb-Clark et al. (2018). The finding is interesting as, Hammer et al. (2009) finds that daughters are not susceptible to maternal bilingualism as compared to sons where the mothers speak more non-English to their daughters, this can be seen opposite to the findings from table 2.7 that there exists marginal significant impact of current and lagged maternal bilingualism on the daughter's vocabulary score.

Regarding behavioural score as shown in columns (1) and (2) of table 2.8, there is no impact of contemporaneous paternal bilingualism on the behavioural scores of sons and daughter. But, lagged maternal bilingualism marginally improves behavioural conduct of daughtersand lagged paternal bilingualism improves behavioural conduct for sons.This is explained by the findings of Kaushanskaya et al. (2013) on children aged between 3 and 7. The findings show that in bilingual households, girls have

[42]There were other groups such as ethnicity, whether a family has a young migrant, the parent responsible for caring. But, the observations of the subgroups were not comparable because of the nominal sample size of nonwhite, young migrants compared to their base group. There were preliminary observations for the "Father" caring in the variable "parent responsible for caring most of the time".

stronger phonological and lexical knowledge of familiar pictures than boys because they have stronger retention. The lagged vocabulary score have significant positive impact on daughter's current period development scores compared to the sons, indicating an increase in behavioural issues due to lagged vocabulary scores. This is interesting and opposite to the findings by Kaushanskaya et al. (2013) and Sun (2019) showing that girls have better grammar skills and more retention than boys. Whereas, lagged SDQ scores increase behavioural issues in sons. Hence, the above findings highlight that in a bilingual household, daughters have less bilingual penalty due to lagged behavioural development in terms but higher penalty due to lagged vocabulary development of mother tongue language and second language compared to sons.

In tables 2.7 and 2.8, as shown in columns(3) and (4) there exists no significant impact of current and lagged parental bilingualism in higher degree and lower degree mother's household. Whereas, there is a marginal negative impact of approximately 0.17 standard deviations lagged maternal bilingualism on vocabulary and 0.15 sd of contemporaneous paternal bilingualism on behavioural development incase of higher-degree mother.

In case of lower degree father there is a significant negative impact of 0.24 standard deviation current period maternal bilingualism on the child's vocabulary score, whereas, a marginal decrease of 0.28 vocabulary score due to lagged paternal bilingualism. For the behavioural development there is no impact of current and lagged parental bilingualism on the child's behavioural scores. This is explained by Duursma et al. (2007)'s findings that the language of higher education is significantly related to the language input to the child. De Houwer (2007) & De Houwer (2015) suggest a significant impact of father's language input, whereas, in households where the father has a lower degree, the mother and both parents speaking non-English have a marginally significant and highly statistically negative impact on the vocabulary scores respectively. This may be because the negative impact of the confusion is more prominent in case of lower degree father household. Also, households where father has

TABLE 2.7: Heterogeneous effect of a person speaking in non-English for Z vocabulary scores.

	Child's Gender		Mother's Edu.		Father's Edu.	
	Male	Female	High Deg.	Oth. lower	High Deg.	Oth. lower
	(1)	(2)	(3)	(4)	(5)	(6)
Person speaking (n) (Base=Eng. only)						
Mother	-0.115	0.085	-0.001	0.029	0.128	-0.236**
	(0.106)	(0.111)	(0.094)	(0.124)	(0.104)	(0.115)
Father	0.056	-0.078	-0.104	0.053	-0.095	0.128
	(0.117)	(0.123)	(0.104)	(0.133)	(0.111)	(0.130)
Person speaking (n-1) (Base=Eng. only)						
Mother	-0.009	-0.199*	-0.166*	0.017	-0.250***	0.196
	(0.102)	(0.112)	(0.095)	(0.128)	((0.094)	(0.125)
Father	-0.182	-0.050	-0.084	-0.173	-0.037	-0.277**
	(0.118)	(0.122)	(0.109)	(0.134)	(0.106)	(0.137)
Z Voc. score (n-1)	0.270***	0.291***	0.263***	0.299***	0.283***	0.275***
	(0.013)	(0.013)	(0.014)	(0.013)	(0.014)	(0.012)
Z SDQ score (n-1)	-0.078***	0.000	-0.062***	-0.102***	-0.068***	-0.106***
	(0.012)	(0.000)	(0.012)	(0.011)	(0.013)	(0.012)
Child's age (months)	-0.015***	-0.026***	-0.016***	-0.017***	-0.015***	-0.012***
	(0.004)	(0.005)	(0.005)	(0.004)	(0.005)	(0.004)
Nonwhite (Base=White)	0.125**	0.135**	0.135**	0.063	0.084	0.200***
	(0.063)	(0.059)	(0.060)	(0.060)	(0.059)	(0.061)
Constant	0.631	1.598***	0.714*	1.164***	0.374	0.871**
	(0.392)	(0.400)	(0.426)	(0.379)	(0.484)	(0.364)
R-squared	0.079	0.125	0.103	0.085	0.084	0.088
Observations	6,641	6,619	6,421	6,839	5,710	7,550

Notes.- The analysis consists of unbalanced panel in CVA model shown in table 2.3. The robust standard errors are in parenthesis, clustered at individual level. The controls are same as used in tables 2.3 for CVA model. Outcome is standardised Z vocabulary score with mean 0 and standard deviation 1 for CVA model. The R –squared estimators are within. Lag SDQ score for female child omitted due to collinearity.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

lower degree, the children are exposed to fewer books, less reading habits and fewer language development activities and the father's low degree may negatively impact their child's vocabulary development. Hence, a father's lower degree inflicts significant bilingual penalty on the child's vocabulary scores as compared to the mother having lower degree. Therefore, current-period maternal bilingualism has a significant penalty on the vocabulary scores, whereas, there exists no impact of parental current and lagged bilingualism on behavioural issues in case of lower degree father. Lagged vocabulary development has greater magnitude of significant positive impact on the lower degree mother compared to higher degree mother on the vocabulary score, whereas, a higher positive impact on the higher degree father's household compared to lower degree father for the vocabulary development. But there is an increase in behavioural issues due to higher lagged SDQ scores in case of higher degree mother and father households compared to the lower degree parental households.

Therefore, in a bilingual household, the current study did not find any significant parental bilingualism impact on the child development outcomes in terms of child gender heterogeneity. But, there exists lagged negative marginal maternal and paternal bilingual penalty on behavioural scores but only maternal penalty on vocabulary score.

In table 2.8, there is no significant heterogeneous effect of contemporaneous individual parental bilingualism on behavioural scores but there is a marginal decrease of 0.16 standard deviation in the behavioural issues in boys as a result of lagged paternal bilingualism compared to girls. The decrease in girls' behavioural issues is accounted through lagged maternal bilingualism by 0.20 s.d. This could be explained by the fact that boys are more influenced by paternal bilingualism whereas, girls by maternal. Also in comparison to girls, boys are more susceptible to anxiety and confusion in the early age of learning language.

There exists no significant impact of parental bilingualism in case of maternal education on the behavioural scores. But, there exists a marginal significant increase in behavioural issues as a result of current paternal bilingualism for higher degree mothers. Hence, children having bilingual fathers in higher degree mother households, show more behavioural issues. Whereas, having bilingual mother in lower degree father households have more vocabulary decline. Therefore, the increase in behavioural issues in the case of higher-degree mothers could be due to the absence of maintaining a harmonious language environment at home where both the mother tongue language and English are spoken, even in the presence of learning environment. But, there is a vice-versa impact on lower-degree father households where mother speaking non-English declines behavioural scores in the absence of learning environment. Therefore, bilingual impact is more negatively significant incase of vocabulary development.

TABLE 2.8: Heterogeneous effect of a person speaking in non-English for Z SDQ scores.

	Child's Gender		Mother's Edu.		Father's Edu.	
	Male	Female	High Deg.	Oth. lower	High Deg.	Oth. lower
	(1)	(2)	(3)	(4)	(5)	(6)
Person speaking (n) (Base=Eng. only)						
Mother	0.071	0.082	-0.007	0.074	0.029	0.009
	(0.088)	(0.110)	(0.081)	(0.094)	(0.086)	(0.086)
Father	0.054	-0.076	0.150*	0.032	0.081	0.026
	(0.099)	(0.122)	(0.086)	(0.105)	(0.086)	(0.101)
Person speaking (n-1) (Base=Eng. only)						
Mother	0.053	-0.204*	0.062	-0.115	-0.085	0.058
	(0.079)	(0.112)	(0.077)	(0.098)	(0.082)	(0.092)
Father	-0.163*	-0.040	-0.111	-0.034	-0.051	-0.073
	(0.087)	(0.121)	(0.086)	(0.102)	(0.090)	(0.100)
Z Voc. score (n-1)	-0.033***	0.291***	-0.022**	-0.056***	-0.025**	-0.056***
	(0.011)	(0.013)	(0.011)	(0.011)	(0.012)	(0.010)
Z SDQ score (n-1)	0.385***	0.000	0.390***	0.385***	0.410***	0.397***
	(0.014)	(0.000)	(0.014)	(0.013)	(0.015)	(0.014)
Child's age (months)	-0.014***	-0.020***	-0.015***	-0.010**	-0.017***	-0.010**
	(0.005)	(0.006)	(0.004)	(0.004)	(0.004)	(0.004)
Nonwhite (Base=White)	0.0571	0.133**	0.018	-0.008	0.031	0.0185
	(0.058)	(0.059)	(0.056)	(0.058)	(0.053)	(0.061)
Constant	1.834***	2.112***	2.275***	1.637***	2.706***	1.637***
	(0.407)	(0.421)	(0.388)	(0.387)	(0.429)	(0.371)
R-squared	0.122	0.124	0.110	0.126	0.084	0.112
Observations	6,641	6,619	6,421	6,839	5,710	7,550

Notes.- The analysis consists of an unbalanced panel in CVA model shown in table 2.4. The robust standard errors are in parenthesis, clustered at individual level. The controls are same as used in table 2.4 for CVA model. Outcome is standardised Z SDQ score with mean 0 and standard deviation 1 for CVA model. The R –squared estimators are within.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

2.6 Discussion

The study sets out to assess the impact of parental decision to speak non-English at home on children's vocabulary and behavioural outcomes. Cobb-Clark et al. (2018) & Clifton-Sprigg (2016) have found, on the impact of bilingualism on child outcomes, that bilingual children fare worse at age three but gradually catch up with their monolingual counterparts. The current study found a significant age 5 maternal bilingual penalty with no impact of age 3 maternal and paternal bilingualism. This also holds true for age 7 vocabulary development. Over the time, maternal bilingualism shows an insignificant negative impact of 0.12(-0.091-0.029) standard deviations which is consistent with ages 3 and 5. Also, there exists significant positive impact of lagged vocabulary outcomes and a significant negative impact of lagged behavioural scores on the vocabulary development. In sum, the under-performance of bilingual children from age 3 to 5 is due to the significant negative maternal bilingualism and insignificant lagged age 3 paternal bilingualism, but, these children do catch up on an average by age 7.[43]

The contemporaneous paternal bilingualism leads to a significant increase in age 3 behavioural issues accompanied by a marginally significant positive impact of 0.1 standard deviation of lagged maternal bilingualism. But by age 7, there is an emergence of significant maternal bilingualism leading to an increase in behavioural issues by 0.26 standard deviations and a significant decline of 0.1 s.d. of paternal bilingualism. But, there is a significant decline of age 5 maternal bilingualism on the behavioural development. Also, there is an increase in the behavioural issues due to lagged behavioural scores and a decrease in the negative impact of lagged vocabulary scores between ages 5 to 7. Throughout ages 3 to 7, more child behavioural issues are observed in households where at least mother speaks non-English at current time. In contrast to the previous studies by Cobb-Clark et al. (2018) & Clifton-Sprigg (2016), the age-specific findings in the current study show positive significant impact of maternal, paternal bilingualism on the behavioural issues. In case of behavioural outcomes, there is a constant significant increase of contemporaneous maternal bilingualism on the behavioural issues at ages 3 and 7, but there is a sudden insignificant decline of vocabulary development due to maternal bilingualism at age 5. However, over time there is no signifciant impact of parental bilingualism on behavioural development. Also, there is a decline in behavioural issues due to increase

[43]The data does not provide the information on the proportion of time parents speak non-English.

in lagged vocabulary scores and an increase due to increase in lagged behavioural scores which are consistent with age 3 to 7 behavioural development.

The model specification comes from various past literature studies on the impact of bilingualism on child development. The current study utilises past findings and approaches by Bono et al. (2016), Cobb-Clark et al. (2018), Clifton-Sprigg (2016) & Todd & Wolpin (2007) to account for econometric challenge arising due to the unobserved child abilities as well as confounding parental inputs. I assume that the current analysis has a constant unobserved child and parental impact throughout ages 3 to 7 and uses the contemporaneous and lagged variables to account age specific development outcomes. The use of lagged development scores account for the cross-complementarity between verbal and SDQ development. The current and previous studies (e.g.Cobb-Clark et al., 2018; Clifton-Sprigg, 2016) highlight persistence endeavour to report unbiased results using fixed-effects approach, relying on exogeneity assumption. The strict exogeneity is partially fulfilled in the current study. The partial fulfilment is explained as there was no significant impact of current period test scores on the later periods decision of parents to speak non-English (for mother or father). However, the main variables (individual parental bilingualism) partial effects were biased because of the strong joint significance of the variable of interest. Therefore, to make the current study precise and minimise bias, I adopt the cumulative model approach. The cumulative model approach reduces biased estimates which arise due to joint significance. The contemporaneous estimates are observed by including the lagged main variable of interest. The inclusion of lagged main variable of interest has a significant joint estimate in the analysis of a child's outcomes (vocabulary and behaviour) at current age (age 5 and 7).[44]

The sample size for bilingual households is smaller than the monolingual overall, leading to a smaller sample size of the mother, father speaking non-English that

[44] I have done the joint significance estimation for the lag main variable of interest and found that the current parental descision to speak non-English depends upon the previous period child's outcomes. Tables A.4 and A.5 show the shift in individual parental home language input across the sweeps.

constitute the bilingual households. The insignificant observations in the variable of interest (excluding the base category) provide a reason for the preliminary observations in the second period lead (excluding the base category). Hence, I checked the exogeneity of the main variable of interest for only one period lead. Therefore, the insufficient bilingual sample poses a limitation to the findings of the study.

I have selected English as the base category because the children in the MCS study are born in the UK. There are few families not speaking English at all and most of the families in the sample speak English at home.

2.7 Conclusion

Few studies have analysed child development in the households where at least one parent is bilingual compared to the studies on comparing bilingualism across different nationalities or languages. The present research mainly contributes to the literature on child vocabulary and behavioural development by examining the heterogeneity of bilingual households. One of the important features of bilingual households is where more than one language is spoken to the child. As a result of exposure to more than one language, there lies an importance of the parental role to understand the dynamics within the bilingual households. The current study brings into light the contributions made by the previous literature in the field of psychology and linguistics that have highlighted the role of parents influencing the language spoken at home and its resultant impact on child's cognitive and behavioural performance. Hence, given the bilingual externality in the form of anxiety, problems in adjustment with peers, and confusion arising due to non-harmony in household language use are some of the factors that motivated the current research to evaluate the impact on children's behaviour. Apart from the bilingual externality in the form of child's behaviour, there lies an important role of the parental efforts to develop the vocabulary of the mother tongue and the second language. This can be seen in the form of a child's vocabulary development scores. Therefore, the present study is motivated to capture the vocabulary and behavioural development gaps arising in bilingual children compared to their monolingual counterparts.

By using the MCS dataset in the current study, the bilingual heterogeneity is seen in the form of a significant negative impact of bilingual parents on a child's vocabulary score at ages 3 to 5. However, this negative significance declines at age 7 with an increase in vocabulary score compared to monolingual counterparts but not significantly, indicating the catching up of bilingual children with their monolingual counterparts by Years 2 and 3. Hence, highlighting the importance of the parental efforts such as reading books or taking child to foundation classes on the vocabulary scores. The study highlights the individual maternal significant negative bilingual impact at age 5 but turns significant and positive at age 7 due to the catching up with monolingual peers. For the behavioural outcomes at age 7, there is an emergence of significant maternal bilingual penalty in the form of increase in behavioural issues. But, over time paternal bilingualism inflicts decrease in behavioural issues. Overall, for vocabulary scores, contemporary maternal bilingualism has persistent significant negative impact through which is consistent with age 3 to 5 vocabulary development and turns positive at age 7. The age 3 significant paternal bilingual penalty on the child's behaviour shows that the bilingual fathers should ensure harmonious use of both languages at home throughout at the early ages. But, over time there exists no impact of parental bilingualism on the behavioural development. It is always better for bilingual households to realise the importance of mother tongue and second language and therefore ensure their harmonious use as early as possible in order to minimise child behavioural issues that may surface later in the child's life.

The impact of gender differentials are slightly different from the previous literature studying the impact on child's vocabulary and behaviour conduct. The present findings highlight that girls have more marginally significant negative impact on the vocabulary scores than the boys for mother speaking in non-English. But, boys have less marginal significant decline in behavioural problems than girls due to lagged paternal bilingualism. Also, lagged behavioural scores decline vocabulary development of boys compared to girls. Hence, highlighting that boys are more susceptible to lower vocabulary development in their early life compared to the girls. There is a significant impact of cross-complementarity between the vocabulary and behavioural scores for boys and girls. For boys, there is a significant decline in the behavioural issues due to improvement in lagged vocabulary scores. Paternal education plays a significant role in vocabulary and behavioural development. Maternal bilingualism inflicts a significant bilingual penalty in lower degree father household for vocabulary score whereas, in higher degree father's household there remains a significant lagged maternal bilingual penalty. For the behavioural scores, there is no impact of lagged and contemporaneous parental bilingualism on behavioural issues. Therefore, highlighting the fact that there is a significant impact of father's education on the vocabulary test scores in terms of providing children with books and foundation classes in their early years which are only influenced by maternal bilingualism.

The most prominent finding to emerge from this study is that, over time, there exists no significant impact of bilingual mother and father on the child's vocabulary and behavioural development. Also, girls'/boys' vocabulary reduces if mother/father speaks non-English and decrease in behavioural issues due to lagged maternal/paternal bilingualism. Therefore, parental heterogeneity has a significant role for the vocabulary and behavioural development in terms of child's gender. In response to the bilingual household heterogeneity, the future studies could explore the dynamics of single-parent bilingual households on child outcomes. A small sample size of the father in bilingual families limits our estimates in the study. Including cultural indicators and detailed ethnicity [45] could improve the scope of the study. Also, extending the age group to above seven years when the child enters Key Stage 2 and above could give a more comprehensive knowledge of the impact of bilingual

[45]The study has attempted to find whether there exists any difference on the development outcomes due to paternal bilingualism. The corresponding results shown in Tables 2.5 and 2.6 for the development outcomes excluding parental ethnicity are compared to the analysis including ethnicity in Tables 2.3 and 2.4.

family heterogeneity as the child reaches higher school grades. This study also brings into light the government policies that need to be more focused on the bilingual fathers having lower degree consisting of mother speaking in non-English to the child and are unable to provide their child the necessary resources to achieve the vocabulary knowledge. Policies should focus in encouraging bilingual fathers to maintain harmonious environment in lower degree paternal household. Overall, bilingual parents should be provided with phonics and vocabulary training so as to inculcate the correct usage of language and transfer the vocabulary knowledge to their children. In this way, the above factors would ensure the harmony in language usage and will lead to reduction in vocabulary and behaviour gaps emerging in the early years of bilingual children.

Chapter 3

The Impact of Parental Residential Mobility Descisions on Child Development

3.1 Introduction

Child development is influenced by household and parental inputs. The household inputs are defined in terms of household resources and parental time investment, education and parental decisions (Beck et al., 2016; Bialystok et al., 2012; Del Boca et al., 2017; Bono et al., 2016; Cunha & Heckman, 2007). The present study focuses on the importance of parental decisions to change residence on child development. The geographical mobility of a household is affected by the neighbourhood, school availability, and housing and employment potential in a particular area. Other important factors influencing family mobility are finances and health. Therefore, if households change residence for one reason, this may in turn, affect other factors influencing child development later on in life.

Children's early years can be greatly impacted by parental decision to change residence, and the impact on children can vary from family to family. Some families move due to housing requirements in the early years of a child. As per the Ministry of Housing Communities & Local Government (2022), the 2001 UK census shows that the mobility due to housing accounts for around one in five families with a new birth i.e., indicating an increase in the demand for bigger house as the family size increases. The parental decisions to move to a new area can lead to changes in child's behaviour, as the family processes stressful events in adjusting to a new environment, and this adjustment is for both parents and children (Clair, 2019; Russell & Lister, 2020). In the UK, the Department for Work & Pension (2020) scheme data on Households Below Average Income, for the year 2019/20, reports 3.2 million children before housing costs rose and 4.3 million children after housing costs living in relative poverty.[1]

It is crucial to highlight the importance of early-life parental decisions acting as inputs to the child development in subsequent years, as highlighted by Beck et al. (2016). The early-life parental decisions have a profound impact on child development in terms of cognitive and non-cognitive skill development as they shape the later life outcomes of the child (Heckman, 2011). This study aims to explore the impact in child development from early life to the subsequent years as a result of these early-life parental decisions. Previous studies by Gambaro & Joshi (2016) & Gambaro et al. (2017) on parental mobility utilised the MCS dataset to study the families moving to improve housing when children were aged 5. The current study focuses on the quality of mobility which is captured by the change in the IMD deciles (Index of Multiple Deprivation) from the current residence IMD deciles. This is achieved by interacting the change in the IMD. The IMD is a measure of deprivation index. It ranges from 0 to 10. In the IMD scale, 0 (0-10%) stands for the 'most deprived' and 10 (90-100%) stands for 'least deprived' areas. The current study observes the impact on the child's development outcomes with the households moving to lower/higher IMD deciles than the current IMD decile level and hence, is the first study to utilise the movements along the IMD deciles to find the impact on child development outcomes. By observing the movements along

[1] As per the DWP definition, relative poverty refers to people living in households with income below 60% of the median in that year.

the deciles of the IMD, the current findings highlight whether there exists any impact on the child development outcomes as a result of moving to areas which are more deprived than the current residence. The analysis has highlighted consideration of the fact that the impact of change in IMD decile on the development scores is not linear, the effect varies for households and thus the study gives the average impact of these changes in the IMD decile as a result of parental mobility descisions on the child development outcomes. Hence, the findings indicate the adverse impact of the moves at this age after accounting for family structure, employment status and insecure tenure. Families who moved to 30% poorer areas shows adverse outcomes in terms of vocabulary and internalising behaviour for children. The present research also uses the UK MCS dataset and finds the marginal impacts of moving to a higher decile or lower decile IMD on child development scores and contributes to the mobility literature by studying the impact of household change via deprivation indices.

The chapter's analysis follows Cunha & Heckman (2007)'s approach on the importance of early stage child development outcomes and utilises the parental information of IMD movements. Throughout the current analysis, I have found the change in the z vocabulary and behavioural scores between ages 3 & 7; 3 & 11; and 3 & 14 age groups based on the parental descisions to move or stay at age 5 years. Therefore, the chapter evaluates the age-specific impacts of residential mobility on child development outcomes.

To achieve the impact of parental residential mobility on child development outcomes, I have used the longitudinal MCS data to analyse the age-specific impact of child residential mobility. The residential mobility is captured by the movements in overall IMD deciles, and individual (sub-IMDs) [2] IMD deciles, spanning between ages 3 to 7; 3 to 11 & 3 to 14. There are noticeable changes in vocabulary scores when households move (lower and higher) around five units of overall IMD deciles. The households moving higher up by one decile overall IMD show a slight decrease of 0.14 standard

[2] The detailed description of the sub-IMDs are shown in the subsection **??** of this chapter.

deviations in the vocabulary scores. At age 7, The maximum positive impact is 0.34 standard deviations when the households move from the most deprived to the least deprived overall IMD deciles. Among all the movements along different deciles of the individual IMDs, the health IMD shows a maximum increase in the vocabulary scores at ages 11 and 14. For the internalising scores when households move down the seven units of overall IMD decile, the behavioural issues increase by 0.28 s.d. The internalising score takes time to subside as households move from deprived crime IMD deciles to less deprived deciles. At age 11, there is a continuous rise in behavioural issues as the child moves to areas having higher deciles of overall IMD, and decreases when households move down the 3, 4 and 5 deciles of the overall IMD. But by age 14, the children show a decrease in the internalising score (decrease in behavioural issues) when moving higher up overall IMD decile and an increase when moving down the lower overall IMD deciles. The child behavioural issues decrease as one moves higher up the crime and education IMD deciles. There is a maximum increase of behavioural issues of 0.14 standard deviations when moving to the areas having lower deciles of overall IMD at age 14. For prosocial skills, there is an improvement in prosocial conduct as households move to higher IMD decile units through ages 7 to 14. Therefore, the findings in this chapter contribute to the literature on the impact of household movements on vocabulary and behavioural development. Any household movements lead to lower vocabulary scores and more behavioural issues at an early age of 7, but by age 14 these negative impacts minimise.

The remainder of the chapter is organised as follows: section 3.2 discusses some related studies; section 3.3 describes the data and variables used for analysis; section 3.4 shows the econometric framework used in the analysis; section 3.5 presents the results; section 3.6 discusses the main findings and compares them with previous studies; and section 3.7 concludes.

3.2 Literature Review

There has been burgeoning literature on the impact of residential mobility on child development in the UK (Clark et al., 2014; Gambaro & Joshi, 2016; Lupton, 2016). The studies by Beck et al. (2016) & Clark & Huang (2003) suggest that residential mobility is influenced by family demographics. Movements of the household are most often due to changes in partnership status, changes in jobs, housing or neighbourhood changes (Beck et al., 2016). Household mobility patterns have undergone changes over the past two decades in the UK. Clark & Huang (2003), by using the British Household Panel Survey (BHPS) data (1991-99), suggests that household movements were successful in terms of gaining more space to live following major household changes such as change in marital status, or birth. There was a decrease in the household mobility with an increase in homeownership, implying that households consider a number of factors before buying a home. Beck et al. (2016) studied the impact of reasons for household move on children's vocabulary and behavioural scores in the US. The impact on child well-being depended upon the circumstances that led to moving, and the mobility forms a characteristic of households where the child was aged 5 (after controlling for family and neighbourhood circumstances). The findings highlight a decrease in vocabulary scores due to changes in the partnership status, parent unemployment, and living in tenancy. For behavioural scores there was a significant adverse impact if hardships were faced in the tenancy compared to other factors.

In his study, Clark (2016) characterises moves according to multiple motives, and whether they are forced or chosen by households. Recent studies by Gambaro & Joshi (2016) & Gambaro et al. (2017) have focused on the negative impacts of pre-school children's residential mobility on schooling and peer relations. The estimates show that children belonging to households facing multiple transitions, within the bottom 30% of a disadvantaged area, show that there is a 36% chance of increase in peer problems. The increase in peer problems could be a result of other factors which can indirectly impact a child's behaviour, such as change in environment and increasing stress of adjustment to a new environment. There have been psychological studies by Jelleyman & Spencer (2008) & South et al. (2007) suggesting that there are adverse effects of residential mobility on the overall behavioural development of a child when controlling for the change in household composition, tenancy, parental unemployment.

Studies by Massey et al. (1993), Oishi & Schimmack (2010), & Shonkoff et al. (2000) focus on the importance of the purpose of changing residence. The household decision to move to make a better future has positive impacts, and moving to disadvantaged areas, which is marked by downward mobility, has negative impacts on child development, characterised by psychological distress and changing learning environments. This suggests the importance of neighbourhoods on child development outcomes. In their study, Bailey & Livingston (2008) highlight the importance of neighbourhood of origin, which influences the household decisions leading to segregation of neighbourhood, where households select neighbourhoods suitable for them. In response to these moves, the government could design policies in terms of resource allocation to deprived areas which many households move out of. As pointed by Gambaro & Joshi (2016) there exists an important role of parental composition such as the age of the parents while changing residence in the early years of child by affecting their supporting network. The neighbourhood quality was captured by the moves were classified as improvement, deterioration or no change within and above the bottom three deciles. The families who have moved to the bottom 30% of the areas showed a significant decline with respect to the households that stayed in the better 70% areas at age 5 years. The households moving to bottom 30% of deprived areas showed a significant impact of 0.37, at the same time showed a decline of 0.13 on vocabulary and increase of 0.36 standard deviations on internalising scores. Also, highlighted the adverse impact on the development outcomes because of the failure to move out of the poor neighbourhoods. Conversely, the study find that the households that moved to poor neighbourhoods from better off areas did not show significant worse outcomes could be due to the reason that the families had compensated by improving the living space. Hence, the study highlights that family moves involves stress and is transferred in the form of worse child outcomes for the families moving into the poor households as well as those who are not able to move out of these deprived areas. Lupton (2016) reviewed the housing policies since 1980 in the UK and 2018 in England, and suggested the importance of government policy in changing residence, describing 'advantaging moves', where households move to better areas, and 'disadvantaging moves', which are made of necessity rather than choice.

In contrast to the study by Rumbold et al. (2012) focusing on the number of moves and finding no association between the moves to either upward or downward mobile housing trajectory on the behavioural problems, the current study focus on the neighbourhood quality of mobility. To achieve this the current analysis highlights the impact of moving to lower deciles of IMDs than the current residence and shows that the movement to different deciles of IMD is nonlinear and does not show an improvement in the development scores when moving to better IMDs than the current. Some studies on residential mobility in England have used Lower Super Output Areas (LSOAs) as a base unit to classify areas as advantaged or disadvantaged. Gambaro & Joshi (2016) evaluates the importance of household moves by studying the predictors of moving at age 5 and the impact of household moves on the cognitive and behavioural development of the child. The analysis by Gambaro & Joshi (2016) uses logistic regression by characterising mobility as a binary variable which is coded as one to at least one move after considering employment transitions, area, and housing variables and family vulnerabilities and capabilities. To compare the moving locations the area lying in the bottom 30% of IMD is the poorest. The study finds a significant negative impact on the vocabulary scores of households moving within and staying in the bottom 30% of the IMD and also households there is an increase in behavioural problems as households move within the bottom 30% of IMD.

While Gambaro & Joshi (2016) and this chapter use the MCS datset, the current analysis uses the IMD indices provided in the MCS data using the ONS Census 2001 classification of geographical areas, and does not classify them as the poorest 30% or 70% better IMD areas (Noble et al., 2004). The IMD deciles range from most deprived to least deprived. Earlier studies, Gambaro & Joshi (2016), Gambaro et al. (2017) & Flouri et al. (2013), have used the MCS data to study the impact of residential mobility on child development outcomes between 9 months old to 5 years old, across the UK. To achieve this, they used IMD deciles provided in the dataset and combined the information provided in Census 2001 with the LSOAs indicator. Gambaro et al. (2017) uses the MCS dataset to examine the factors that lead households to move from infancy to age 5 and find that moves ended up in improvements to housing conditions. The findings also indicate that families with fewer resources showed trade-offs by remaining in the 30% of the lowest IMD in search of bigger housing, whereas families losing employment or experiencing partnership changes lead to less favourable moves. Flouri et al. (2013) has observed mobility between sweep 1 (age 9 months) to sweep 2 (age 3 years) as a basis for analysing child development at age 5 years. Flouri et al.'s (2013) findings show that there is no negative impact of residential mobility on child behavioural development. Hence, unlike the findings by the above authors, the current study utilises the movements along the deciles of the IMDs to find the impact on child development outcomes.

In the present chapter I use the MCS dataset which is also used by (Gambaro & Joshi, 2016; Gambaro et al., 2017; Flouri et al., 2013) to examine factors impacting household decisions to move. The current analysis focuses on children's vocabulary and behavioural outcomes between age 3 to 7; 3 to 11; and 3 to 14, where the child moved/stayed aged 5. The IMD is used as a scale to highlight the moves of the households to a particular area. The marginal analysis studies impact of the parental discisions to move higher or lower deciles of IMD. Where moving to higher decile indicates moving to areas with less deprivation index and vice-versa for moving to lower decile IMD (overall and individual). These IMD deciles movements are captured between ages 3, 7, 11 and 14 on vocabulary and behavioural development. The margins help in studying the unit changes in the household moves rather than to classifying the advantageous and non-advantageous moves as per the factors of household mobility. The findings in the chapter will aid the effectiveness of government policy in terms of resource allocation.

3.3 Data

The analysis uses MCS longitudinal data, consisting of household inputs (reasons for moving, language spoken by the parents, parental time investment, socioeconomic status, etc.) into children from an early age. The MCS is a collection of approximately 19,000 children born between 1 September, 2000, and 31 August, 2001, in England and Wales and between 24 November, 2000, and 11 January, 2002, in Scotland and Northern Ireland. The current research focuses on the mobile children (cohort members) within England [3], by collecting information on parental decisions to move, and household characteristics, siblings and neighbours. On a household level, the data set consists of information on the reasons for moving, such as whether the household wants to buy a larger home, move to a better area, move away from crime, and forced mobility reasons (also known as pull reasons) like problems with neighbours, unaffordability, or eviction. Respondents gave a set of reasons for movement from their current residence but they were not asked about the main reason for the mobility.

The MCS data set consists of seven sweeps to date. The children's ages range from 9 months (sweep 1) to 17 years (sweep 7). This study utilises the second to sixth sweeps of the MCS spanning children aged 3 to 14.[4] The information on the household socioeconomic status, parental heterogeneity, and child development components (vocabulary and behaviour) are collected from the parent and child

[3] Index of Multiple Deprivation (IMD) measures differ between countries of the UK.

[4] The analysis is age-specific, consisting of the age group- age 3 to age 7, 11 and 14; this age group consists of families also moving between age 3 and age 5.

interview questionnaires. The current study utilises the MCS surveys, mainly conducted with parents/parent-like figures when children are aged 3 to 7, and additionally from age 11 up to age 14 (early teenage). Most of the household movements in this study are observed during ages 5 and 7. In light of the importance of early-age development on to later child outcomes (Cunha & Heckman, 2007) it is interesting to know the impact of the household mobility decisions on various IMD measuring the income, crime, and educational characteristics on children's vocabulary and behavioural development.

3.3.1 Sample Representation-Index of Multiple Deprivation (IMD)

The MCS survey contains the 2004 IMD, ranging from most deprived ('0' IMD decile) to least ('10' IMD decile) deprived areas, where 0 consists of 0-10% and 10 consists of 90-100% IMD deciles. Apart from the overall IMD measure in the MCS dataset, there are six types of sub-IMDs which are referred to as individual IMDs in the current chapter. These individual IMDs are based on a set of factors such as income, health, education, housing, crime, and living environment. The Income IMD measures the proportion of population experiencing deprivation relative to low income; Health IMD measures the premature death and impairment of quality of life through poor physical or mental health; Education IMD measures the lack of attainment and skills in the local population; Housing IMD measures the physical and financial accessibility of housing and local services; Crime IMD measures the risk of personal and material victimisation at local level; and the Environment IMD measures the quality of both indoor and outdoor environment (Penney, 2019). The individual IMD measures are measured in a similar way as the overall IMD. The value of the individual IMDs ranges from 0 to 10 deciles. The 0 is the 'most deprived' decile IMD which consists of bottom 0-10%, and the value 10 is the 'least deprived' decile IMD which consists of top 90-100%. Hence, in the current chapter, the parental residential mobility descision is captured by the movements to areas having lower IMD decile or higher IMD decile relative to their current IMD decile of residence.

The IMD measures the relative deprivation for each small area (lower super output area (LSOA)). The deciles are provided in the MCS dataset on the basis of ranking the small areas in England from most deprived to least deprived and dividing them into 10 equal groups at population level. There are 32,844 LSOAs in England. The LSOAs are designed to have similar population sizes with an average of 1500 residents in each. There are differences in the size of LSOAs, and in the indicators and time periods used for the measure of deprivation indices between the countries in UK, hence it is impractical to have a direct comparison of IMDs between countries (Department for Communities & Local Government, 2015). In the MCS, IMDs are population standardised and vary from 1 to 10, from the most deprived to least deprived deciles.

To adequately represent ethnic minorities and disadvantaged groups, the population of electoral wards were classified as disadvantaged, advantaged and ethnic minority.[5] The current study uses the MCS technical report for information on the formation of strata and clustering of the sample (Plewis et al., 2007).

In the MCS, a total of 11,532 families (including twins and triplets) in England were selected for interview, of which, 691 families entered the MCS at sweep 2. This study uses information on sweep 1 and new families. The analysis includes only the first cohort member in the twins' family to account for statistical complexity, while clustering at family level and using the non-response weights for England. Triplets are not included in the study as the data availability is secured access for them, unlike the single and twin families whose data is available under end-user licence.

Sample selection

[5]The MCS technical report by Plewis et al. (2007) characterises Disadvantaged ward where the population lies on the upper quartile of the 1998 Child Poverty Index (CPI) of England and Wales. These families were characterised as receiving either Income support, Jobseekers allowance, Family Credit and Disability Working Allowance. The Advantaged stratum consists of the population of the ward lying in the lower quartile of the 1998 CPI. Ethnic minority stratum was only present in England where 30% of the total population was "Black" or "Asian".

The MCS data consists of information on the socioeconomic measure of the household, such as income in bands, availability of materials to determine material deprivation, and OECD income. For analysis, the current study utilises the residential mobility responses of parents, and whether they move or stay at the same level of IMD decile as their current residence IMD decile. For the analysis, residential mobility is captured by the variable mobility, which is categorised as 0 "stayed at same address with same/different IMD decile" 1 "moved to area having higher/lower IMD decile" with change in the address. Movers are further categorised as moving to areas having either higher IMD decile or lower IMD decile. Staying families did not change their previous sweep residence or, if changed, remained at the same IMD.[6]

The current study observes the mobile households of children throughout ages 3 to 14. There are three categories of the households: stayers (who did not change residential address from previous sweep or moved within same IMD), those who moved to areas having higher IMD deciles, and those who moved to areas having lower IMD deciles. The current research considered households who did not move as a base category. I have coded the households who moved to area having same IMD decile, a base category, as they constitute only 2% of the entire sample. In table 3.1, there are approximately 41% of movers in the entire sample throughout sweep 2 (age 3) to 6 (age 14) who have moved at least once from age 3 to age 14, irrespective of the change in the IMD decile.[7] Hence, the majority of families (approximately 59%) stayed in the same address throughout ages 3 to 14. Moving to a higher/lower IMD decile is observed for each sweep. The sample is pooled with 45,090 observations, consisting of observations that contain information for the overall and individual IMDs. For sweep 2, there are 10,079 observations; 9,755 for sweep 3; 8,881 for sweep 4; 8,664 for sweep 5 and 7,711 for sweep 6.

[6]I have dropped families who did not move but their IMDs changed.

[7]The number of stayers and movers are based on the information provided for overall IMD change i.e., it does not include information on individual IMD changes.

TABLE 3.1: Family mobility status throughout age 3 to 14

	N	Percent
Stayed throughout	6309	59.40
Mobile Families(anywhere between sweeps)	4313	40.60
Moved once	3255	30.64
Moved twice	844	7.95
Moved thrice or more	214	2.01
Total families	10,622	100.0

Note: The sample captures households that are present on one of the sweeps between ages 5 to 14 and report the responses as 1 for families who have moved when asked if there is any change from the previous residence.

3.3.2 Attrition

The current chapter addresses the issue of attrition similar to chapter 1. Parents could not respond to all the household mobility decision questions and hence specific weights were assigned to address the attrition. Attrition weights for single country analysis were used, consequently, the current analysis consists of only England as an area of study. Mobility questions asked only to main respondents were considered in households where both main parental figure and partner responded, provided the child was currently living with them. I have used the MCS technical report (Plewis et al. (2007)) information to accounting for non-response weights. For analysis, the overall longitudinal UK population weights estimate outcome variables of corresponding sweeps in regression models.[8] The overall weights are used because they are adjusted for non-response and have a broader range than the sample weights, as the former combines with sample weights and non-response weights for the UK as a whole.[9]

3.3.3 Descriptive statistics-IMD

The present analysis uses the ’overall’ (combined IMD deciles) and the ’individual’ (conducting individual impact of each IMD categorised in the MCS dataset such as:

[8]The current study focuses on the child development outcome for the UK as a whole consisting of input information from households of each country (England, Wales, Scotland and Northern Ireland).

[9]The attrition is accounted for by the overall UK weights standardised with mean 1 and multiplied together to give an overall unit non-response weight. The non-response is caused by a change in the respondent from the previous sweep.

income, housing, education, health, crime, and environment) components of the IMD to visualise the change (improvement or decline) in the vocabulary and behavioural scores at ages 4, 7 and 11, given the households moved between ages 3 and 5. The change in the IMDs, i.e., moving to higher or lower deciles of the IMDs is captured by the dummy variable, which equals 1 if households move to lower deciles of the IMDs and 0 if households move to higher deciles of the IMDs, otherwise.

Child outcomes

Various cognitive and non-cognitive measures are used to study child development. The age-standardised cognitive assessments, such as BAS naming vocabulary at age 3 and 5, measure the knowledge of nouns of the cohort member. The BAS Word Reading is conducted at age 7, and the test assesses the cohort member on expressive spoken vocabulary, and a set of coloured pictures of objects are shown to the child one by one, which they are asked to name. Success depends on the cohort member's previous knowledge of a vocabulary of nouns.[10] At age 11, a BAS verbal similarities test is conducted; at age 14 years, word activity tests are conducted. In the BAS word activity test, respondents were provided with a list of 20 target words; the interviewer asked the respondents to match the target word with the word provided next to them with the same or approximately the same meaning. The child's score is based on accuracy and the speed with which the task is completed.

The age-standardised test scores are used for the purpose of evaluation of child development. The current analysis uses the age-standardised raw test scores. Raw scores are the number of correct answers. The cohort members did not complete the same set of test items, as it was based on the child's experience with testing and the time taken by the child (Connelly, 2013; Elliott et al., 1996; Hansen et al., 2010).[11]

[10]The score ranges from 0-80 for BAS naming vocabulary, and BAS word reading ranges from 55-145.

[11]In the age-standardised evaluation of vocabulary scores, the children are presented with test items that are suitable for their age and are not too easy or too difficult. An initial set of items is given to each child for completion and based on the performance on the initial set the children are asked to stop if they have scored sufficient correct and incorrect answers. Hence, this age-specific test determines a child's ability to progress to more difficult problems if the initial set of problems is too easy or return to

Another component of child outcome is non-cognitive development. From sweep 2 (age 3) to sweep 6 (age 14), internalising and externalising behaviours are measured by four sub-items in the SDQ index and externalising by prosocial score. The behavioural outcomes are measured in the sub-scales of mood (5 items), adaptability (2 items), approach-withdrawal (3 items) and regularity (4 items). The SDQ measurement for internalising behaviour consists of five scales; emotional symptoms, conduct problems, hyperactivity/Inattention, peer relationship problems and prosocial behaviour. Each scale consists of five items. The parents answer each item based on three components: 1 "not true", 2 "somewhat true," and 3 "certainly true". The components scores ranges from 0 to 2. "Somewhat true" is consistently scored as 1, but "not true" and "certainly true" vary with each item. For each of the five scales, the score can range from 0 to 10 if all items are completed. The total difficulties score is calculated as the sum of the scores of only the first four scales, excluding the prosocial score. Total difficulties score ranges from 0 to 40, ideally when all the scales are present with complete scores. As per the item description under each scale, I assume that a higher total difficulty score for a child tends to correspond to more behavioural issues (Goodman, 1997; Johnson et al., 2015). Conversely, higher values of prosocial score show fewer behavioural problems in child.

In the current study, I have computed the changes in the outcome variables for families move to higher or lower IMDs by standardising the changes in the vocabulary and behavioural test scores with mean 0 and standard deviation 1 for a given sample

more simple problems if the initial set of problems is too difficult. Hence, the raw scores are not directly comparable and raw scores are converted to ability scores, which consider only a specific sets of items which the cohort members were presented. Rasch model-Item response theory was used to convert the raw scores into ability scores. The model consists of the probability of successful completion of the test item, based on the difficulty of the items and the ability of the child. This model then predicts the ability of the child based on the level of difficulty of the completed items. Ability test scores overcome the problem of comparison of raw test scores among children who have finished different set of test items. Hence, standardised test scores are used for the current analysis which are age -adjusted and take account of the child's age and the difficulty level of test item completed by the child. Look up tables are used for the conversion to standardised test scores as provided in the BAS II manual (Connelly, 2013; Elliott et al., 1996; Hansen et al., 2010).

means for a particular age group before weighting to facilitate comparison across the children's ages.

3.3.4 Control variables

The selected control variables are a collection of the household socioeconomic environment, parental inputs comprising of reasons to change residence and child-specific variables. Unlike Gambaro et al. (2017) who has used family income to control for residential mobility, the current study uses the OECD income-adjusted household level size as a socioeconomic indicator, since the outcome variable in the present study is change in children's vocabulary and behavioural development and the study observes per unit change in the IMD deciles, hence the OECD income simplifies the analysis by considering the broader classification of household income groups after controlling for factors leading to household mobility. Parental heterogeneity is captured by the variable for whether it is a single or two-parent household. Single and two-parent household proxies the inputs child receive from the parents in the form of the decision related to resource allocation for child's development, assuming that more resources are allocated to a child two-parent households, also the influence of mother and father in two-parent households impacts the parental decisions to allocate resources to child a compared to single-parent household. The positive impact of higher maternal education at the birth of child is indicated by the household availability of resources such as books, taking child to library that inculcates the habit of reading in the child, and reading to child more often which provides positive environment for vocabulary development. Additionally, a higher socioeconomic status has a strong correlation of 0.29 with the household literacy environment, such as highly educated mother, having more books, and parents frequently reading (Duursma et al., 2007). The presence of siblings also provides an environment for a child's development through help in vocabulary, which again depends on the age of the sibling. Elder siblings, after controlling for subject specific school characteristics and inputs, tend to have 0.110 significant positive impact on children's development (Nicoletti & Rabe, 2019). Maternal characteristics such as maternal age, whether born in the UK, suffering from long-term illness [12], maternal ethnic background (0-white and 1-nonwhite) are some of the factors that impact child's vocabulary and behavioural development. Byers-Heinlein & Lew-Williams (2013) highlighted the causal impact of household mobility due to neighbourhood origins using the qualitative case study in two different Canadian neighbourhoods through examining the impact of social exclusion and attachment. Hence, a neighbourhood attachment plays an important role in household mobility decisions. A neighbourhood attachment could be defined in terms of households' decisions to move to areas that are similar to their household characteristics.

Pre-determined child specific inputs, ethnicity (0-white, 1-nonwhite), birth weight in kgs, child's age in months, whether child suffers from long-term illness, are also included to find the changing vocabulary and behavioural scores when moving from age 3 to 7, 11 and 14.

3.4 Method

The current analysis uses the child production function approach introduced by Todd & Wolpin (2003, 2007).[13] The child production function is widely used in educational research (Bono et al., 2016; Cobb-Clark et al., 2018; Del Boca et al., 2017; Dickerson & Popli, 2018). The function evaluates a child's current performance based on the lagged period child test scores, which provides a child's performance trajectory across the school years for subjects taught in class but some factors could lead to endogeneity in the measure of the child development, such as inherited abilities that are not available in the dataset and cannot be measured, leading to a not causal impact of parental mobility decisions on child vocabulary and behavioural outcomes. The

[12] This also includes the variable whether mother suffers from maternal depression

[13] Todd & Wolpin (2007) has found a significant association of previous period child endowments, parental inputs, household characteristics on the cognitive performance at the school-going age.

following analysis attempts to find the causal impact of parental mobility decisions by addressing the endogeneity issue.

In this chapter, the child development is modelled as being influenced by parental decisions to move to a lower or higher IMD or stay at the same IMD. The following equation provides the child development for a given age:

$$Y_{it} = \alpha + \beta M_{it} + \rho C_i + \gamma H_i + \psi cage_{it} + \eta cage_{it}^2 + \mu_i + \varepsilon_{it}. \quad (3.1)$$

In the equation (3.1) Y_{it} is the child, 'i"s, vocabulary and behavioural outcomes at sweep, t (t consists of sweep 2 (age 3) to sweep 6 (age 14)) for household 'h'. M_{it} measures the overall deprivation decile where its values range between 1 to 10 for the child's age at time 't', here '1' stands for the 'most deprived' and '10' for least deprived. IMD values 1 through 10 measure the deprivation index for an area in the 10th percentile. IMD value of 1 consists of an area between 0%-<10% level of deprivation, 2 consists of an area between 10%-<20%, etc for successive IMD levels. IMD 10 denotes the area at the upper 10th percentile of the IMD at given sweep t. Hence, M_{it} is treated as a continuous variable in the current analysis. The term C_i, is a set of child-specific inputs for cohort 'i'. The child specific inputs consist of time-invariant inputs such as ethnicity, birth weight and whether child suffering from long-term illness. The term H_i, is a set of household-specific inputs for cohort 'i'. The household inputs consist of time-invariant inputs such as whether mother was born in the UK, mother suffering from long-term illness at age 3, language spoken in household at age3, one or two parents, age of the mother at the birth of the child, OECD income of household at age 3, mother's education at age 3 and the IMD of the household at age 3. $cage_{it}$ indicates the child's age in years(non-squared) and $cage_{it}^2$ as squared at time t which is the only time-variant child-specific input in the analysis. μ_i captures the unobserved child-specific inputs such as ability, genetic endowments, and factors affecting parental decisions to move which are not captured in the data. ε_{it} captures the idiosyncratic error, which consists of unobserved effects of omitted inputs such as previous age child cognitive ability, child endowments and measurement error, depending on the factors affecting parental decision to move, in turn affecting cohort child outcomes. The MCS data is rich in child-level variables and parental inputs (decisions) accounting for child development for each age group in our analysis. However, in the analysis of child outcomes where children differ in ethnicity, parental heterogeneity, and household inputs, information on the child's ability, child's attitude, and changing parental decisions which constantly impact the child outcomes at each age group, are not available. Hence, equation (3.1) is not valid since the β estimate is not consistent as the omitted variables capturing unobserved impacts on child development should ideally be is not orthogonal to the C_i, and H_i, and $cage_{it}$, and $cage_{it}^2$ co-variates after addressing for the term μ_i + ε_{it} i.e.,

E((C_i,H_i,$cage_{it}$,$cage_{it}^2$ | μ_i + ε_{it})$\neq$ 0

To account for the estimation issues arising from the impact of unobserved inputs in equation (3.1), the following equation (3.2) uses the first differencing of (3.1), showing the impact of the previous period (sweep2) parental decision to move to different IMD on vocabulary and behavioural scores at time t.

$$Y_{it} - Y_{i,s2} = \alpha + \beta \Delta M_i + \gamma scr_{iY,s2} + \theta M_{it} + \rho M_{i,s2} + \zeta C_i + \phi H_i + \psi cage_{it} + \eta cage_{it}^2 + \xi_{it}. \tag{3.2}$$

In equation (3.2), the child development for sweep 't' is observed as a difference between current sweep 't' child development and sweep 2 (age 3) denoted by subscript 's2' child development scores, the current sweep 't' development is determined by parental decision at a previous period (sweep 3 (age 5)) to move to different IMD (e.g. moving from IMD 1 to 2 or moving from IMD 6 to 5). The $scr_{iY,s2}$ is sweep 2 test score of 'Y' for measuring 'Y' outcome (vocabulary and behaviour) which captures the impact of an age 3 child's development on current sweep 't' child development. The sweep 't' stands for sweep 4 (age 7), sweep 5 (age 11) and sweep 6 (age 14).

Subsequently, the analysis is carried on for sweeps 5 (age 11) and 6 (age 14) where the impact of parental decisions of moving (between age3 and age5) on the changing child development scores from sweep 2 (age 3) to sweep 3 (age 5) is observed. This can be explained as the impact of residential mobility at t-1 (between age 3 and 7 test scores), t-2(between age 3 and 11 test scores) and t-3 (between age 3 and 14 test scores) lag periods on sweep 4 (age 7), sweep 5 (age 11) and sweep 6 (age 14) change in child development outcomes respectively. This approach of inclusion of lagged period inputs of child development is also adopted by studies on child development (Bono et al., 2016; Cobb-Clark et al., 2018; Fiorini & Keane, 2014) to measure current period child development. The sweep 't' stands for sweep 4 (age 7), sweep 5 (age 11) and sweep 6 (age14) for the current analysis. The resultant parental decision to move is captured by ΔM_i which consists of the change in IMD (measured as a difference of moving between different IMDs) when moving between age 3 (sweep 2) for households who moved/stayed at age5; the impact of change in residence between ages 3 & 7, 3 & 11 and 3 & 14 and age 5 (sweep 3) to capture the impact of changing residence in between ages 3 and age 7. M_{it} is the current period IMD and $M_{i,s2}$ is IMD at sweep 2(which forms a part of household inputs at age 3).

In equation (3.3), the quadratic term, $\Delta M^2{}_{it}$ captures the nonlinear impact of moving to different IMD deciles from the current, on the development scores and highlights that moving to IMD deciles can improve or decline the development scores, hence, the current analysis providing the average impact of the change in the development scores as a result of moving to different IMD deciles. Also, the term $\Delta M_i * lM_i$ shows the impact of moving to lower IMD decile, captured by the term 'l' whose value ranges from 0 'higher IMD decile' to 1 'lower IMD decile' from the current residence on the development outcomes. Also, the term ΔM_i captures the change in the IMD deciles when households move from their current residence. Hence, the terms $\Delta M^2{}_{it}$ and $\Delta M_i * lM_i$ show that the impact of moving to different IMD deciles is not the same and therefore, the term $\Delta M^2{}_{it}$ gives the average impact of movement. In summary, to segregate the impact of movement to the lower IMD decile, the term, $\Delta M_i * lM_i$ provides the estimate of development scores when households move to lower decile of the IMD index. Therefore, the equation (3.3) provides an overall analysis of the impact of the impact of the movement to different deciles of the IMD index on the development scores.

$$\begin{aligned} Y_{it} - Y_{i,s2} = \alpha + \beta \Delta M_i + \gamma \Delta M^2{}_i + \eta \Delta M_i * lM_i + \rho \Delta M^2{}_i * lM_{i,} + \kappa scr_{iY,s2} \\ + \theta M_{it} + \chi M_{i,s2} + \zeta C_i + \phi H_i + \psi cage_{it} + \omega cage^2_{it} + \xi_{it}. \end{aligned} \tag{3.3}$$

The equation (3.3) is an extension to the equation (3.2) which consists of similar controls. In addition, the equation (3.3) includes the squared change in IMD shown by $\Delta M^2{}_i$ captures the non-linear relation between the change in IMD and the child outcome. The interaction term $M_i * lM_i$ with the negative change in IMD highlights the relative increase in the child's change in test scores for the households moving to lower IMD compared to those moving to higher IMD. Hence, the interaction term explains the non-linear impact of the change in IMD for the families moving to lower IMD at age 7,11 and 14 from age 3. The lM_i stands for low IMD; it is a dummy variable which consists of 0-not moving to lower IMD at ages 7,11 and 14 from age 3 and 1-moving to lower IMD at ages 7,11 and 14 from age 3.[14]

3.5 Results

The age-specific analysis on the impact of moving to each IMD (ranging from the combination of all indicators as well as individual indicators of deprivations) are shown in tables 3.2 to 3.4 for vocabulary, tables 3.5 to 3.7 for internalising (higher internalising scores indicate more behavioural issues), and tables 3.8 to 3.10 for externalising scores (higher externalising scores indicate less behavioural issues).

[14]The F-test of joint significance of M_i and $M^2{}_i$ has a test-statistic of 2.56 hence provides sufficient evidence to reject the null at 95% CI.

3.5.1 Vocabulary scores

The combined impact of movements to different deciles of overall IMD on vocabulary outcomes is shown in column (1) of tables 3.2 to 3.4. For age 7, when households move to areas having different deciles of overall IMD, there is a decrease in the vocabulary score of around 0.1 standard deviations when moving to higher IMD decile and an increase of 0.1 standard deviations when moving to the lower IMD decile than the current. Columns (2), (4) and (7) of table 3.2 also show that overall, the change in vocabulary scores bore a negative impact in moving to the lower income, education and environment IMD deciles with a significant decline of 0.1 standard deviations when moving down the education IMD decile compared to movement in higher decile IMD areas. The negative impact of changing IMD are shown in column 1 of tables 3.3 for age 11 & 3.4 for age 14. There is no overall impact of change in IMD deciles due to household mobility at ages 11 and 14. At age 11 there is a significant increase of 0.08 s.d. on vocabulary scores as a result of households moving down the lower environment IMD deciles. But by age 14 there is no impact of change in individual IMD deciles due to household mobility. There is a constant increase in the significant negative impact of age 3 z-vocabulary scores on the vocabulary score at ages 7, 11 and 14 for overall and individual change in IMD deciles. Hence, vocabulary scores are more responsive to the household socioeconomic and health conditions of the child where the later is influenced by the air quality of that area at age 11. But, this impact later subsides at age 14.

Marginal impact of the change in the vocabulary scores.

The marginal impact of moving to areas having higher and lower IMD deciles for vocabulary scores are shown in figures 3.1 to 3.3 for ages 7 to 14 in panel(a). The panel (a) of figure 3.1 shows that there is a significant decline in z vocabulary scores when households move to areas which are higher up by 1,2,3 & 4 IMD deciles. For age 11, z vocabulary scores decline when households move to areas with 4,5,6 & 7 higher IMD deciles. But, there is no marginal impact of changing residence at age 14. At age 14, there is a significant increase in the vocabulary scores when housholds move to areas higher up by 1,2,3 & 4 IMD deciles as shown in panel(b) of figure 3.3. There is a significant decline in vocabulary score at age 7 when households move higher up by 1 and 2 health IMD deciles and by 5 IMD decile at age 11. For the education IMD, there is a significant decline in vocabulary scores when moving higher up by 1 IMD decile at age 7. But a significant increase in vocabulary scores when households move lower down by 9,8,7 & 6 education IMD deciles. Also, at age 11, there is an increase in vocabulary scores when moving lower down by 9,8 & 7 IMD deciles. For housing IMD at age 7 decline in vocabulary score, when moving higher up by 1 & 2 IMD deciles. At age 11, decline in vocabulary score when moving higher by 4,5,6,7,8 & 9 IMD deciles. For crime IMD, decline in vocabulary scores at age 7 when moving higher up by 1 & 2 IMD deciles. Also, a decline when moving down by 9,8,7 & 6 crime IMD deciles at age 7. At age 11, decline when moving higher up by 1,2,3,4,5 & 6 crime IMD deciles. At age 14 increase in vocabulary when moving higher up by 1 IMD decile. For environment IMD, decline at age 7 when moving higher up by 1,4,5,6,7,8 & 9 IMD deciles. At age 11, decline when moving higher up by 1,2 & 3 environment IMD deciles.

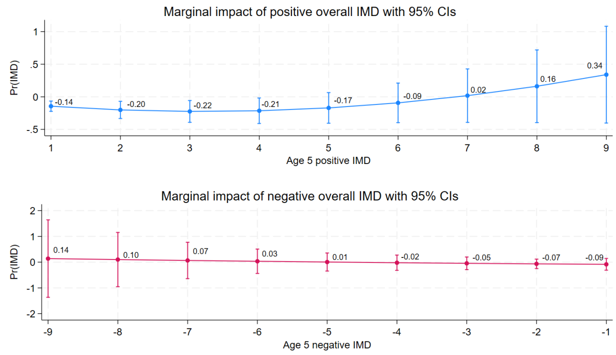
TABLE 3.2: The change in child vocabulary scores between sweeps 2 (age 3) and 4 (age 7)
for families moving/staying at sweep 3 (age 5).

	Overall Indv. IMDs						
	All reason (1)	Inc. (2)	Hlth. (3)	Edu. (4)	House (5)	Crime (6)	Env. (7)
OLS with controls							
Δ(IMD) (Base=Stayed at same IMD or different IMD)							
Change in IMD	-0.107**	0.011	-0.076*	0.012	-0.035	-0.030	0.003
	(0.048)	(0.041)	(0.046)	(0.020)	(0.039)	(0.031)	(0.022)
Change in IMD^2	0.017**	-0.004	0.014*	-0.007**	0.002	0.002	-0.006**
	(0.008)	(0.008)	(0.008)	(0.003)	(0.008)	(0.005)	(0.003)
-ve Δ in IMD	-0.046	-0.013	0.016	-0.120	0.003	0.072	-0.019
movement(Base=+ve change)	(0.234)	(0.108)	(0.105)	(0.095)	(0.091)	(0.091)	(0.102)
Change in IMD for lower IMD (ΔM#low IMD)	0.093	-0.048	0.052	-0.076***	0.053	0.051	0.004
	(0.172)	(0.066)	(0.087)	(0.025)	(0.062)	(0.045)	(0.048)
Change in IMD^2 for	-0.015	-0.003	-0.025*	0.009**	0.001	-0.010	0.005
lower IMD (ΔM#ΔM#low IMD)	(0.025)	(0.013)	(0.014)	(0.004)	(0.012)	(0.008)	(0.010)
Lagged z test scores (sweep2)	-0.291***	-0.292***	-0.290***	-0.289***	-0.292***	-0.289***	-0.289***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Constant	9.835	9.384	9.864	9.566	9.654	9.327	9.019
	(9.529)	(9.553)	(9.554)	(9.474)	(9.475)	(9.463)	(9.519)
R^2	0.165	0.165	0.165	0.167	0.165	0.165	0.164
N	6,058	6,058	6,058	6,058	6,058	6,058	6,058

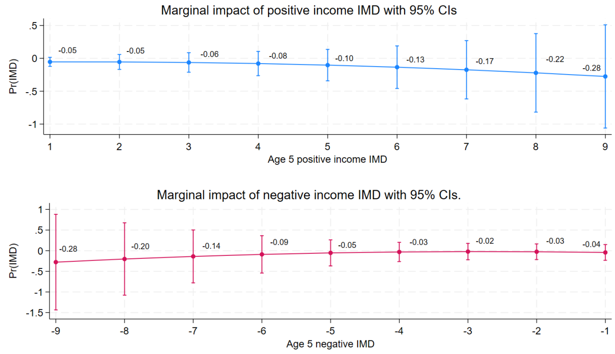
Notes.- The robust standard errors are in parenthesis. Child's age in months is controlled. Outcome is difference in standardised Z vocabulary scores between sweeps 2 (age 3) to 4 (age 7) with mean 0 and standard deviation 1. The variable of interest is the movement of families between IMDs for sweeps 2 (age 3) and 3 (age 5). The analysis also controls for child's age in months at sweeps 4(age 7), current period IMD level. The lagged period controls used consist of respective lagged period (sweep 2) outcomes, sweep 4 IMD. Child and household level controls consist of child's and mother's ethnicity, number of siblings, OECD weighted household income, mother's highest educational qualification, language spoken in household, mother having long term illness, child having long term illness, whether mother is born in the UK, child birth-weight, whether single or two parent household, age of mother at birth of child.
*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

FIGURE 3.1: The change in child vocabulary scores between sweeps 2 (age 3) and 4 (age 7) for families moving/staying at sweep 3 (age 5).

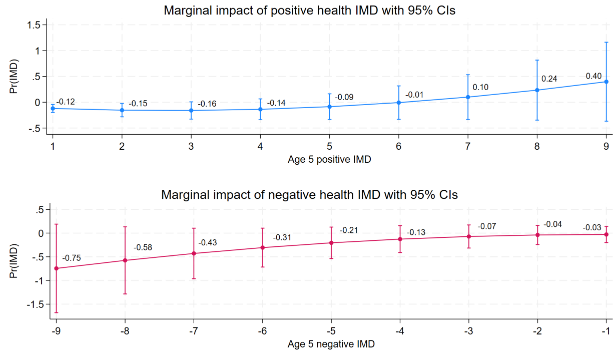
(a) Overall IMD



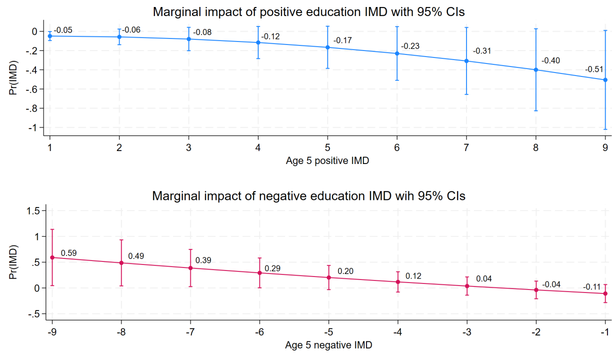
(b) Income IMD



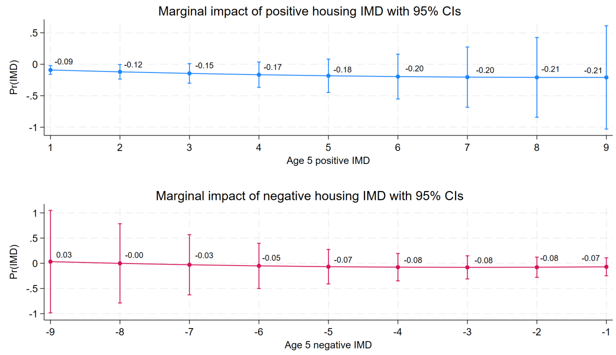
(c) Health IMD



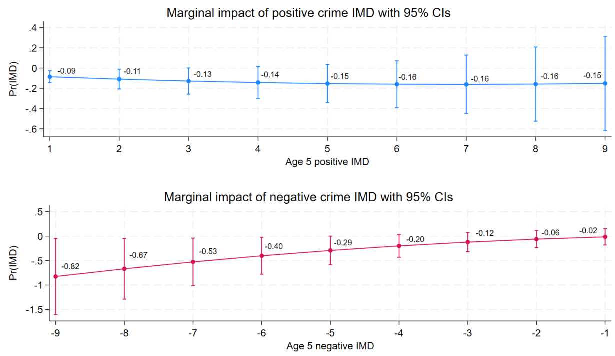
(d) Education IMD



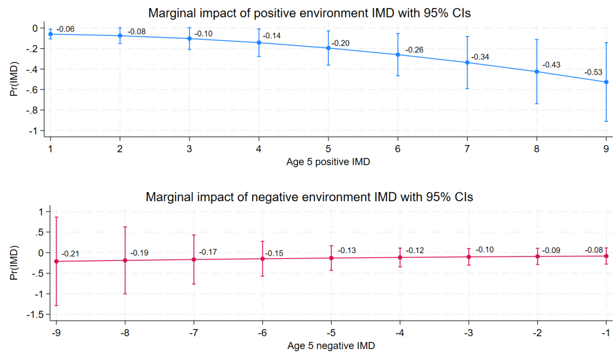
(e) Housing IMD



(f) Crime IMD



(g) Environment IMD



Notes.-The author's calculations are based on the families moving between different IMDs or staying at same IMD at age 5. The coefficients capture the change in IMD as a function of families moving between different levels of IMD. The difference in IMD scale is converted to absolute values so as to compare it with the families having positive difference in IMD at age 5. The outcome variable is the change in the Z vocabulary scores between ages 3 and 7 years.

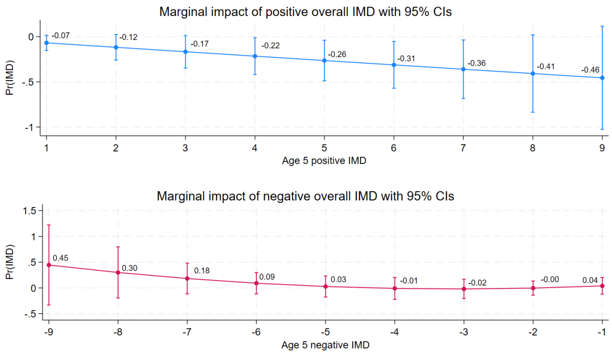
TABLE 3.3: The change in child vocabulary scores between sweeps 2 (age 3) and 5 (age 11) for families moving/staying at sweep 3 (age 5).

	Overall Indv. IMDs						
	All reason (1)	Inc. (2)	Hlth. (3)	Edu. (4)	House (5)	Crime (6)	Env. (7)
OLS with controls							
Δ(IMD) (Base=Stayed at same IMD or different IMD)							
Change in IMD	-0.050	0.041	-0.044	0.012	0.005	-0.041*	-0.049*
	(0.050)	(0.044)	(0.062)	(0.033)	(0.035)	(0.025)	(0.027)
Change in IMD^2	0.000	-0.016	0.002	-0.009	-0.014*	0.003	0.007*
	(0.008)	(0.010)	(0.011)	(0.008)	(0.006)	(0.004)	(0.004)
-ve Δ in IMD	0.128	0.032	0.033	-0.014	0.067	0.051	0.047
movement(Base=+ve change)	(0.171)	(0.077)	(0.072)	(0.075)	(0.069)	(0.071)	(0.077)
Change in IMD for lower IMD	0.133	-0.013	0.025	-0.027	0.008	0.030	0.084**
(ΔM#low IMD)	(0.132)	(0.052)	(0.078)	(0.045)	(0.035)	(0.033)	(0.035)
Change in IMD^2 for	0.013	0.022**	-0.006	0.013**	0.013	-0.005	-0.001
lower IMD (ΔM#ΔM#low IMD)	(0.018)	(0.011)	(0.015)	(0.006)	(0.010)	(0.006)	(0.006)
Lagged z test scores (sweep2)	-0.620***	-0.621***	-0.619***	-0.619***	-0.621***	-0.619***	-0.618***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Constant	-1.622	-1.733	-1.838	-1.979	-2.127	-1.827	-1.366
	(7.266)	(7.247)	(7.416)	(7.429)	(7.325)	(7.385)	(7.357)
R^2	0.420	0.421	0.419	0.422	0.422	0.419	0.420
N	5,274	5,274	5,274	5,274	5,274	5,274	5,274

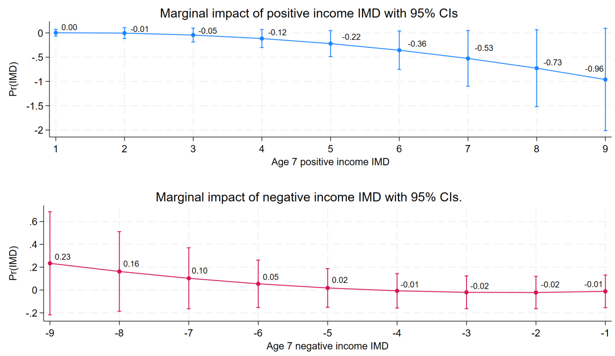
Notes.- The robust standard errors are in parenthesis. Child's age in months is controlled. Outcome is difference in standardised Z vocabulary scores between sweeps 2 (age 3) to 5 (age 11) with mean 0 and standard deviation 1. The variable of interest is the movement of families between IMDs for sweeps 2 (age 3) and 3 (age 5). The analysis also controls for child's age in months at sweeps 5 (age 11), current period IMD level. The lagged period controls used consist of respective lagged period (sweep 2) outcomes, sweep 5 IMD. Child and household level controls consist of child's and mother's ethnicity, number of siblings, OECD weighted household income, mother's highest educational qualification, language spoken in household, mother having long term illness, child having long term illness, whether mother is born in the UK, child birth-weight, whether single or two parent household, age of mother at birth of child.
*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

FIGURE 3.2: The change in child vocabulary scores between sweeps 2 (age 3) and 5 (age 11) for families moving/staying at sweep 3 (age 5).

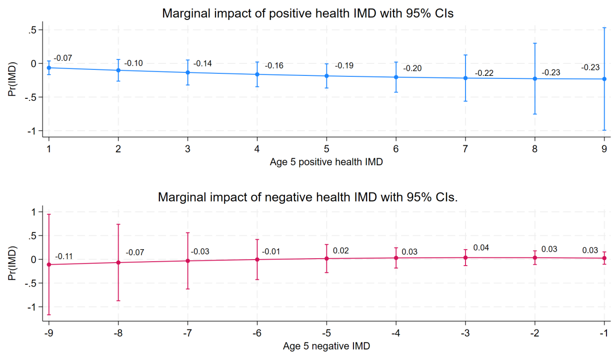
(a) Overall IMD



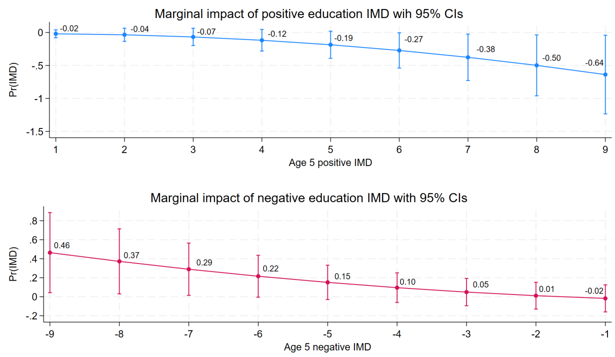
(b) Income IMD



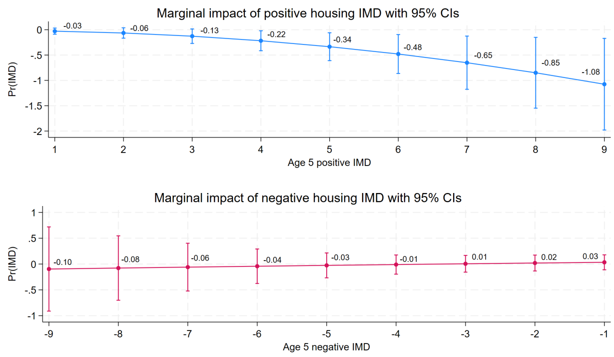
(c) Health IMD



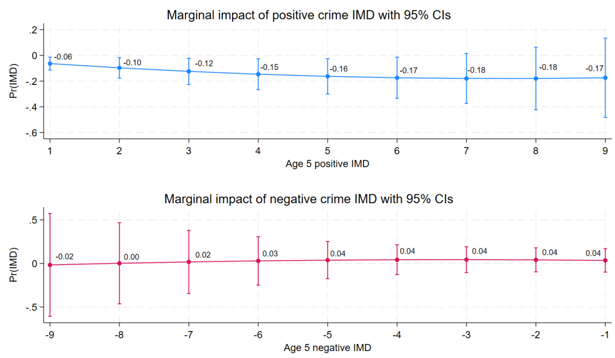
(d) Education IMD



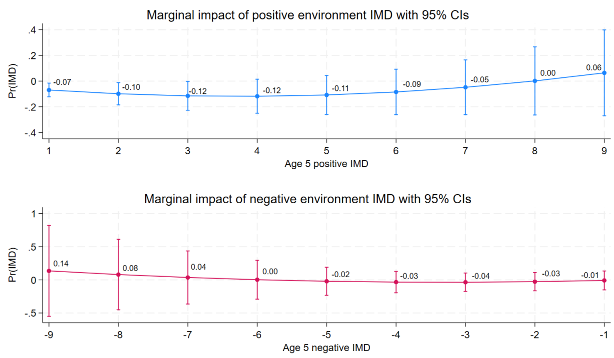
(e) Housing IMD



(f) Crime IMD



(g) Environment IMD



Notes.-The author's calculations are based on the families moving between different IMDs or staying at same IMD at age 5. The coefficients capture the change in IMD as a function of families moving between different levels of IMD. The difference in IMD scale is converted to absolute values so as to compare it with the families having positive difference in IMD at age 5. The outcome variable is the change in the Z vocabulary scores between ages 3 and 11 years.

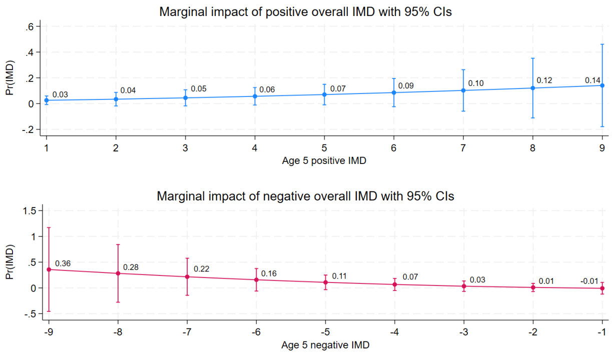
TABLE 3.4: The change in child vocabulary scores between sweeps 2 (age 3) and 6 (age 14) for families moving/staying at sweep 3 (age 5).

	Overall Indv. IMDs						
	All reason (1)	Inc. (2)	Hlth. (3)	Edu. (4)	House (5)	Crime (6)	Env. (7)
OLS with controls							
Δ(IMD) (Base=Stayed at same IMD or different IMD)							
Change in IMD	0.006	0.018	-0.001	-0.004	0.000	0.015	0.004
	(0.021)	(0.016)	(0.022)	(0.010)	(0.020)	(0.015)	(0.018)
Change in IMD^2	0.001	-0.002	0.002	-0.001	0.001	-0.002	-0.001
	(0.004)	(0.003)	(0.004)	(0.002)	(0.003)	(0.002)	(0.003)
-ve Δ in IMD	-0.033	-0.016	-0.055	0.007	0.001	0.034	-0.009
movement(Base=+ve change)	(0.120)	(0.048)	(0.049)	(0.042)	(0.045)	(0.042)	(0.045)
Change in IMD for lower IMD	-0.010	-0.057*	-0.065	-0.006	0.012	0.002	0.003
(ΔM#low IMD)	(0.088)	(0.032)	(0.047)	(0.015)	(0.032)	(0.020)	(0.024)
Change in IMD^2 for	0.003	-0.004	-0.010	0.002	0.002	0.002	0.004
lower IMD (ΔM#ΔM#low IMD)	(0.013)	(0.006)	(0.010)	(0.003)	(0.005)	(0.004)	(0.004)
Lagged z test scores (sweep2)	-0.882***	-0.878***	-0.878***	-0.878***	-0.878***	-0.878***	-0.878***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Constant	-3.782	-3.868	-3.847	-3.880	-3.776	-3.539	-3.513
	(4.321)	(4.342)	(4.337)	(4.338)	(4.332)	(4.332)	(4.332)
R^2	0.842	0.842	0.842	0.842	0.842	0.842	0.842
N	4,433	4,433	4,433	4,433	4,433	4,433	4,433

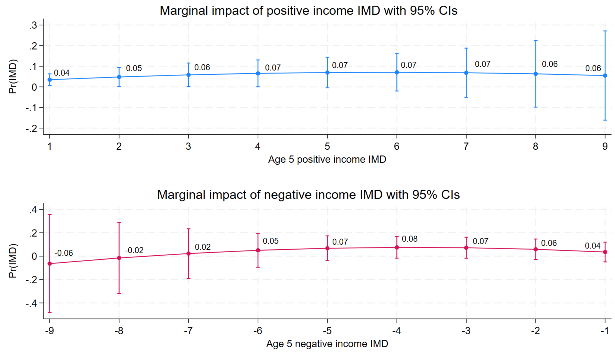
Notes.- The robust standard errors are in parenthesis. Child's age in months is controlled. Outcome is difference in standardised Z vocabulary scores between sweeps 2 (age 3) to 6 (age 14) with mean 0 and standard deviation 1. The variable of interest is the movement of families between IMDs for sweeps 2 (age 3) and 3 (age 5). The analysis also controls for child's age in months at sweeps 6 (age 14), current period IMD level. The lagged period controls used consist of respective lagged period (sweep 2) outcomes, sweep 6 IMD. Child and household level controls consist of child's and mother's ethnicity, number of siblings, OECD weighted household income, mother's highest educational qualification, language spoken in household, mother having long term illness, child having long term illness, whether mother is born in the UK, child birth-weight, whether single or two parent household, age of mother at birth of child.
*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

FIGURE 3.3: The change in child vocabulary scores between sweeps 2 (age 3) and 6 (age 14) for families moving/staying at sweep 3 (age 5).

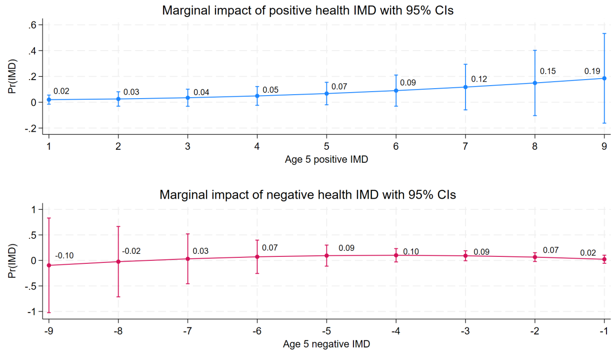
(a) Overall IMD



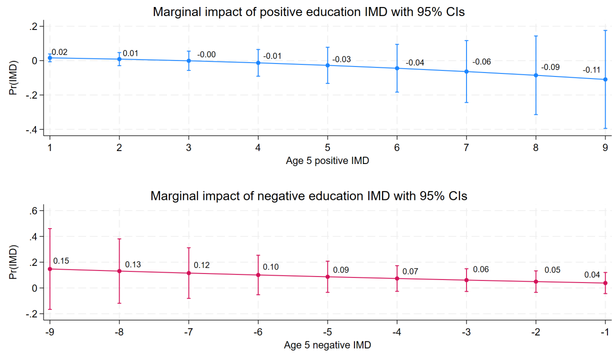
(b) Income IMD



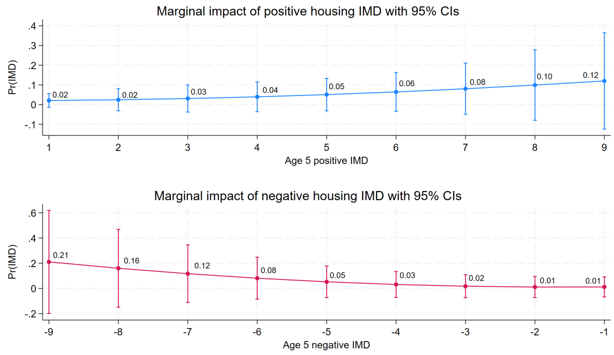
(c) Health IMD



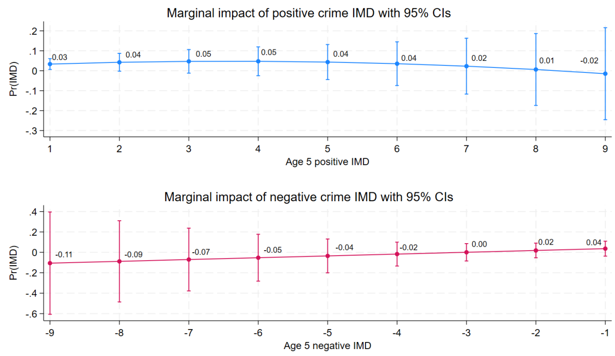
(d) Education IMD



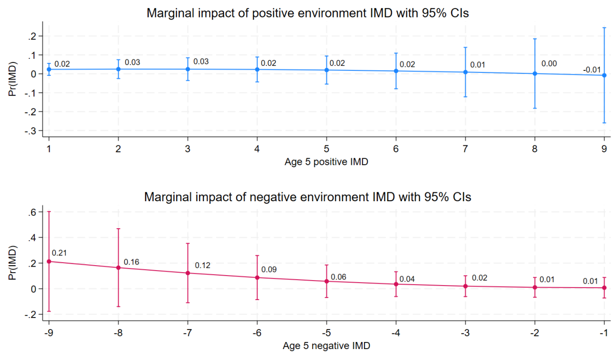
(e) Housing IMD



(f) Crime IMD



(g) Environment IMD



Notes.-The author's calculations are based on the families moving between different IMDs or staying at same IMD at age 5. The coefficients capture the change in IMD as a function of families moving between different levels of IMD. The difference in IMD scale is converted to absolute values so as to compare it with the families having positive difference in IMD at age 5. The outcome variable is the change in the Z vocabulary scores between ages 3 and 14 years.

Overall, there is an increase in vocabulary scores when one moves down the overall IMD deciles and by age 11. But, there is a marginal decline in the vocabulary performance of the movement of the households. This negative impact on the vocabulary scores is attributed to the change in the learning environment availability of schools, which makes it difficult for a child to adapt to a new school curriculum; this holds for health, housing and crime IMDs. Movements along the lower deciles of crime IMD have negative impact on child's vocabulary development as households are reluctant to send their children to socialise and, hence, limits their interaction to being at home, which could play a key role as this negative impact decreases as one moves to lower crime IMDs, which holds true throughout ages 7 to 14. Conversely, as households move to lower deciles of income IMD there is a decrease in the vocabulary scores compared to those moving to areas having higher decile of the income IMD. This is in line with previous studies by Dearden et al. (2011) & Kelly et al. (2011) suggest a positive impact of income on vocabulary development, highlighting that households moving to higher deciles of income IMD arrange extra classes and activities that contribute to the child's vocabulary development. There is a negative impact on the vocabulary scores of moving to areas having higher deciles of health IMD; the decline in the vocabulary scores is attributed to the negative impact of longstanding maternal poor health on the child vocabulary development. Also, households changing residence and moving to 1 unit higher decile of education IMD witness a decline in the vocabulary scores at age 7. This is due to the induced significant negative impact of poor school education on the vocabulary development of a child at age 7. Hence, there exists a negative impact on vocabulary development as one moves to areas having lower decile of overall IMD than the current. The negative impact is prominent for the individual IMDs. The, household movements (upwards/downward IMD) from 4 IMD decile units onwards lead to a noticeable change (positive /negative) on vocabulary scores throughout ages 7 to 14.

3.5.2 Internalising scores

The combined impact of movement to different deciles of overall IMD on internalising outcomes is shown in column(1) of tables 3.5 to 3.7 for ages 7,11 and 14. Throughout the household mobility between ages 3 and 7,11 and 14, there is no significant impact of change in overall IMD deciles for internalising scores and also no significant impact of moving between lower overall IMD deciles for behavioural scores. But a significant negative impact of 0.08 standard deviations on age 7 behavioural scores when moving down the overall IMD deciles at age 7. At age 11, there is a significant negative impact of 0.08 s.d. when moving across environment IMD deciles but a significant positive impact of 0.10 s.d. on behavioural scores when moving to areas having lower decile of the environment IMD. There is no impact of household mobility on behavioural scores as a result of movement along individual IMD deciles. Hence, moving to a lower environment IMD decile with poorer housing conditions and a favourable area for child development creates behavioural problems in the child as shown by the increase in the total difficulties score when moving to lower IMDs. However, by age 14, this negative impact on behavioural issues subsides. Also, by age 11 and 14 there is no significant impact of the change in the health IMD deciles. The insignificant impact of movements to areas along different deciles of health IMD could be because by the ages of 11 and 14, the child has acquired certain behavioural traits which are less dependent on health conditions than in younger years.

TABLE 3.5: The change in child internalising(total difficulty) between sweeps 2 (age 3) and 4 (age 7) for families moving/staying at sweep 3 (age 5).

	Overall Indv. IMDs						
	All reason (1)	Inc. (2)	Hlth. (3)	Edu. (4)	House (5)	Crime (6)	Env. (7)
OLS with controls							
Δ(IMD) (Base=Stayed at same IMD or different IMD)							
Change in IMD	0.062	0.062	-0.085*	0.012	-0.035	-0.014	0.005
	(0.047)	(0.044)	(0.049)	(0.022)	(0.041)	(0.033)	(0.024)
Change in IMD^2	-0.008	-0.009	0.015*	-0.007**	0.002	-0.000	-0.006*
	(0.008)	(0.008)	(0.009)	(0.003)	(0.008)	(0.005)	(0.003)
-ve Δ in IMD	0.201	0.050	0.018	-0.135	0.026	0.079	-0.016
movement(Base=+ve change)	(0.221)	(0.108)	(0.096)	(0.095)	(0.090)	(0.090)	(0.105)
Change in IMD for lower IMD	0.119	0.065	0.054	-0.082***	0.081	0.057	0.007
(Δ M#low IMD)	(0.148)	(0.064)	(0.084)	(0.027)	(0.062)	(0.044)	(0.042)
Change in IMD^2 for	0.025	0.023**	-0.032**	0.010**	0.002	-0.008	0.004
lower IMD (Δ M# ΔM#low IMD)	(0.020)	(0.011)	(0.013)	(0.004)	(0.012)	(0.007)	(0.009)
Lagged z test scores (sweep2)	-0.467***	-0.469***	-0.068***	-0.068***	-0.068***	-0.070***	-0.069***
	(0.016)	(0.016)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Constant	10.955	11.129	11.789	11.948	11.945	11.876	11.194
	(9.101)	(9.096)	(9.706)	(9.639)	(9.647)	(9.638)	(9.666)
R^2	0.186	0.187	0.085	0.088	0.085	0.086	0.085
N	6,058	6,058	6,058	6,058	6,058	6,058	6,058

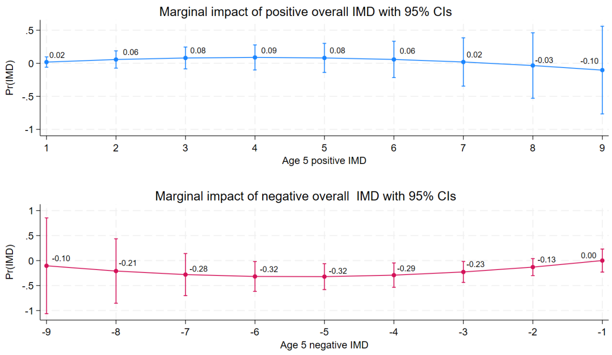
Notes.- The robust standard errors are in parenthesis. Child's age in months is controlled. Outcome is difference in standardised Z total difficulties scores between sweeps 2 (age 3) to 4 (age 7) with mean 0 and standard deviation 1. The variable of interest is the movement of families between IMDs for sweeps 2 (age 3) and 3 (age 5). The analysis also controls for child's age in months at sweeps 4 (age 7), current period IMD level. The lagged period controls used consist of respective lagged period (sweep 2) outcomes, sweep 4 IMD. Child and household level controls consist of a child's and mother's ethnicity, number of siblings, OECD weighted household income, mother's highest educational qualification, language spoken in household, mother having long term illness, child having long term illness, whether mother is born in the UK, child birth-weight, whether single or two parent household, age of mother at birth of child.
*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Marginal impact on the change in the internalising scores.
The marginal impact of moving to higher and lower IMDs on internalising scores is shown in figures 3.4 to 3.6 for ages 7 to 14, for each unit change in the IMD decile. Panel (a) of figure 3.4 shows that as the households move further down 3,4,5 & 6 overall IMD deciles, there is a significant decline in the behavioural scores indicating an improvement in behavioural development at age 7. But, there exists no impact of movements along deciles of overall IMD at ages 11 and 14 behavioural development outcomes. At age 7 decline in behavioural scores when moving down by 5,4,3 & 2 income IMD deciles. Next, moving up by 1, 2 deciles of health IMD has negative impact on age 7 behavioural outcomes, also moving down by 9,8,7 & 6 deciles of health IMDs has declining impact on age 7 behavioural scores. At age 14, there is a decline in behavioural score when moving down 9,8,7 & 6 health IMD deciles. At age 7 increase in behavioural scores as households move to areas having lower 9,8,7 & 6 deciles of education IMD. There is a significant decline in age 7 behavioural score when moving up by 1 unit housing IMD decile. Also, at age 11 there is a decline in behavioural score when moving to areas having higher 5,6,7 & 8 deciles of housing IMDs. For crime IMDs, at age 7, there is a decline in behavioural score when moving higher up by 1 unit decile, and also when moving down by 9,8,7,6,5 & 4 units of crime IMD decile. Also, decline in behavioural score at age 14 when moving to areas having lower 8,7,6 & 5 deciles of crime IMD. At age 7, decline in behavioural score when moving to areas having higher 1,6,7,8 & 9 deciles of environment IMD. Similarly, at age 11, decline in behavioural score due to moving higher up by 1,2 & 3 environment IMD deciles.

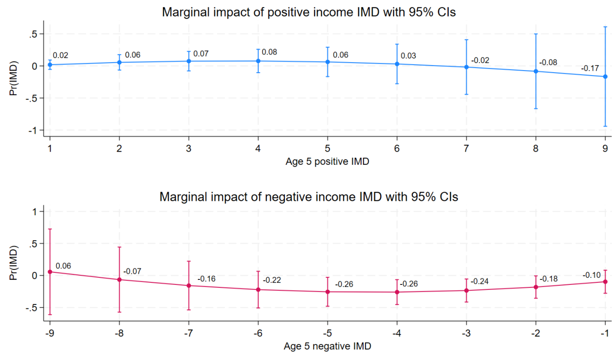
Overall, there is a negative impact on the behavioural scores as a child moves down the lower IMD, but, in case of moving to lower health IMD deciles there is a positive marginal significant impact on the increase in behavioural scores, highlighting more behavioural issues, in some households. For all the individual IMD deciles the behavioural scores increase as one moves down the lower IMD deciles and slightly increase when moving down the education IMD deciles. Also, households moving down the education IMD deciles experience a significant decline in behavioural scores. Gambaro & Joshi (2016) finds a significant impact of 0.37 standard deviations increase in the behavioural issues of moving within bottom 30% of deprived area. Gambaro & Joshi (2016) highlighted the impact of maternal depression that could result in a negative impact on the behavioural development of the child, although some families show a negative impact on behavioural issues when moving into lower deciles of overall IMD, which could be due to the fact highlighted by Gambaro & Joshi (2016) that children moving to poor areas compensate for the negative impact of being in lower decile of overall IMD by witnessing an improvement in their home environment, initiated by their parents. Throughout ages 7 to 14, there is a decrease in behavioural issues as the households move further up the higher IMD deciles or further down the lower IMD deciles.

FIGURE 3.4: The change in child internalising(total difficulty) scores between sweeps 2 (age 3) and 4 (age 7) for families moving/staying at sweep 3 (age 5).

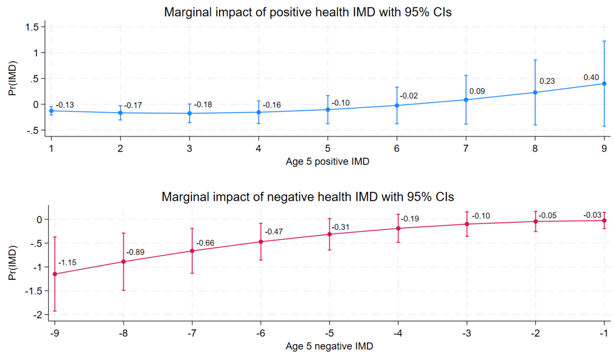
(a) Overall IMD



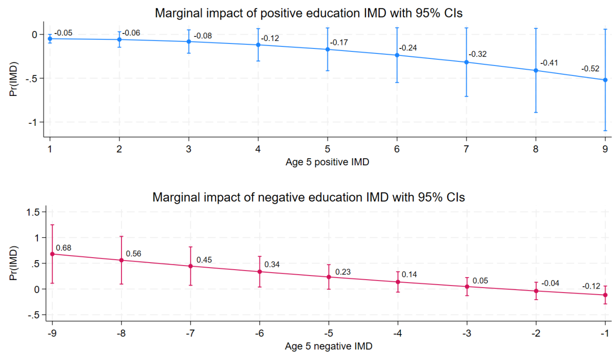
(b) Income IMD



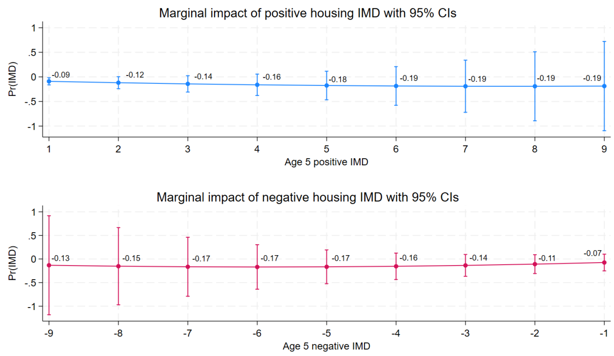
(c) Health IMD



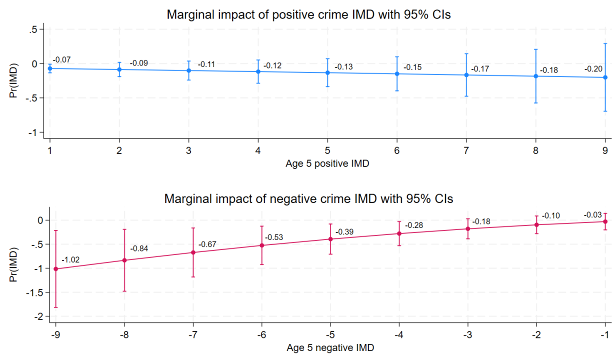
(d) Education IMD



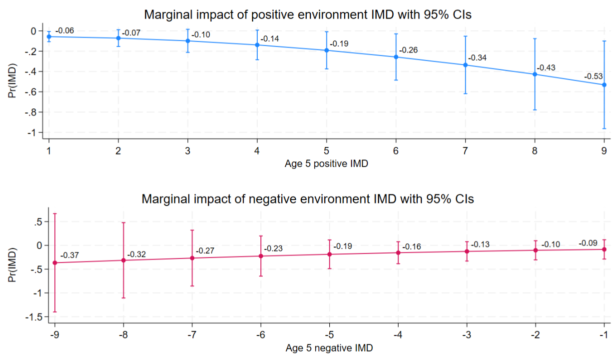
(e) Housing IMD



(f) Crime IMD



(g) Environment IMD



Notes.- The author's calculations are based on the families moving between different IMDs or staying at same IMD at age 5. The coefficients capture the change in IMD as a function of families moving between different levels of IMD. The difference in IMD scale is converted to absolute values so as to compare it with the families having positive difference in IMD at age 5. The outcome variable is the change in the Z total difficulties scores between ages 3 and 7 years.

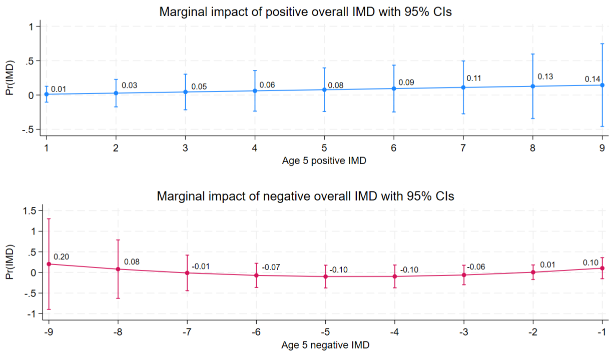
TABLE 3.6: The change in child internalising(total difficulty) between sweeps 2 (age 3) and 5 (age 11) for families moving/staying at sweep 3 (age 5).

	Overall Indv. IMDs						
	All reason (1)	Inc. (2)	Hlth. (3)	Edu. (4)	House (5)	Crime (6)	Env. (7)
OLS with controls							
Δ(IMD) (Base=Stayed at same IMD or different IMD)							
Change in IMD	0.016	0.054	-0.070	0.003	0.002	-0.019	-0.076**
	(0.068)	(0.052)	(0.071)	(0.038)	(0.042)	(0.031)	(0.030)
Change in IMD^2	0.000	-0.020*	0.005	-0.009	-0.013	-0.000	0.011**
	(0.009)	(0.011)	(0.013)	(0.007)	(0.009)	(0.004)	(0.005)
-ve Δ in IMD	0.238	0.041	0.036	-0.039	0.074	0.047	0.088
movement(Base=+ve change)	(0.265)	(0.095)	(0.094)	(0.091)	(0.083)	(0.090)	(0.097)
Change in IMD for lower IMD	0.130	-0.020	0.012	-0.015	0.014	0.027	0.099**
(ΔM#low IMD)	(0.189)	(0.066)	(0.098)	(0.043)	(0.057)	(0.039)	(0.040)
Change in IMD^2 for	0.016	0.025*	-0.024	0.015*	0.009	-0.001	-0.013
lower IMD (ΔM#ΔM#low IMD)	(0.025)	(0.013)	(0.019)	(0.008)	(0.013)	(0.007)	(0.008)
Lagged z test scores (sweep2)	-0.491***	0.050***	0.049***	0.050***	0.048***	0.049***	0.047***
	(0.018)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Constant	0.045	-3.704	-3.320	-3.611	-4.049	-3.077	-2.534
	(11.835)	(10.712)	(10.885)	(10.855)	(10.906)	(10.798)	(10.765)
R^2	0.200	0.076	0.075	0.077	0.077	0.075	0.077
N	5,274	5,274	5,274	5,274	5,274	5,274	5,274

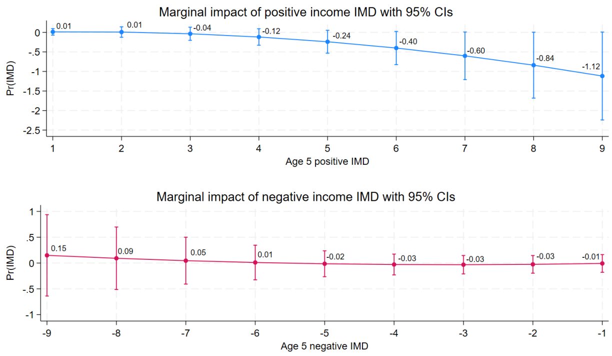
Notes.- The robust standard errors are in parenthesis. Child's age in months is controlled. Outcome is difference in standardised Z total difficulties scores between sweeps 2 (age 3) to 5 (age 11) with mean 0 and standard deviation 1. The variable of interest is the movement of families between IMDs for sweeps 2 (age 3) and 3 (age 5). The analysis also controls for child's age in months at sweeps 5 (age 11), current period IMD level. The lagged period controls used consist of respective lagged period (sweep2) outcomes, sweep 5 IMD. Child and household level controls consist of a child's and mother's ethnicity, number of siblings, OECD weighted household income, mother's highest educational qualification, language spoken in household, mother having long term illness, child having long term illness, whether mother is born in the UK, child birth-weight, whether single or two parent household, age of mother at birth of child.
*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

FIGURE 3.5: The change in child internalising(total difficulty) scores between sweeps 2 (age 3) and 5 (age 11) for families moving/staying at sweep 3 (age 5).

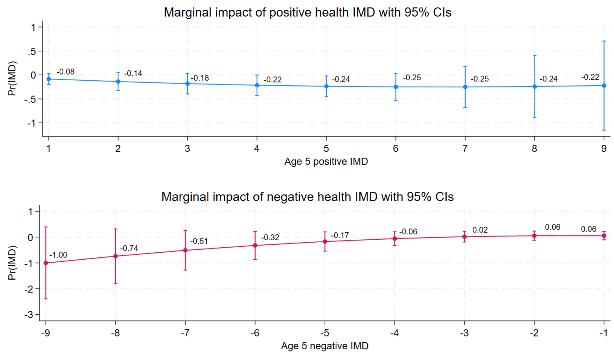
(a) Overall IMD



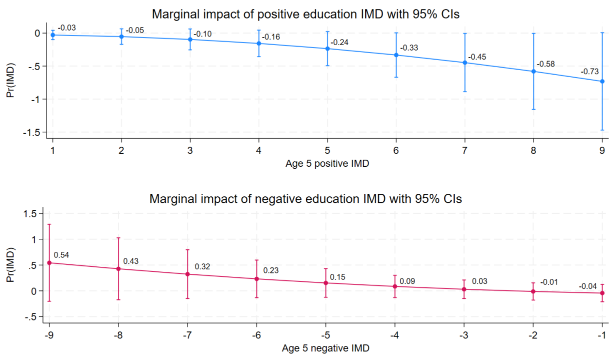
(b) Income IMD



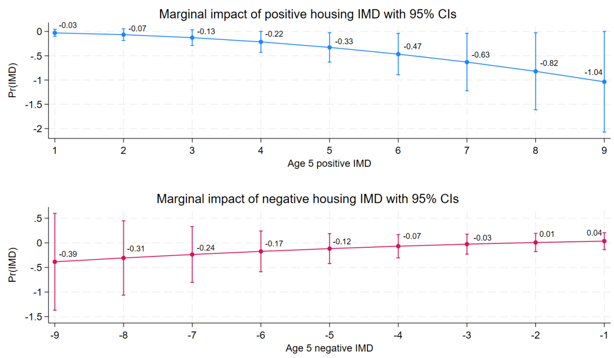
(c) Health IMD



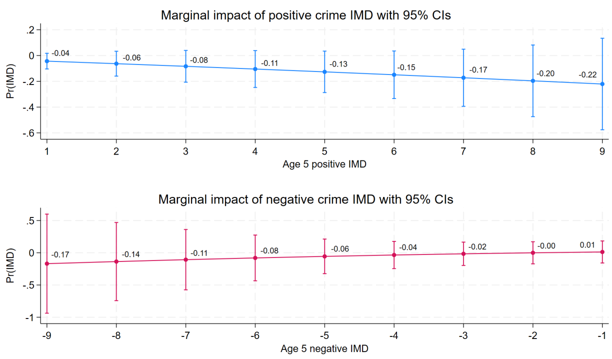
(d) Education IMD



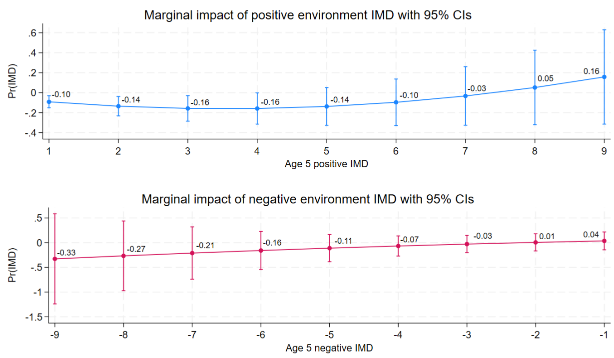
(e) Housing IMD



(f) Crime IMD



(g) Environment IMD



Notes.- The author's calculations are based on the families moving between different IMDs or staying at same IMD at age 5. The coefficients capture the change in IMD as a function of families moving between different levels of IMD. The difference in IMD scale is converted to absolute values so as to compare it with the families having positive difference in IMD at age 5. The outcome variable is the change in the Z total difficulties scores between ages 3 and 11 years.

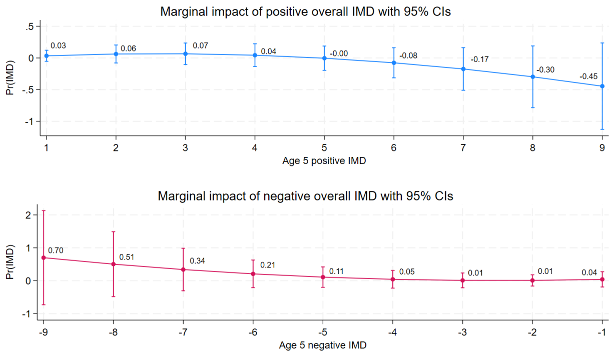
TABLE 3.7: The change in child internalising(total difficulty) between sweeps 2 (age 3) and 6 (age 14) for families moving/staying at sweep 3 (age 5).

	Overall Indv. IMDs						
	All reason (1)	Inc. (2)	Hlth. (3)	Edu. (4)	House (5)	Crime (6)	Env. (7)
OLS with controls							
Δ(IMD) (Base=Stayed at same IMD or different IMD)							
Change in IMD	0.066	0.026	-0.026	-0.015	-0.045	-0.013	-0.019
	(0.054)	(0.041)	(0.050)	(0.022)	(0.042)	(0.036)	(0.035)
Change in IMD^2	-0.013	-0.004	0.003	-0.009	0.006	0.001	0.004
	(0.009)	(0.007)	(0.009)	(0.007)	(0.008)	(0.006)	(0.007)
-ve Δ in IMD	0.127	-0.095	-0.101	-0.039	-0.010	0.060	-0.044
movement(Base=+ve change)	(0.241)	(0.099)	(0.094)	(0.091)	(0.090)	(0.093)	(0.097)
Change in IMD for lower IMD	0.015	-0.064	-0.133*	0.010	0.124*	0.055	0.007
(ΔM#low IMD)	(0.176)	(0.055)	(0.077)	(0.031)	(0.064)	(0.045)	(0.048)
Change in IMD^2 for	0.029	0.002	-0.044***	0.007	0.004	-0.006	-0.006
lower IMD(ΔM#ΔM#low IMD)	(0.025)	(0.010)	(0.015)	(0.006)	(0.011)	(0.008)	(0.010)
Lagged z test scores (sweep2)	-0.480***	0.117***	0.118***	0.120***	0.116***	0.118***	0.118***
	(0.017)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Constant	13.375	0.533	-0.100	-0.354	0.454	0.543	1.075
	(15.647)	(14.692)	(14.730)	(14.666)	(14.745)	(14.759)	(14.667)
R^2	0.203	0.189	0.190	0.188	0.190	0.189	0.189
N	4,433	4,433	4,433	4,433	4,433	4,433	4,433

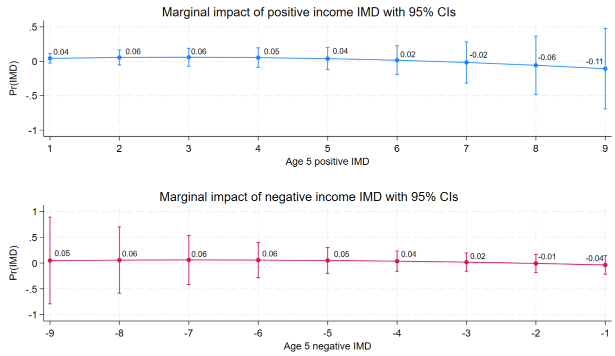
Notes.- The robust standard errors are in parenthesis. Child's age in months is controlled. Outcome is difference in standardised Z total difficulties scores between sweeps 2 (age 3) to 6 (age 14) with mean 0 and standard deviation 1. The variable of interest is the movement of families between IMDs for sweeps 2 (age 3) and 3 (age 5). The analysis also controls for child's age in months at sweeps 6 (age 14), current period IMD level. The lagged period controls used consist of respective lagged period (sweep 2) outcomes, sweep 6 IMD. Child and household level controls consist of a child's and mother's ethnicity, number of siblings, OECD weighted household income, mother's highest educational qualification, language spoken in household, mother having long term illness, child having long term illness, whether mother is born in the UK, child birth-weight, whether single or two parent household, age of mother at birth of child.
*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

FIGURE 3.6: The change in child internalising(total difficulty) scores between sweeps 2 (age 3) and 6 (age 14) for families moving/staying at sweep 3 (age 5).

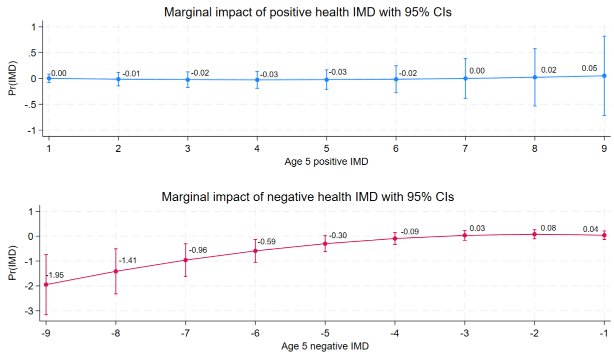
(a) Overall IMD



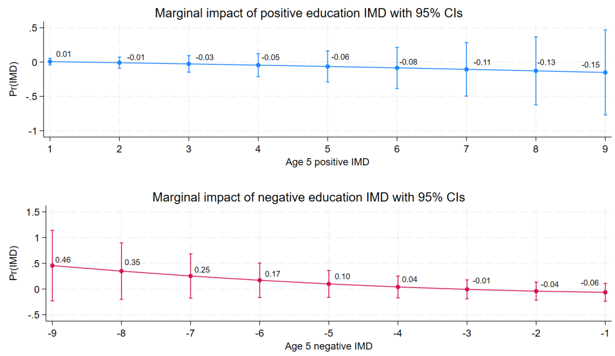
(b) Income IMD



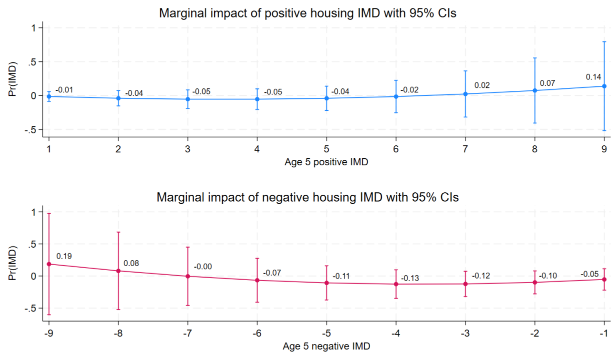
(c) Health IMD



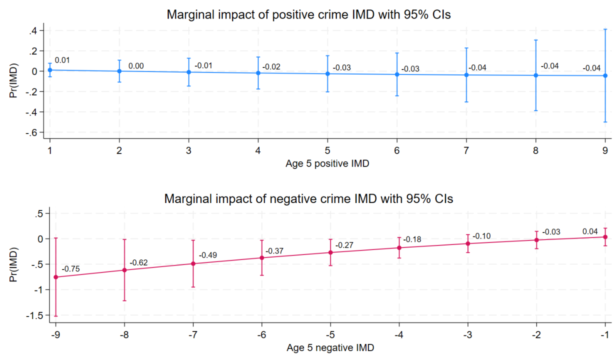
(d) Education IMD



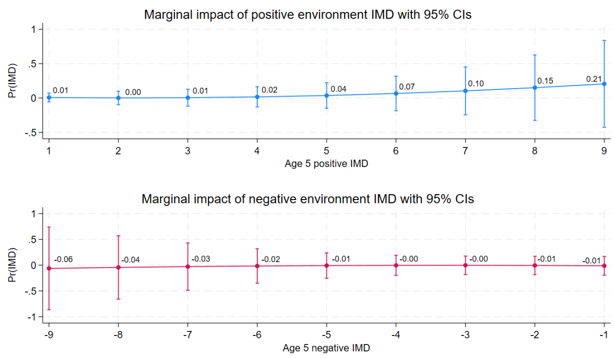
(e) Housing IMD



(f) Crime IMD



(g) Environment IMD



Notes.- The author's calculations are based on the families moving between different IMDs or staying at same IMD at age 5. The coefficients capture the change in IMD as a function of families moving between different levels of IMD. The difference in IMD scale is converted to absolute values so as to compare it with the families having positive difference in IMD at age 5. The outcome variable is the change in the Z total difficulties scores between ages 3 and 14 years.

3.5.3 Externalising scores

The combined impact of movements to different deciles of overall IMD is shown in column(1) of tables 3.8, 3.9 & 3.10 for the prosocial scores. At age 7, there is no significant impact of change in the overall IMD deciles for externalising scores and also there is no significant impact of moving between lower deciles of overall IMD. This also corresponds for the household mobility impact between ages 3 to 11 and 3 to 14 years. In case of the individual IMD impact, there is a significant negative impact of 0.08 s.d. on externalising scores when moving down deciles of education IMD at age 7 years. But, there is a significant decline of 0.08 s.d. on prosocial behaviour when moving across the deciles of environment IMD, whereas, a significant improvement in prosocial scores by 0.10 s.d. when moving down deciles of the environment IMD. At age 14, there is a significant improvement in externalising behaviour by 0.13 s.d. when moving down deciles of the housing IMD. The significant improvement while moving down deciles of the environment IMD for the prosocial skills indicate that the prosocial behaviour inculcated at an early age remains prominent and is less impacted by the change in the surroundings due to household mobility this is confirmed that by age 14, there exists no impact of household mobility along deciles of the environment IMD, impacting prosocial development.

TABLE 3.8: The change in child externalising(prosocial) between sweeps 2 (age 3) and 4 (age 7) for families moving/staying at sweep 3 (age 5).

	Overall						
	Indv. IMDs						
	All reason (1)	Inc. (2)	Hlth. (3)	Edu. (4)	House (5)	Crime (6)	Env. (7)
OLS with controls							
Δ(IMD)							
(Base=Stayed at same IMD							
or different IMD)							
Change in IMD		-0.050	-0.066*	-0.089*	0.012	-0.038	-0.015
0.005							
	[0.039]	[0.035]	[0.049]	[0.021]	[0.041]	[0.034]	[0.024]
Change in IMD^2	0.007	0.009	0.016*	-0.007**	0.003	-0.000	-0.007**
	[0.006]	[0.006]	[0.009]	[0.003]	[0.008]	[0.005]	[0.003]
-ve Δ in IMD	0.090	0.036	0.016	-0.143	0.022	0.076	-0.017
movement(Base=+ve change)	[0.202]	[0.088]	[0.098]	[0.095]	[0.091]	[0.090]	[0.106]
Change in IMD for lower IMD	0.083	0.042	0.064	-0.081***	0.085	0.057	0.004
(ΔM#low IMD)	[0.141]	[0.047]	[0.086]	[0.027]	[0.062]	[0.045]	[0.041]
Change in IMD^2 for	-0.002	-0.013	-0.032**	0.010**	0.002	-0.008	0.004
lower IMD(ΔM#ΔM#low IMD)	[0.021]	[0.009]	[0.013]	[0.004]	[0.012]	[0.008]	[0.009]
Lagged z test scores (sweep2)	-0.656***	-0.656***	0.001	-0.001	0.000	0.001	0.000
	(0.014)	(0.014)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Constant	-6.933	-7.055	11.119	11.164	11.254	11.120	10.585
	(7.762)	(7.746)	(9.755)	(9.681)	(9.690)	(9.688)	(9.718)
R^2	0.363	0.363	0.082	0.084	0.081	0.082	0.082
N	6,058	6,058	6,058	6,058	6,058	6,058	6,058

Notes.- The robust standard errors are in parenthesis. Child's age in months is controlled. Outcome is difference in standardised Z externalising scores between sweeps 2 (age 3) to 4 (age 7) with mean 0 and standard deviation 1. The variable of interest is the movement of families between IMDs for sweeps 2 (age 3) and 3 (age 5). The analysis also controls for child's age in months at sweeps 4 (age 7), current period IMD level. The lagged period controls used consist of respective lagged period(sweep2) outcomes, sweep4 IMD. Child and household level controls consist of child's and mother's ethnicity, number of siblings, OECD weighted household income, mother's highest educational qualification, language spoken in household, mother having long term illness, child having long term illness, whether mother is born in the UK, child birth-weight, whether single or two parent household, age of mother at birth of child.

*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Marginal impact of the change in the externalising scores.
The marginal impact of moving to higher and lower IMDdeciles in case of externalising scores is shown in figures 3.7 to 3.9 for ages 7 to 14 for each unit change in the IMDs. Panel (a) of figures 3.7 to 3.9 shows that overall at ages 7 there is a significant decline in the prosocial scores when households move higher up by 1,2,3,4 & 5 IMD deciles. There is no impact in age 11 prosocial scores, but there is a significant decline in externalising conduct when households move lower by 7,8 & 9 overall IMD deciles. At age 7, significant decline when moving higher up by 1,2,3,4 & 5 income IMD deciles. At age 7, a significant decline in externalising behaviour when moving higher up by 1 & 2 health IMD deciles. Also, decline in age 7 prosocial scores when moving down by 9,8,7 & 6 health IMD deciles. At age 11, decline when moving higher up by 5 health IMD deciles. At age 14, decline in externalising behaviour when moving down by 9,8,7 & 6 health IMD deciles. At age 7, improvement in prosocial conduct when moving down by 9,8,7 & 6 deciles of education IMD. At age 7, decline in prosocial scores when moving up by 1 & 2 deciles of housing IMD. At age 11, decline in prosocial conduct when moving up by 5,6,7 & 8 deciles of housing IMD. At age 7, decline in prosocial behaviour when moving higher up by 1 crime IMD decile. Also, a decline when moving down by 9,8,7,6,5 & 4 crime IMD decile. At age 14, decline in prosocial conduct when moving down by 8,7,6 & 5 deciles of crime IMD. At age 7, decline in prosocial behaviour when moving higher by 1,5,6,7,8 & 9 deciles of environment IMD. At age 11, decline in externalising conduct when moving higher up by 1,2,3 & 4 deciles of environment IMD. Hence, throughout, ages 7 to 14, the households moving down the lower deciles of IMDs show an improvement in prosocial behaviour except for moving down the crime and environment IMDs. Among the individual IMDs, the movements along the deciles of crime and environment IMDs play a significant role in determining the prosocial behaviour of the child. It is also validated by the fact that prosocial behaviour is determined by interaction with peers hence, moving down to lower deciles of crime and environment IMDs increases the negative impact on prosocial behaviour.

Consequently, for overall prosocial IMDs, there is a positive impact on the prosocial skills as households move to lower IMD deciles. Among the individual IMD decile movements, the income and health IMDs have a significant declining prosocial behaviour as a child moves to higher decile of income and lower decile of health IMD. Whereas, relative to the households moving into higher deciles of IMDs, there is a positive impact on the externalising scores. The declining externalising scores highlights the increasing external behaviour issues when moving to lower deciles of education IMD. Gambaro & Joshi (2016) & Gambaro et al. (2017) find a significant negative association of 0.30 standard deviations with the externalising behaviour of the child moving to bottom 30% of area (deprived) at age 7. Although the households moving to areas having lower deciles of income and education IMDs show good behaviour conduct as shown by the higher prosocial scores. This is because prosocial behaviours are not affected by the availability of the resources but those are the skills that the child inculcates at home and is reflected throughout their life.

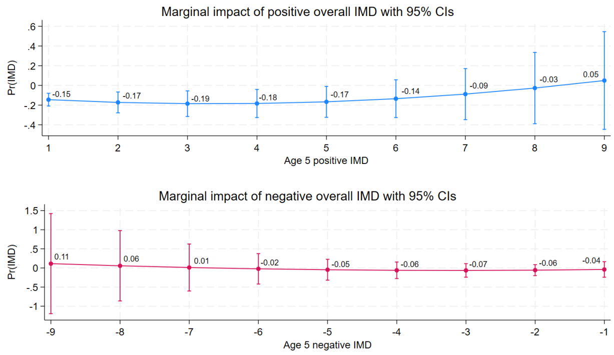
TABLE 3.9: The change in child externalising(prosocial) between sweeps 2 (age 3) and 5 (age 11) for families moving/staying at sweep 3 (age 5).

	Overall **Indv. IMDs**						
	All reason	Inc.	Hlth.	Edu.	House	Crime	Env.
OLS with controls							
Δ(IMD) (Base=Stayed at same IMD or different IMD)							
Change in IMD	-0.079	0.050	-0.072	0.004	0.001	-0.017	-0.078**
	(0.063)	(0.051)	(0.072)	(0.039)	(0.042)	(0.031)	(0.030)
Change in IMD^2	0.014	-0.019*	0.006	-0.009	-0.013	-0.000	0.011**
	(0.009)	(0.011)	(0.013)	(0.007)	(0.010)	(0.004)	(0.005)
-ve Δ in IMD	0.199	0.049	0.041	-0.034	0.072	0.050	0.095
movement(Base=+ve change)	(0.192)	(0.096)	(0.095)	(0.092)	(0.084)	(0.091)	(0.097)
Change in IMD for lower IMD	0.170	-0.013	0.015	-0.017	0.010	0.029	0.102**
(ΔM#low IMD)	(0.140)	(0.067)	(0.100)	(0.045)	(0.058)	(0.039)	(0.041)
Change in IMD^2 for lower IMD (ΔM#ΔM#low IMD)	-0.001	0.025*	-0.025	0.015*	0.009	-0.000	-0.014*
	(0.020)	(0.014)	(0.019)	(0.008)	(0.013)	(0.007)	(0.008)
Lagged z test scores (sweep2)	-0.721***	-0.053***	-0.052***	-0.052***	-0.052***	-0.051***	-0.053***
	(0.015)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Constant	-12.901	-1.891	-1.507	-1.756	-2.308	-1.310	-0.747
	(12.491)	(11.007)	(11.165)	(11.164)	(11.187)	(11.089)	(11.039)
R^2	0.417	0.076	0.076	0.077	0.077	0.076	0.077
N	5,274	5,274	5,274	5,274	5,274	5,274	5,274

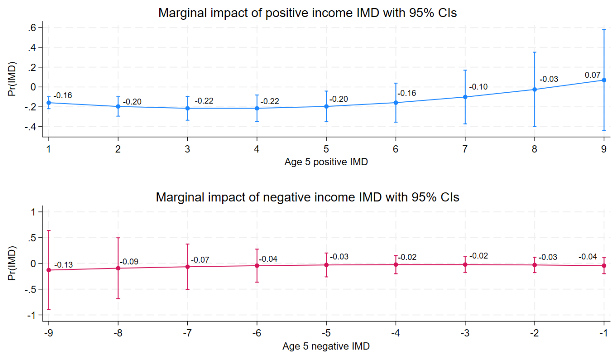
Notes.- The robust standard errors are in parenthesis. Child's age in months is controlled. Outcome is difference in standardised Z externalising scores between sweeps 2 (age 3) to 5 (age 11) with mean 0 and standard deviation 1. The variable of interest is the movement of families between IMDs for sweeps 2 (age 3) and 3 (age 5). The analysis also controls for child's age in months at sweeps 5 (age 11), current period IMD level. The lagged period controls used consist of respective lagged period (sweep 2) outcomes, sweep 5 IMD. Child and household level controls consist of child's and mother's ethnicity, number of siblings, OECD weighted household income, mother's highest educational qualification, language spoken in household, mother having long term illness, child having long term illness, whether mother is born in the UK, child birth-weight, whether single or two parent household, age of mother at birth of child.
*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

FIGURE 3.7: The change in child externalising(prosocial) scores between sweeps 2 (age 3) and 4 (age 7) for families moving/staying at sweep3 (age 5).

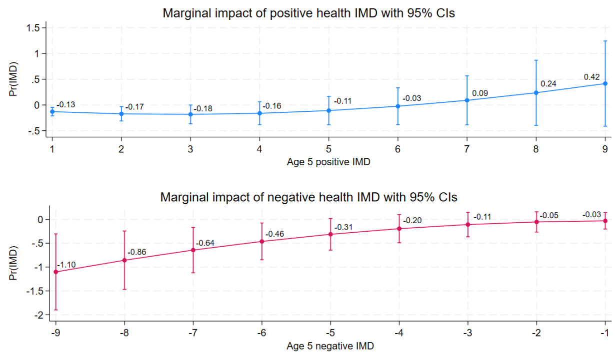
(a) Overall IMD



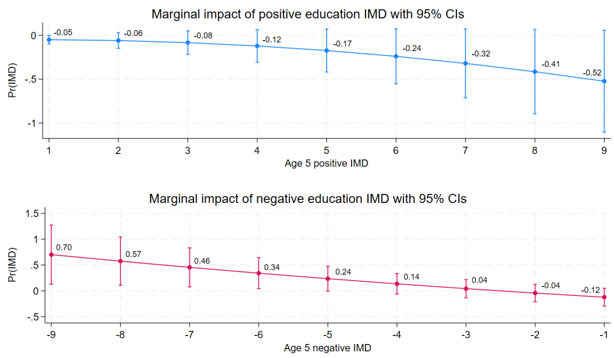
(b) Income IMD



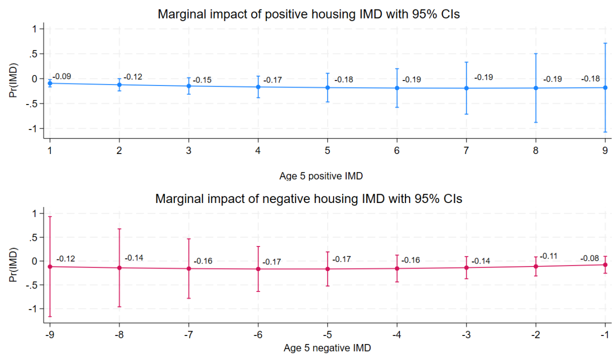
(c) Health IMD



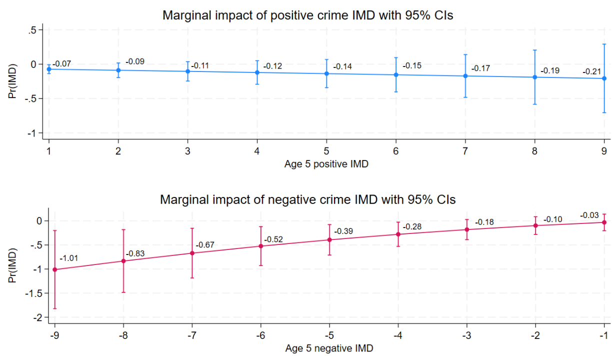
(d) Education IMD



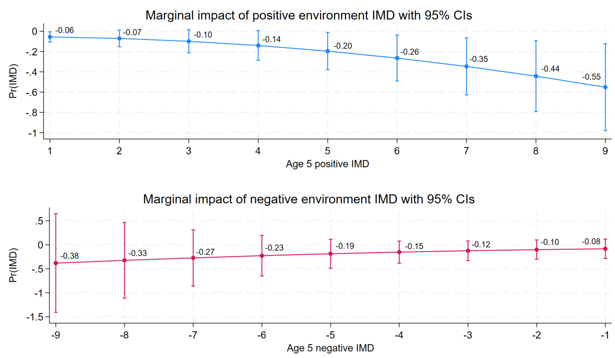
(e) Housing IMD



(f) Crime IMD



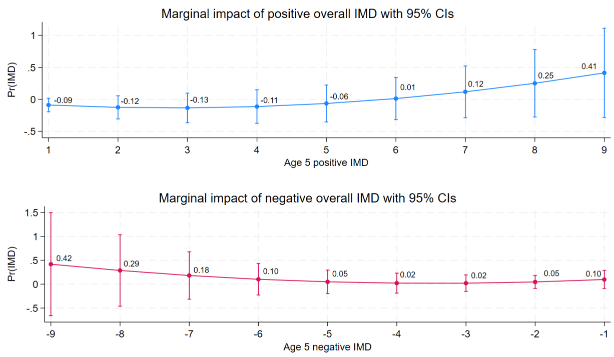
(g) Environment IMD



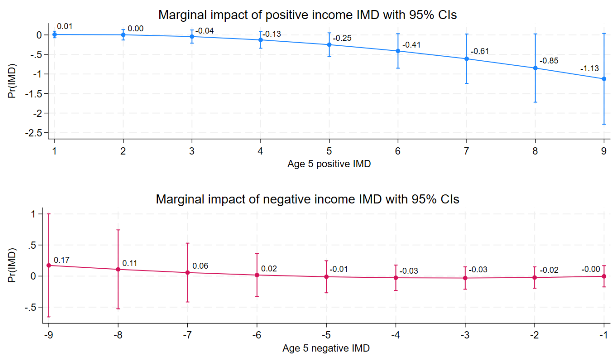
Notes.- The author's calculations are based on the families moving between different IMDs or staying at same IMD at age 5. The coefficients capture the change in IMD as a function of families moving between different levels of IMD. The difference in IMD scale is converted to absolute values so as to compare it with the families having positive difference in IMD at age 5. The outcome variable is the change in the Z externalising scores between ages 3 and 7 years.

FIGURE 3.8: The change in child externalising(prosocial) scores between sweeps 2 (age 3) and 5 (age 11) for families moving/staying at sweep3 (age 5).

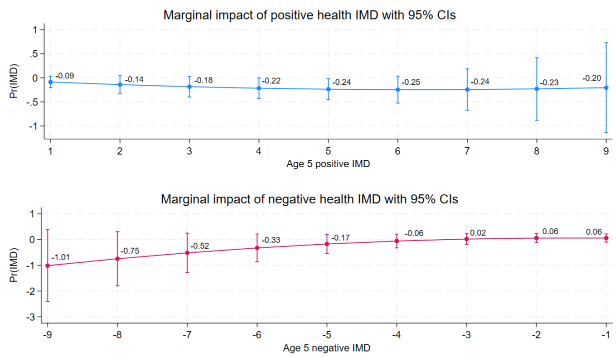
(a) Overall IMD



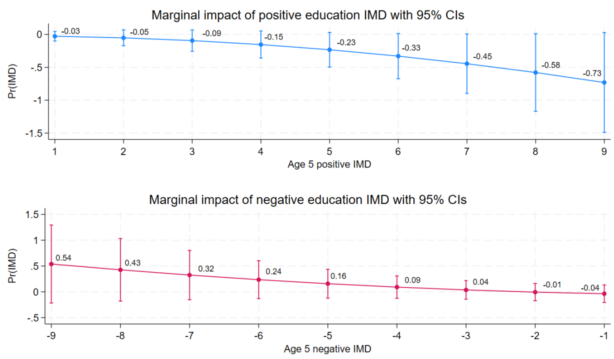
(b) Income IMD



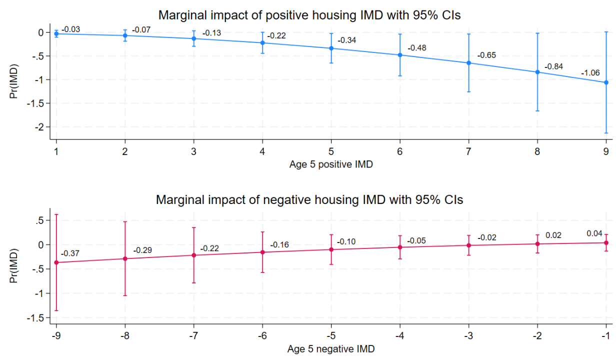
(c) Health IMD



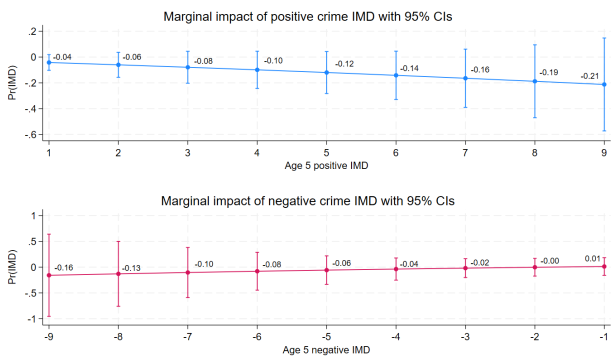
(d) Education IMD



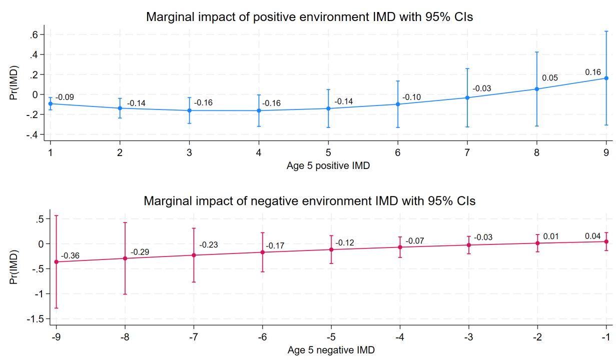
(e) Housing IMD



(f) Crime IMD



(g) Environment IMD



Notes.- The author's calculations are based on the families moving between different IMDs or staying at same IMD at age 5. The coefficients capture the change in IMD as a function of families moving between different levels of IMD. The difference in IMD scale is converted to absolute values so as to compare it with the families having positive difference in IMD at age 5. The outcome variable is the change in the Z externalising scores between ages 3 and 11 years.

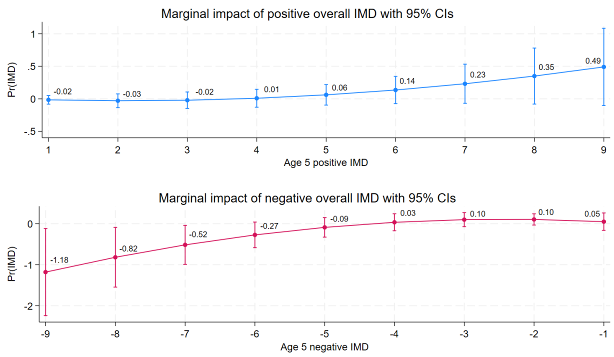
TABLE 3.10: The change in child externalising(prosocial) between sweeps 2 (age 3) and 6 (age 14) for families moving/staying at sweep 3 (age 5).

	Overall Indv. IMDs						
	All reason (1)	Inc. (2)	Hlth. (3)	Edu. (4)	House (5)	Crime (6)	Env. (7)
OLS with controls							
Δ(IMD) (Base=Stayed at same IMD or different IMD)							
Change in IMD	-0.047	0.021	-0.023	-0.015	-0.051	-0.010	-0.017
	(0.040)	(0.040)	(0.049)	(0.022)	(0.042)	(0.035)	(0.035)
Change in IMD^2	0.011	-0.004	0.003	-0.001	0.007	0.001	0.004
	(0.007)	(0.007)	(0.009)	(0.004)	(0.008)	(0.005)	(0.007)
-ve Δ in IMD	-0.086	-0.067	-0.075	-0.080	0.001	0.080	-0.020
movement(Base=+ve change)	(0.212)	(0.098)	(0.093)	(0.092)	(0.091)	(0.092)	(0.097)
Change in IMD for lower IMD	-0.097	-0.047	-0.121	0.009	0.127*	0.060	0.007
(ΔM#low IMD)	(0.143)	(0.053)	(0.076)	(0.031)	(0.066)	(0.044)	(0.048)
Change in IMD^2 for	-0.041**	0.002	-0.042***	0.007	0.003	-0.006	-0.007
lower IMD (ΔM#ΔM#low IMD)	(0.020)	(0.010)	(0.015)	(0.006)	(0.011)	(0.008)	(0.010)
Lagged z test scores (sweep2)	-0.578***	-0.098***	-0.098***	-0.098***	-0.098***	-0.098***	-0.098***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Constant	-23.248*	1.364	0.689	0.300	1.239	1.276	1.898
	(13.120)	(14.876)	(14.928)	(14.858)	(14.931)	(14.950)	(14.861)
R^2	0.316	0.185	0.186	0.184	0.186	0.185	0.185
N	4,433	4,433	4,433	4,433	4,433	4,433	4,433

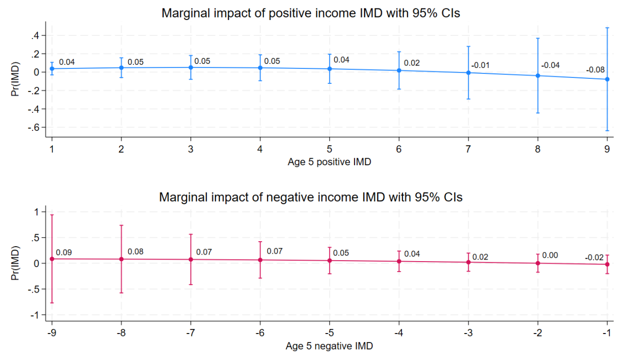
Notes.- The robust standard errors are in parenthesis. Child's age in months is controlled. Outcome is difference in standardised Z externalising scores between sweeps 2 (age 3) to 6 (age 14) with mean 0 and standard deviation 1. The variable of interest is the movement of families between IMDs for sweeps 2 (age 3) and 3 (age 5). The analysis also controls for child's age in months at sweeps 6 (age 14), current period IMD level. The lagged period controls used consist of respective lagged period (sweep 2) outcomes, sweep6 IMD. Child and household level controls consist of child's and mother's ethnicity, number of siblings, OECD weighted household income, mother's highest educational qualification, language spoken in household, mother having long term illness, child having long term illness, whether mother is born in the UK, child birth-weight, whether single or two parent household, age of mother at birth of child.
*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

FIGURE 3.9: The change in child externalising(prosocial) scores between sweeps 2 (age 3) and 6 (age 14) for families moving/staying at sweep3 (age 5).

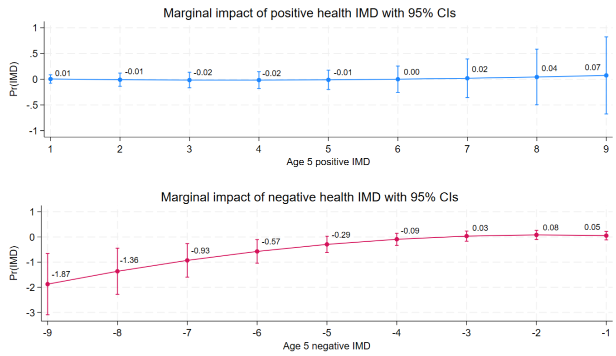
(a) Overall IMD



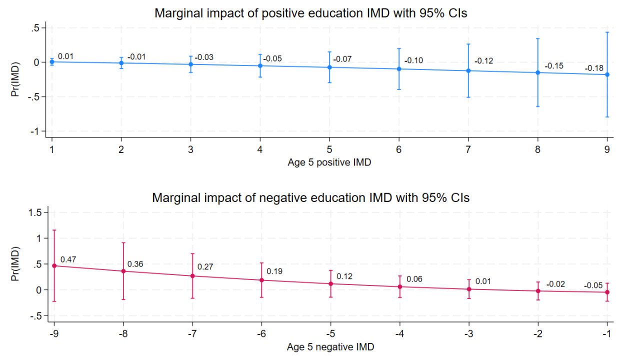
(b) Income IMD



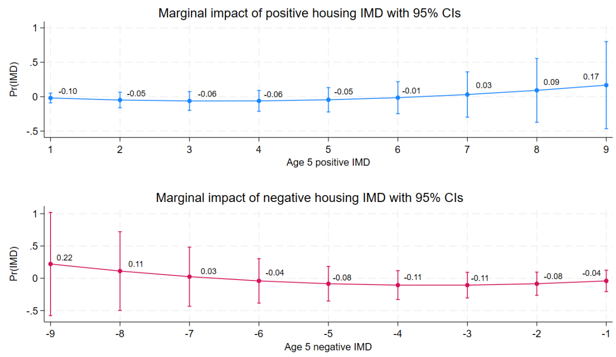
(c) Health IMD



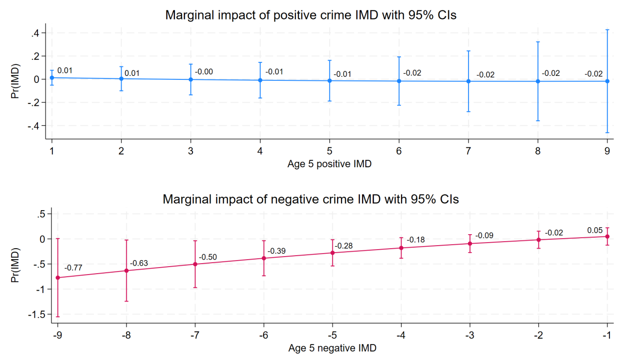
(d) Education IMD



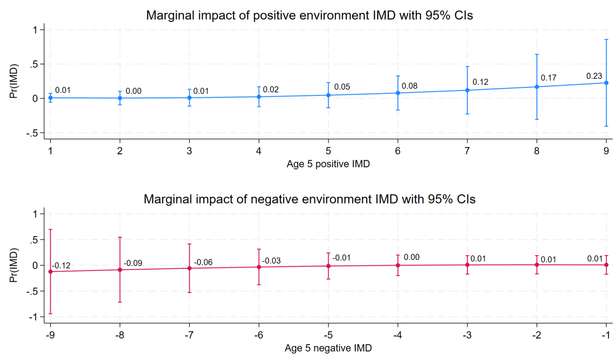
(e) Housing IMD



(f) Crime IMD



(g) Environment IMD



Notes.- The author's calculations are based on the families moving between different IMDs or staying at same IMD at age 5. The coefficients capture the change in IMD as a function of families moving between different levels of IMD. The difference in IMD scale is converted to absolute values so as to compare it with the families having positive difference in IMD at age 5. The outcome variable is the change in the Z externalising scores between ages 3 and 14 years.

3.6 Discussion

The study assess the age-specific changes in the vocabulary and behavioural development of the child in the form of observing the changes in child development scores between two periods at a time. The current study uses household movements along deciles of combined and individual IMDs. The overall age-specific analysis shows that there lies a significant negative impact of the changing residence at age 5 on the vocabulary scores. After observing the changing patterns as a result of movements along different deciles of individual IMD it is worth highlighting that there is a negative impact on the vocabulary scores as households move down the lower IMD deciles. Among the individual IMD analysis, the health IMD shows diverging impacts of moving to higher decile and lower decile IMDs on vocabulary scores. Maternal longstanding illness in the households moving down the lower decile of health IMD inflicts negative impact on the internalising and externalising scores. In the analysis the variable mother suffering from illness is included as control. For the internalising outcome, there is a significant positive impact of 0.11 s.d. of mothers suffering from long-term illness when households move along the deciles of overall and income IMDs for age group 3 to 7 years; a significant positive impact of 0.18 s.d. when moving along deciles of overall IMD for age groups 3 to 11 years; and a positive significant impact ranging from 0.08 to 0.11 for movements along deciles of overall and individual IMDs for age groups 3 to 14 years. In the case of externalising outcomes, the maternal illness impact ranges from 0.10 to 0.11 for movement along deciles of individual IMDs for age 3 to 14 years (the corresponding coefficients are available in tables B.10 to B.12 for internalising score and in table B.15 for externalising score.).

The movements along decile of overall IMD in case of internalising outcomes showed a negative impact on the internalising scores as one moves down the lower deciles of the IMD. The health IMD shows an adverse impact on behavioural issues by increase in scores when households move to lower deciles of IMD. The reasons for this increase in the internalising scores are accounted for by the negative impact of maternal long-term illness that has a significant negative impact on the child's internalising conduct at age 7. The internalising issues remain significant, shown by positive margins outcomes of internalising behaviour when households move to higher deciles of health IMD, this is due to the impact of the maternal longstanding illness that has a persisting positive impact on the internalising issues of the child throughout ages 7 to 14. There is a marked decrease in internalising issues as households move to area having lower deciles of crime IMD and exhibits a negative impact on the internalising behaviour where the child adopts new internalising issues when interacting with peers and with their external environment.

Similar to the rise in internalising issues with households movements to lower overall and individual IMD deciles, children also exhibit an increase in the prosocial behavioural issues. Factors such as poor maternal health or low household income exhibit a declining prosocial score, indicating an increase in the prosocial issues when moving to lower deciles of overall and individual IMDs. Even the households moving into higher deciles of overall and individual IMDs show a little improvement over their prosocial scores that highlights the long-term impact of the externalising issues at child age 7, and from age 11 to 14 moving higher up the deciles of overall IMD leads to improvement in prosocial conduct.

The current study corroborates with Gambaro & Joshi (2016) in terms of the movement from more deprived to less deprived or vice versa. The current study utilises the unitary movements of households to and from more deprived to less deprived IMD deciles. Unlike classifying the top 70% areas as better, the study uses the continuity in the household movements. The analysis takes into account households' relative movements to the current IMD deciles, and the findings highlighted that there is a negative impact on vocabulary scores as one moves down the lower IMD. The analysis consists of the households witnessing improvement and decline of vocabulary and behavioural scores when moving across the overall and individual IMD deciles. Hence, the analysis captures these changes and highlights the mobility impact on vocabulary and behavioural scores for households moving to different IMD deciles. The overall IMD decile movements' impacts differ from the individual IMD index for given vocabulary and behavioural scores, highlighting the marginal impacts of the households moving to different IMDs. Also, the individual IMDs provides the insights into the interaction of the specific factor acting as an input to a child's vocabulary and behavioural development at age 7. The studies of (Gambaro & Joshi, 2016; Gambaro et al., 2017) have characterised the households based on situations allowing them to move to different areas of residence. The mobility situations included the housing space, neighbourhood quality, partnership changes, parental employment transitions and family demographics, health, and economic resources. The current study excludes such household situations and attempts to find out the impact on the improving or declining vocabulary and behavioural development scores. This highlights the impact of the IMD movements alone that can shape a child's vocabulary and behavioural transitions while moving to different IMDs, focusing on the fact that the impacts of changing IMDs do not always remain same for development outcomes of the child. Also, unlike (Beck et al., 2016; Gambaro & Joshi, 2016; Gambaro et al., 2017), the present study has used the questions directly asked to the respondents about their reasons to change resident. This directly captures the household decisions which impact the child development.

The MCS data is a panel of mobile and immobile families from age 3 to age 14. Gambaro & Joshi (2016) & Gambaro et al. (2017) have also used the panel data in their respective studies. The present study utilises the two time period cross-sectional analyses. The aim of which being to observing the dynamics of child development between two periods consisting of families who also entered new or remained at same residence aged 5. This captures the change in child development of families moving between age 3 and successive age groups (7,11 & 14) against the families who definitely move/stayed between age 3 and 5.

The model specification in this study is based on Bono et al. (2016) & Todd & Wolpin (2007) focusing on the importance of lag period child development. This also minimises the issues of unobserved child and household specific endogeneity. The impact is observed for the change in child vocabulary and behavioural development that includes child development measures for both time periods. This cancels out the effect of any unobserved factors impacting the current period child development. Child development is influenced by the factors which led the households to decide to move, and the difference in child development between two given time periods is captured in the age-specific factors of previous and current sweep.

The majority of families have not moved throughout sweeps 2 to 6, i.e., ages 3 to 14. From the sample of families who have actually moved, the majority have moved only once between ages 3 to 14. This provides the ground for the age-specific analysis that could predict the child's development.

Apart from the purposes of movement by the households, further analysis on the impact of changing government policies on housing, provision of education, health, environment and income distribution in different areas could also impact decisions on household mobility. This requires large-scale government policy to address these issues which could influence household mobility. The early years policy could further its potential by focusing on the household demographics of both the moving and staying families. This study could also develop the analysis for the staying families who want to move but are not able to because of a variety of reasons, which could include the inability to afford new housing, problems in adapting to a new area, taking care of parents, good schools in the current area of residence, and so on. Gambaro & Joshi (2016) highlights the consequences on families who choose to stay as a negative impact on their development. Further analysis on the circumstances and the reasons for why staying families could not move can provide a wider picture.

3.7 Conclusion

Parental mobility decisions are influenced by family situations. These decisions are made as a result of situations faced by parents and have a significant impact on the child vocabulary and behavioural development. The change could be beneficial or detrimental to a child's vocabulary and behavioural development. This is of particular interest to the field of education economics where the household environment plays a crucial role in shaping the child's early years of development. A child's early vocabulary and behavioural measures form a basis for future achievements in the child's life. The current study utilises the changing household situations to track an improvement or decline of the child development measures. There is an important role of the deprivation index of the area recognised by the current study as well as by the Gambaro & Joshi (2016) & Gambaro et al. (2017) studies in shaping child development outcomes.

The IMD deciles given in the MCS dataset provide an important measure of the household's mobility decisions which could be used to measure the resultant impact on the child's development indicators. The unified measure of area categorisation, apart from the household socioeconomic measure, gives a collective overview which helps measure the impact of household decisions on shaping child development. Additionally, there could be some measurement errors in the analysis arising from the fact that the household mobility decisions to a particular IMD decile are influenced by a set of correlated factors, although there is a gap in the current study to measure the impacts of the factors individually. The MCS dataset lacks information on the schools and related information on the parental mobility reasons influenced by the proximity of schools to their residence that could influence parental decisions to change residence while moving to areas having different deciles of IMD.

By using the household mobility decisions to move to higher and lower deciles of the IMDs of residence at ages 7,11 and 14. The current study raises the interaction of household mobility decisions with the government policies that could help ameliorate the increasing negative impact on vocabulary and behavioural development as children move from higher deprivation areas. This could help government agencies identify these area, and design area-focused policies, and reach out to the households.

Chapter 4

The Impact of Covid on the Youth's Preferences to Continue Education

4.1 Introduction

COVID-19 has widened educational inequalities across children and young people of different socioeconomic backgrounds (Anders et al., 2021; Blanden et al., 2022; Major et al., 2021; Rose et al., 2021; Werner & Woessmann, 2021), and young people (aged 14-17) are on the most vulnerable end of this. They have witnessed imbalances induced by school closures, a switch to online learning, inequality in access to online learning resources, and an absence of interaction with friends. The human capital theory by Heckman & Mosso (2014) gives importance to the dynamic complementarity which states that investments made in child's early years' should be accompanied by later years development. This chapter views COVID-19 as a shock to the typical educational path of youth, which is shaped by early-life inputs such as parents and household environment. Heckman & Mosso (2014) has highlighted the importance of dynamic complementarities of the skills, and in the present context, irrespective of timely investment from parents into children's skills in early age, there is a gap in the educational achievement of the child during COVID-19 that could lead to an incomplete accumulation of skills, since dynamic complementarity is not reached.

Recent findings by Anders et al. (2021) has highlighted the negative impact of socio-economic inequalities in terms of home learning experiences, returning to school, compared to more advantaged youth. Hence, the socio-economic inequalities negatively impact less advantaged young people's educational and subjective well-being. In another study, Macmillan & Tominey (2023) examined the impact of compulsory schooling reform in England. Their study found that family resources and parental investments are impacted by the educational reforms that raised female school leaving age in England. This change led to an increase in human capital, and had an additional benefit in increased second-generation cognitive skills by 0.12-0.14 standard deviations higher than those not covered by the raising of the school leaving age. Therefore, the above findings, highlight the cumulative nature of skills and emphasises importance of previous year's intervention into youth educational preferences at the end of schooling. Motivated by the aforementioned studies, the current study addresses the parental impact through observing the dynamics of parental job and socioeconomic well-being on the youth educational preferences, as a result of COVID-19. To achieve this, I utilise the pre-COVID and COVID-19 Millennium Cohort Study datasets (MCS).[1]

The existing pre-COVID-19 studies capturing the change in youth educational preferences by Anders (2017) have utilised duration modelling. Using the duration modelling, Anders (2017) highlighted the change in preferences of 14-year-olds on whether they plan on going to college, and asked again when they turn 21. The duration modelling technique helped to shed light on the significant role of socioeconomic factors at ages 14 and 21 on changes in expectations from being "likely to apply" to "unlikely to apply", the findings highlighted that more disadvantaged youth change

[1]The MCS survey was conducted pre-COVID-19 and during COVID-19. In the present research, I am using both pre and during COVID-19 MCS datasets. In the pre-COVID-19 MCS dataset, there are GCSE and other equivalent exams grades. I have used the exam grades as lagged inputs into child's educational preferences during COVID-19.

their decisions from "likely to apply" at age 14 to "unlikely to apply" at age 21. The policy implications could be to support these disadvantaged young people and promote access to education, the government could provide assistance to academically able youth from disadvantaged backgrounds.[2]

In the context of present study, given the interview questions asked of the youth in the COVID-19 MCS dataset, I use the probit model to capture the youth (aged 18-23) career preferences of whether to "continue education" or "not continue education". The choices provide an overview of youth career prospects for those who experienced the pandemic. The findings provide a holistic view of the impact of parental dynamics by using the ONS SOC sectoral information as a result of COVID-19, on the youth (aged 18-23) educational preferences. The motivation behind using the ONS sectoral information lies in the fact that it captures the changes in the working hours due to the pandemic. Also, previous studies by Rege et al. (2011) and Huttunen & Riukula (2019) has utilised parental job transition to study the impact on educational outcomes of the children and adolescence, respectively. Hence, there lies an importance of parental job transitions on the educational outcomes of the children. Therefore, the current study utilises the pre and COVID-19 MCS datasets to find out the impact of parental pre-COVID SOC transitions on the youth educational preferences during COVID-19. The MCS dataset consists of the pandemic induced changes in work hours that led to different impact on various parental professional sectors. This parental job insecurity raises the question of the youth preferences to continue further in education. Therefore, to find out the impact of change in parental employment status on the youth career preferences, the study uses a set of previous-age parental, child and household inputs in the pooled cross-sectional data.

The research findings show an insignificant impact of parental work transitions on youth educational preferences after addressing the possible factors affecting the youth

[2] Anders (2017) study is not focused on the effects of the pandemic on youth preferences but is helpful to understand the young people's career prospects since it has used the duration modelling, giving importance to the young people's preferences of how likely or unlikely they will go to the university.

career preferences. The sectoral analysis of the occupations showed a positive significant maternal impact in the case of single-parent households on the probability to continue education, suggesting that COVID impacted the ONS occupations overall and encouraged youth to continue in education. There is an increase in the socioeconomic adversities of the youth in shaping their career. Hence, the findings of the study provide valuable insights into the measures to be taken by policymakers in shaping the youth career paths. One of the important insights is to help policymakers focus on the household inputs that lead to changes in the youth educational aspirations. It could be achieved by designing a occupational sectoral policy (classifying young people as "most exposed" and "less exposed") to help ameliorate the negative impacts of the pandemic-induced job insecurity.

The remainder of the chapter is organised as follows: section 4.2 discusses some related studies; section 4.3 describes the data and variables used for analysis; section 4.4 analyses the impact of change in parental working status and socioeconomic(SES) status on the youth's preferences to continue education during COVID by using the econometric framework and presenting the results; section 4.5 shows the impact of pre-COVID parental ONS Standard Occupational Classification(SOC) on the youth's preferences to continue education; section 4.6 discusses the main findings and compares them with previous studies; and section 4.7 concludes.

4.2 Literature Review

Due to COVID-19 in the UK, the 2021 university cohort were deprived of face-to-face course delivery (Major & Machin, 2020; Murphy & Wyness, 2020). These students were disadvantaged, vulnerable to unequal availability of learning resources, and faced set backs in grades and performance, negatively impacting their career potential. A -levels that year were awarded by teacher-graded assessments, many of whom predicted low grades, and Murphy & Wyness (2020) commented on the reduction of disadvantaged children in top universities following cancelled exams in 2020. Ray-Chaudhuri & Xu (2023) & Major & Machin (2020) reported the scarring effect that the reduction in total number of hours worked per year by a fifth had on young people, and how it led to reduced work experience and career progression. There has also been a reduction in apprenticeships for young adults (aged 16-24) that led to a decline in the proportion of youth with fewer qualifications in well-paid jobs. During COVID, many young people missed out on work experience, training, and learning, hence lead to the "scarring" effect mentionned above, which is estimated to have been economically worse than the 2008 global financial crisis. There is no exact pattern of young peoples' education as a result of the pandemic in terms of whether they are able to stay in education, choose to leave education, or choose to delay entry to the labour market in the expectation of a better future. Overall, unlike during the labour market shocks in 2020 and early 2021, youth entering the labour market post-COVID do not tend to start in lower-quality occupations; also, career progression regarding occupation-level pay does not appear to be any slower as there is no difference in the occupation-level pay in the first year after graduating for 2018, 2019 and 2020 cohorts is in line with other pre-pandemic cohorts since 2013. Using Quarterly Labour Force Survey data, the 2019 and 2020 cohort were less likely to change jobs after finishing school or university for 0-1 and 1-2 years after graduating. The on-the-job training received in early 2020 was slightly less common than that received by previous cohorts, but the on-the-job training after one to two years after graduation was no longer lower than the previous cohorts. Using the Office for Budget Responsibility data (2022), concerns arise that the youth entering the labour market in the next few years will be exposed to a fall in disposable incomes by 4.3% for 2022-23 (Ray-Chaudhuri & Xu, 2023), which may mean youth decide to spend an extra year in education to gain higher qualifications and enter the labour market at a higher income.

The pandemic in the UK had differing impacts on different occupations which had various implications for youth entering the labour market, additionally compounded by parental social-ties of job selection where the social ties capture the parental experience at work, and whether they experienced any period of unemployment. Kramarz & Skans (2014) finds a significant impact of the parental social-ties in the determination of youth first job. The current study, using median income distribution from Standard Occupation Classification (SOC), as per Government office regions, finds regional disparities in youth career choices depending on the sectoral exposure to the pandemic. In the context of the present study, the maternal and paternal working transitions during the pandemic help to account for the still growing insecurity and lack of clarity among youth regarding higher education or entering the labour market as parental job transitions act as an indicator of parental job insecurity. Consequently, youth leaving school or in university show an inclination to remain in education if they witness less parental job transition compared to those who witness more parental job transition.

The focus of the current study on youth preferences to continue education shows that parents' changes in working hours indicate job insecurity. This job insecurity raises the question of the youth preference to continue further in education. The youth in the present study are aged 18-21 years old, which is the age young people typically enter the labour force. Previous research by Ruiz-Valenzuela (2015) highlights that the combination of previous educational attainment and cumulative parental experiences during financial recessions and opinions on the importance of education can influence youth preference for educational attainment. Using the impact of parental unemployment during the recession in Spain as an example, students experienced a significant negative decrease in their grades by 13% of the standard deviation. In terms of parental job loss impacting young people, this study found a negative and significant reduction in average grades of around 15% of a standard deviation due to a father's job loss, however a mother's job loss on school performance is close to zero and insignificant. A Mooi-Reci et al. (2019) study on a sample of Dutch children exposed to their parents' unemployment during the 1980s economic crisis in the Netherlands has stressed the importance of mothers maintaining positive views about work to help support the environment necessary for educational attainment, and finds both a direct negative impact of father's unemployment with children's educational achievement and also a positive relationship between parental pessimistic views regarding job and children's educational attainment. Therefore, it is essential to highlight that the combination of previous academic attainment and cumulative parental views can influence the youth's preferences for educational attainment.

The individual parental transition in the labour market forms an essential base for studying youth career prospects during the pandemic. It is necessary to bring to light the differences in the impact of the pandemic on working hours and pay of mothers and fathers. Hupkau et al. (2020), using the Understanding Society data, has explored the effects of COVID and the restrictions on economic activity for men and women. They found that women were 4 percentage points more likely to be furloughed than men, and that the gender differential is positive in the case of occupation and industry as controls, consistent with lower furlough in female-dominated jobs. This positive gender differential, after including the job-related controls, highlighted differences in working hours, which led to the increase in those with young children working from home. This highlights the adverse shocks of COVID-19 borne by the parents and having a knock-on effect on their children in terms of resource and time investment. There is a further decrease in the earnings and employment for already disadvantaged parents. An indirect impact of gender differentials on children's development through job transitions which could pose a cumulative impact on children's overall development and decision to continue education.

Youth preference to continue education is influenced by past grades and abilities, and parental socioeconomic and educational level. Anders et al. (2021) has identified the impact of the pandemic on children's well-being aged between 11-18 attending state-funded education in England and finds distinct socio-economic inequalities where young people from disadvantaged families facing learning inequalities and gender inequalities, such as females showed significant lower well-being scores than males. Previous child development studies (Anders, 2017; Blanden et al., 2022; Ruiz-Valenzuela, 2015) have highlighted the positive significant impact of parental education and socioeconomic factors in a child's academic performance. The positive impact of parental education is illustrated by the increase in the compulsory post-16 schooling, and an improvement in children's outcomes aged 4 to 16 (Macmillan & Tominey, 2023). Therefore, parents with higher-level education have an impact on their children to continue their own education, and likewise negative outcomes where lower grades are awarded to children who are deprived of learning alternatives.

The unequal availability of learning alternatives during the pandemic can be accounted for by socioeconomic inequalities (Anders et al., 2021). Job insecurity led to a decrease in household income, which is an essential input in the determination of the educational achievement of children. Hence, for a particular household, the SES of fathers has a positive significant impact on the academic performance of the school-age children. A nationally representative survey of youth based on educational record found a significant negative impact of social inequalities in home learning, return to school and exam cancellations in more disadvantaged families, compared to better grades for advantaged youth. By utilising the parental change in financial situations during COVID, this study analyses the impact of SES, however rather than focusing on the impact of learning inequalities due to increased socioeconomic inequality, the current research investigates the impact of maternal and paternal job transition and changed financial situation during COVID on youth career preferences.

Anders (2017) emphasised the role of SES by focusing on the change in the expectations of youth aged between 14 and 17. More advantaged youth showed a small change in the preference to join university by age 18, where around 35% of disadvantaged youth and 50% from advantaged households made a shift from 'unlikely to likely' to apply to university. Hence, social inequalities are more prominent in the disadvantaged households, impacting choices to continue further in education.

In line with Anders (2017) and Anders et al. (2021), this research focuses on youth career paths. Anders (2017) discusses youth career preferences pre-pandemic, and Anders et al. (2021) highlights the impact of socioeconomic inequalities on educational achievement paths of youth during the pandemic. In the latter, there is an increased learning inequality in the form of returning to school, and home schooling which had a negative impact on disadvantaged households compared to advantaged ones. The current research aims to contribute to literature on growing inequality in youth educational achievement during COVID by highlighting the impact of parental job insecurity and the change in the parental SES on youth career prospects. It finds that there is an impact of parental job insecurity and SES in determining youth academic progress. The impact of all these components forms a basis for determining children's progress, developed abilities, and grades achieved at school.

In summary, the chapter analyses the impact of parental job status during the pandemic on the youth educational preferences. To observe this the analysis uses the probit model, which measures youth preferences for education and career i.e, whether the youth wants to continue in education or not and is coded as binary. It is valid to code as binary during the pandemic, as given the small sample size of the youth it is valid to use the probit model. It is important to analyse the impact of SES on youth educational preferences during COVID as the pandemic has widened the learning inequalities among the children and youth. Due to the unavailability of learning resources in terms of some disadvantaged households not able to provide with private tuitions for the child, online learning material, to compensate for the learning losses during the pandemic. However, the findings on SES are not significant and there appears to be no significant link between job change, SES, and career preference of the youth.

4.3 Data

The MCS COVID survey was conducted in April 2020 (first wave), September 2020 (second wave), and February 2021 (third wave). I have considered all the responses from these waves.

The analysis uses the MCS COVID-19 waves longitudinal data, consisting of questions from cohorts aged 18-23 and their respective parents/like figures. The sample size consists of 2,832 observations; wave 1 sample size consisted of 1,452 observations; wave 2 2,257 observations; and wave 3 consisted of 2,272 observations. For youth responding to more than one COVID wave, the earliest response was taken for the analysis. The questions concerned schooling experiences, financial management, and mental and physical health. The first COVID-19 wave consisted of questions about pre- and during-pandemic experiences. Other questions concerned how the cohort members managed financially and their overall health. To keep the questionnaire length short, some questions are absent for new participants entering COVID waves 2 and 3.

The pre-COVID MCS data set consists of seven sweeps to date where the children's ages range from 9 months (sweep 1) to 17 years (sweep 7). This study utilises the third (age 5), fourth (age 7), fifth (age 11), sixth (age 14) and seventh (age 17) sweeps, covering development from child to teenager. The pre-COVID sweeps capture early-years feelings on likelihood of continuing education. Prior waves' child and parental inputs act as controls to address the impact of parental job transition on youth career preferences. Responses to some questions asked of cohorts or parents/parent-like figures were absent in one of the waves, so previous wave responses to similar questions are considered to compensate for the non-response in a particular wave. For the pre-COVID waves, surveys are mainly conducted with parents/parent-like figures when children are young, and later on by both parents and the young people in question. The pre-COVID wave data is linked with the help of the Unique identifier present for each cohort member.

4.3.1 Sample Representation

The MCS pre- and during COVID waves are disproportionately stratified clustered. The sample was selected systematically within each stratum and country. The ward selection (done in MCS pre-COVID wave 1, remaining the same for the MCS COVID wave) is based on the expected number of births in that particular ward. If the number of births in a ward is less than expected for the time period of the survey, it is combined with the adjacent wards to form a larger sampling unit and described as super wards. The sample is clustered in the form of a ward to include all eligible cohorts in the selected ward; this is done to address the socioeconomic disparities and to reduce the survey cost. There were separate sample selections for each country as per their target sample sizes: the target sample sizes for Scotland, Wales and Northern Ireland are larger than their population to address the intra and inter-country comparisons. Finally, the wards were selected as per the ratio of the English regions in England and Scotland and only by the ward size (based on the expected number of births) in Wales and Northern Ireland. The combined weight comprising sample design and survey non-response is used for the MCS COVID-19 sample analyses.

To adequately represent the ethnic minority and disadvantaged groups, the population of electoral wards was classified as a disadvantaged, advantaged and ethnic minority (only in England). The stratification was done based on the four countries of the UK (England, Wales, Scotland and Northern Ireland). The MCS technical report by (Plewis et al., 2007) characterises a disadvantaged ward where the population lies in the lower quartile of England and Wales's 1998 Child Poverty Index (CPI). These families received some income support, for instance jobseeker's allowance, family credit, or disability working allowance. The advantaged stratum consists of the ward's population lying in the lower quartile of the 1998 CPI. Ethnic minority stratum was only present in England, where 30% of the total population was "black" or "asian". The current study uses the MCS pre and during COVID technical reports for information on stratum formation and clustering of the sample. The sample weights present for the COVID survey are used (Plewis et al., 2007).

Attrition

There were three waves of the COVID MCS survey. The first COVID wave survey was initiated in April 2020, the second in September 2020, and the third in February 2021. Cohort members did not necessarily respond in all waves. In order to minimise the impact of non-response during any COVID wave, the current research combines all three COVID wave cohorts to capture information on whether young people plan to continue their course. The total cohort in all three COVID waves and responses in one of the waves is 1,088.[3] The sample consists of only those respondents who responded in at least one COVID wave. There were 1,715 (61% youth) respondents in wave 1, 861 (30% youth) respondents in wave 2 and 256 (9% youth) in wave 3. In the pooled sample, there were in total 855 youth (30%) who wanted to continue/start a course and 1,977 (70%) who did not want to continue a course. In total, the sample distribution of youth preferences to continue or start a course or stop a course at wave 1 consisted of 303 youth (18%) wanting to start or continue a course and 1,412 (82%) wanting to stop a course, at wave 2 there were 534 youth (62%) wanting to start or continue a course and 327 (38%) wanting to stop a course, and at wave 3, there were only 18 youth (7%) wanting to start or continue a course and 238 (93%) wanting to stop a course. The wave-specific marginal impact (controlling the COVID waves) of parental job transition and change in financial situation on the youth likelihood to continue in education is shown in table 4.2 and table 4.3 in the results section.

In the COVID MCS interview, there were respondents (for pre and COVID waves) who did not answer all questions and were coded as "do not know". "Refusal", where respondents declined to answer, led to non-response attrition, and so, under-representation of the new movers (in case of pre-COVID waves), to account for this, weights are assigned to each of the four countries of the UK. The attrition

[3] In case of non-response in a particular wave, the response given in the next wave is recorded.

weights also allow for an over-representation of winter births in Scotland and Northern Ireland. Thirdly, there were sample losses while conducting MCS interviews from Child Benefit records to the families present and issued to the field because some families moved outside the sample ward. Some families targeted for interview moved from the address registered with the National Pupil Database (NPD).

The dependent variable is coded as 0 "stop course" and 1 "continue course". The cohort members were asked whether they accepted any place at college or university that year (coded as "yes" or "no"). Those who opted for "no" or selected "yes", but stopped the course are categorised as 0 "stopped course", and those who selected "yes" opted for the course or want to continue the course are coded as 1 "started course". The question of whether they want to return to the course was asked in the second wave (cohort members were interviewed in September 2020). The detailed description of the outcome variable is shown in the following sub-section 4.3.2

The current research focuses on parental pre- and post-COVID working status. The responses to the questions on economic activity in both periods are considered. Those

TABLE 4.1: Descriptive statistics of change in work and financial situation according to the child's stated intention to continue the course.

Panel A:
(Parental change in work status)

	Single & Two parent						Two parent only					
	Maternal			Paternal			Maternal			Paternal		
Number[%] where	Cont. edu. (1)	Not cont. edu. (2)	N (3)	Cont. edu. (4)	Not cont. edu. (5)	N (6)	Cont. edu. (7)	Not cont. edu. (8)	N (9)	Cont. edu. (10)	Not cont. edu. (11)	N (12)
Remain Working	434	907	1,341	200	411	611	191	385	576	212	417	629
	[32%]	[68%]	[100]	[33%]	[67%]	[100]	[33%]	[67%]	[100]	[34%]	[66%]	[100]
Switch working	5	14	19	4	8	12	3	7	10	1	2	3
	[26%]	[74%]	[100]	[33%]	[67%]	[100]	[30%]	[70%]	[100]	[33%]	[67%]	[100]
Switch not working	60	142	202	34	65	99	34	57	91	28	44	72
	[30%]	[70%]	[100]	[34%]	[66%]	[100]	[37%]	[63%]	[100]	[39%]	[61%]	[100]
Remain not working	85	179	264	37	75	112	35	72	107	22	58	80
	[32%]	[68%]	[100]	[33%]	[67%]	[100]	[33%]	[67%]	[100]	[28%]	[73%]	[100]

Panel B:
(Financial satisfaction)

	Single & Two parent						Two parent only					
	Maternal			Paternal			Maternal			Paternal		
Number[%] where	Cont. edu. (1)	Not cont. edu. (2)	N (3)	Cont. edu. (4)	Not cont. edu. (5)	N (6)	Cont. edu. (7)	Not cont. edu. (8)	N (9)	Cont. edu. (10)	Not cont. edu. (11)	N (12)
Much worse off	55	98	153	34	54	88	21	38	59	31	44	75
	[36%]	[64%]	[100]	[39%]	[61%]	[100]	[36%]	[64%]	[100]	[41%]	[59%]	[100]
A little worse off	120	248	368	70	124	194	61	101	162	59	105	164
	[33%]	[67%]	[100]	[36%]	[64%]	[100]	[38%]	[62%]	[100]	[36%]	[64%]	[100]
About the same	288	649	202	140	307	447	131	274	405	114	236	350
	[31%]	[69%]	[100]	[31%]	[69%]	[100]	[32%]	[68%]	[100]	[33%]	[67%]	[100]
A little better off	101	202	303	64	133	197	43	89	132	50	115	165
	[33%]	[67%]	[100]	[33%]	[68%]	[100]	[33%]	[67%]	[100]	[30%]	[70%]	[100]
Much better off	20	45	65	11	22	33	7	19	26	9	21	30
	[31%]	[69%]	[100]	[33%]	[67%]	[100]	[27%]	[73%]	[100]	[30%]	[70%]	[100]

The author uses the sample of MCS COVID data to collect information on maternal and paternal shifts in working and financial status. Percentages are shown in the square brackets for respective categories of change in parental working status and change in financial situation during COVID.

employed, self-employed and working, in apprenticeship, or voluntary work without equal working hours are classified as working (1). The respondents in education, unemployed, looking after family, or retired are classified as non-working (0).

4.3.2 Youth educational preferences-outcome variable

In the analytical sample, the age composition of the youth ranges from 18-23 years. Out of these youth, approximately 61% of the youth are 20 years old, followed by 25% aged 21 years, and 14% aged 19 years. There is a negligible composition of youth aged 18 (0.14%) and 23 (0.04%) years. I have considered follow on questions asked of youth who had formally accepted an offer for a place in the college/university. The follow on questions in the COVID MCS survey, consisted of whether the youth will start college or university course or apprenticeship programme, which is going to start later this year/ in the autumn term, or have already returned to their course. The responses were coded as "yes", "deferred entry to the course", and "no". The majority of respondents aged 20 years indicate that the sample consists of cohort that are at the end of undergraduate and entering postgraduate level ONS-Edu. (2023). Further, to address heterogeneity in the age composition, and to ensure consistency, the analysis controls for whether the respondents (youth) are currently working or not, which is coded as yes(=1) and no(=0). The panel(a) of figure 4.1 shows the age-composition in the sample, and the panel(b) of the figure 4.1 shows the age-composition of those who want to start or stop the course. Hence, the majority of cohort aged 20 years in the sample, ensures that the outcomes of the analysis are consistence.[4] The present study represents the sample of individuals who are at the end of undergraduate and start of postgraduate level. Hence, the age-composition in the current sample provides a suitable platform to draw insights from the findings by Ray-Chaudhuri & Xu (2023) on whether the youth want to stay in education, or delay employment in the hope of a better labour market in the future.

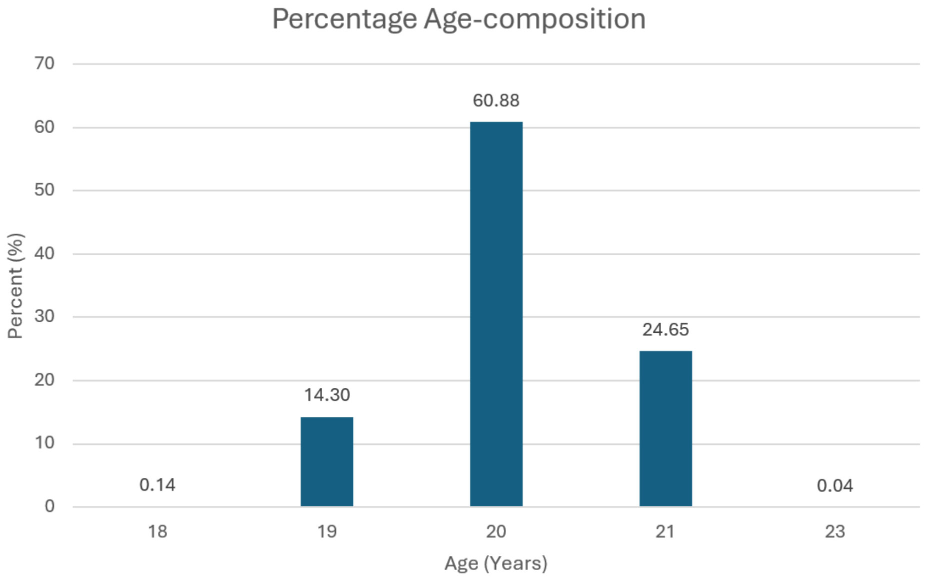
Variables of interest

The variables of interest for the present research are the change in parental employment status (working being 1 and not working 0); the change is observed between pre-COVID and during COVID employment status. Ruiz-Valenzuela (2015), investigate the impact of the Spanish depression on the children's academic performance and used in their analysis a set of definitions regarding father's job loss:

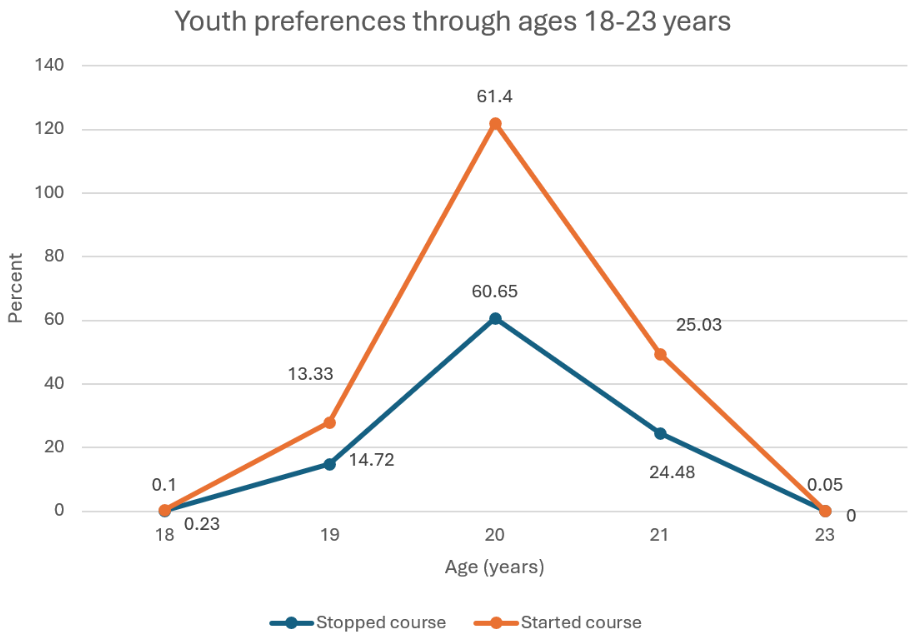
[4]The current sample does not contain youth aged 16 or 17. Therefore, unlike Major & Machin (2020) findings, the analysis does not account for the transitions faced by the school leavers aged 16 in terms of failure to qualify for GCSEs, or aged 17 to enrol in the college education.

FIGURE 4.1: The age composition of the youth and their educational preferences.

(a) Age composition



(b) Youth educational preferences



Notes.-The author's calculations are based on the age composition of the youth as per their educational preferences.

the first being the fathers' involuntary job loss during the crisis; second was fathers' voluntary and involuntary job loss or closing their own business; the third were job losses from downsizing or closure. In contrast, the current study categorises the parent respondents as looking for a job, taking care of the family, retired, in full-time education, or have not worked at all as "not working" (coded 0). At the same time, the parental working status is coded as "working"(1) for those employed, self-employed, key workers, or on an apprenticeship. Also, slightly different from the approach of Mooi-Reci et al. (2019) & Ruiz-Valenzuela (2015), the current study aims to collect employment information from both parents in the survey, regardless of whether they stayed in the same house. The importance of the questions asked highlighted the heterogeneity in parental characteristics during COVID. This helps analyse the individual parental impact of job transitions on the youth's decision to continue in education. It has been well-established in education literature that a positive impact of both mothers and fathers are present in shaping the skills, abilities, and views of the youth. Similar to the current survey, Ruiz-Valenzuela (2015) collected information about employment situation before and after the great recession. Any changes in employment situation after the recession was captured.

To assess the change in the parental financial situation on the likelihood of youth to continue in education, I use the variable change in the parental financial situation from pre-COVID to during COVID which is expected to positively impact the youth's career prospects. Anders (2017), Anders et al. (2021) & Ruiz-Valenzuela (2015) have focused on the SES of the parents. Anders et al. (2021) categorises students belonging to parents having high school education as advantaged as compared to those students belonging to households having a free school meal status. Anders (2017) uses the LSYPE data to capture household income, parental education, and parental occupational status and also the social class in measuring socioeconomic status. To capture the change in the financial situation, the current analysis uses the self-reported variable on the change in the financial situation; the self-reported variables asked of the parents are coded as “much worse off”, “little worse off”, “about the same”, “little better off,” and “much better off”. I have considered the base category as those with the same financial situation. Predictive margins capture the impact of change in the financial situation on the probability of continuing education.

Table 4.1 shows the descriptive statistics of youth who want to continue or not continue the course for the case of single and two-parent households, and two-parent only households analysed for both parents. In the case of "single and two-parent only" and "two-parent only" analyses, maternal job transition from "switching to working from not working" or "switching to not working from working" found a higher composition of youth who do not want to continue in education. Similarly, in the case of "single and two-parent only" and "two-parent only" analysis, paternal job transition from "switching to working from not working" has a higher composition of the youth who do not want to continue in education. But, there is no difference in the youth composition in case of no change in the parental work status except for the "two-parent-only" paternal analysis where fathers choosing to "remain not working" have a higher composition of the youth not continuing in education.

In panel B for "two and single parent-only" analysis, the cohorts belonging to mothers who are much better off or remain the same have the highest composition of those who don't want to continue the course; the youth belonging to fathers whose financial conditions are the same comprise the highest proportion. In the case of "two-parent-only" analysis, cohorts belonging to mothers whose financial conditions are the same comprise the highest proportion and cohorts belonging to fathers whose financial conditions are little or much better off comprise the highest proportion. But, this does not necessarily conclude that the youth belonging to advantaged households encourage the probability of not continuing in education. This may be because financial improvement could be seen more in disadvantaged households than in those which are already advantaged.

ONS SOC occupations

During the pandemic, ONS occupations were exposed to COVID-19 in various ways. Occupation exposure depends on the flexibility of the work, i.e., whether the nature of work allowed working from home. Some industries saw complete shutdown of firms or downsized, while some saw reduced working hours as a result of inefficiency caused by working at home. One of the major reasons for this inefficiency was the time invested in the child's activities, such as completing homework, home-schooling, and helping the child with online learning. Households having at least one child below 6 were the most affected and reported reduced working hours. The impact of working hours differed across the GOR in the UK, with some reporting early school closures and closures of firms. The closures of firms were mostly observed in the GOR, comprising a large proportion of industries where work from home was not possible and required the presence of the employees. However, the GORs comprising of occupations where the work nature allowed working from home witnessed reduced working hours. This varying regional occupational impact significantly influenced the parental views regarding the work and, in turn, the youth. In order to capture this varying regional occupational phenomenon, I have used the median income distribution of the ONS (2023) SOC occupations for each GOR for the UK Census 2021 provided in ASHE (2023) database. The median income used in the SOC rankings helps to capture the impact of the less contact intensive to contact more intensive occupations, which further leads to differences in the flexibility of choosing a workplace, work hours, and furlough which, in turn, acts as a potential explanation of the impact of parents on youth career preferences.

The ONS SOC occupations is a list of all occupations that ranged in contact intensity during the pandemic. The SOC occupations demand the presence of workers, who are unable to run without face-to-face contact with the clients, these are professions requiring laboratory work, health professionals, or those in the hospitality sector. Therefore, the professions where there was a slight possibility of remote working, and those who witnessed a reduction in working hours due to the decreased feasibility of remote working were categorised as contact-intensive SOC occupations.Panel B of table 4.6 provides the list of ONS SOC occupations, which are arranged as per their descending order of contact intensity (ascending median income). The importance of area of residence (GOR in terms of the present study) on the parental occupations is not new to the COVID-19 impact shown in the present study. As it can be seen from the previous study by Rege et al. (2011), that has utilised the Norwegian educational and register database. The educational database consists of graduating secondary graders in Norway from 2003-05. The register database is a longitudinal dataset consisting of records of every Norwegian from 1992 to 2003. The register database consists of parental and child demographic, socio-economic, employment status, industry of employment information. The findings by Rege et al. (2011) has focused on the decline in the school performance of children whose father faced a job loss. Rege et al. (2011) study found a significant decrease in the children's GPA having fathers layed-off in non-booming sectors. Whereas, there is no impact on the grades of the booming municipalities. The results suggest that the grade points is 0.12 points lower for children of fathers employed in closing plants and residing in non-booming municipalities. They have captured the effect of plant closures in the form of paternal income loss, but the estimates did not show a significant impact on the children's GPA. Whereas, the current study has utilised the parental pre-COVID-19 SOC and the change in the financial situation, separately, to analyse the impact on the youth educational preferences. Also, in line with their findings on the analysis outcome, the present study did not find any significant impact on the youth preferences. Their analysis also extends for maternal job loss impact on grade points. Their findings showed a different impact for mothers, where mothers displaced were less likely to search for job in non-booming municipalities. Rather their findings suggested an improvement in the maternal inputs to the child at home in terms of rearing child, increase in maternal time at home in improving child's grades.

This study focuses on the impact of parental job transitions from working to non-working or vice versa on the youth's educational preferences during COVID-19. Hence, this is the first study to find the impact of parental job transitions on youth educational preferences as a result of COVID-19. Previous study by Huttunen & Riukula (2019) has focused on the impact of parental job loss on the child's career in terms of field of study chosen by the child. These studies have provided evidence that father's job loss decreases the likelihood of child choosing the father's field of study. The study focuses on the impact of parental job loss as a result of Great Depression in Finland in the early 1990s. Using the Finnish administrative data 1988-2016. The findings show that displaced father whose earnings reduced due to great depression has no effect on the length of schooling of children. Also, there exists no impact on the study choices for children who had already made their educational descisions since their father's job loss compared to those who are at the end of schooling. Hence, their study have stressed the importance of timing of parental job loss. Huttunen & Riukula (2019) findings indicate that parental job loss at the time where children are making schooling descisions have strong impact on their career choices. Their study found that father's job displacement has a significant impact on child's schooling descisions where the fathers loose their jobs when the child is older than 16 years. The findings show that there is no impact of maternal job displacement on their children. Hence, the present study consisting of youth having average age of 20 years indicate that the parental job loss has an important impact on the youth educational preferences.

Another study by Hilger (2016) utilised administrative tax records on over 7 million fathers layoff between the years 2000-09. Their findings suggest that layoffs during adolescence have very small causal impact on children's long-term outcomes. The parental layoffs reduce the annual college enrolment over the ages 18-22 years by 1 percent. But these layoffs increases the labour supply during adolescence and college. The 10-15 percent of reduction in late childhood income has a causal reduction in the college enrolment. Therefore, motivated by the previous studies on adolescence career preferences, the present study uses separate maternal, paternal analysis to contribute to the the literature of sociology highlighting the importance of paternal role as a primary earner and mothers focusing on rearing the child as a result of job loss.

To find the maternal and paternal SOC impact, the present study uses the MCS pre-COVID-19 and COVID-19 dataset, to find the impact on the youth educational choices in terms of pursuing the course in a college or not. Another study by Fradkin et al. (2019) focus on the impact of parental job loss on the labour market outcomes of young adults. But, the current study focuses on the educational outcomes of the youth in terms of their preferences: to continue in education or not. The study by Fradkin et al. (2019) finds the impact of parental job loss on their children's labour market outcomes. The youth tend to find the first job quicker. Whereas, the present study focuses on the educational preferences of the youth to continue in education or not, as a result of parental job transitions during COVID-19. There have been previous studies focused on the child's career preferences due to parental unemployment but none have focused on the youth educational preferences due to parental job transitions during the shocks such as COVID in the present context. To capture the maternal and paternal impact, the current study uses the MCS data consisting of parental pre-COVID-19 SOC occupations.

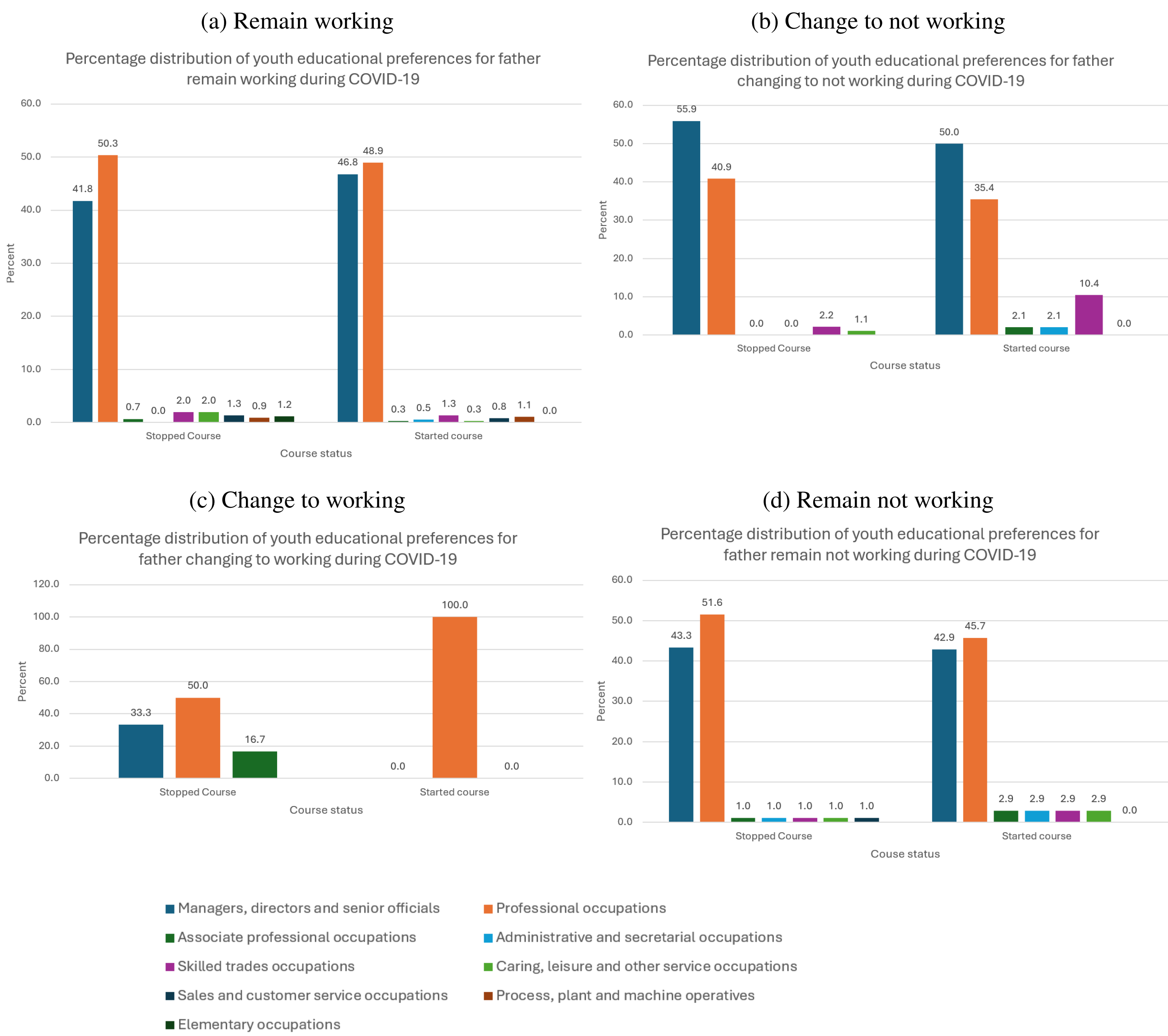
To provide overview of the impact of pre-COVID parental SOC on the youth preferences, figures 4.2 and 4.3 show the impact of maternal and paternal transitions in the SOC occupations. As per figure 4.2, youth belonging to mothers facing transition in managerial and professional occupations show higher likelihood to continue or not in the education compared to other pre-COVID-19 SOC occupations. Out of the maternal transition status, those who change to not working and who remain not working, the mother employed in managerial occupations inflicts higher likelihood to pursue or end the course. Whereas, mothers who remain working or transitioning to working and employed in pre-COVID-19 professional occupations, inflict higher likelihood to pursue education. In sum, majority of educational preferences are explained by the youth belonging to mothers employed in managerial or professional pre-COVID-19 occupations. Similarly, this trend continues for fathers employed as professionals in pre-COVID occupations impacting youth's likelihood of educational preferences. The situation where the father change to not working as shown in panel(b) of figure 4.3, the youth belonging to fathers employed in managerial roles during pre-COVID-19, inflicts highest percentage ranging from 50% of youth who want to start course and 56% of youth who want to stop the course.

Predictive margins are used to capture the change in parental work status as shown in table 4.2 and table 4.4, change in the financial situation as shown in table 4.3 and table 4.5 and the varying GOR regional impact of the ONS SOC rankings as shown in table 4.8, table 4.9, table 4.10 and table 4.11. The change in the work status is categorised as "remaining not working" (0), "switching to work from not working" (1), "switching to not working from working" (2), and "remain working" (3). The base category is "not working during COVID". Another mechanism affecting the likelihood is the change in the financial situation, which is categorised as: "about the same" (0), "much worse" 1), "a little worse" (2), "a little better (3)", and "much better (4)".

FIGURE 4.2: The percentage distribution of youth educational preferences based on change in maternal working condition for pre-COVID maternal SOC.

Notes.-The author's calculations are based on the youth preferences to continue education or not in case of maternal pre COVID SOC impact during COVID-19.

FIGURE 4.3: The percentage distribution of youth educational preferences based on change in paternal working condition for pre-COVID paternal SOC.



Notes.-The author's calculations are based on the youth preferences to continue education or not in case of paternal pre- COVID SOC impact during COVID-19.

Controls

The paper utilises child and parental-specific variables as controls. The pre-COVID cohort-specific controls asked to those aged 17 include whether the cohort wants to attend university in the future.. As mentioned earlier, Anders (2017) uses duration modelling to highlight a change in preferences of 14 year-olds going to college when they turn 21. The duration modelling technique helped to highlight the significant role played by socioeconomic factors from age 14 onwards in the expectations from being "likely to apply" to "unlikely to apply" of young people at age 21. Anders (2017) also used 11 and 16 Z-scores as controls, as well as whether the cohort attended a grammar school and a school with a sixth form. The child's age in the current sample ranges from 18-23, with an average age of 20. The MCS data helps to capture the differences in the education stage when analysing youth career prospects. To capture the cohort differences, whether to choose work or education or address the youth current financial situation and the extent to which the pandemic has shaped their financial situation, I have used the information from the data on whether the cohort was also working during or before the pandemic. To address the previous age educational achievements, I have used the A/A* grades in GCSEs and the equivalent qualifications (1 if higher education, 0, otherwise). The cohort-specific controls that affected the channel of receiving education were the changes in the learning activities experienced by the youth (1 "learning affected", 0 "learning not affected"). The other sociodemographic controls are the ethnicity coded as white(1) and non-white(0) and gender(1 if male and 0 female). Past studies on child achievement have addressed a significant impact of ethnicity that widens with the child's age. The GOR of the cohort's residence helps address the regional inequalities. During COVID, regional inequalities were seen in schools and companies shutting down or downsizing. This varying impact of the pandemic between regions is worth addressing in the case of youth educational preference. Having siblings under the age of 6 helps in analysing the distribution of work hours and the parental time investment that impacted work hours, for example switching to part-time with a reduction in earnings which again varied on the contact intensity of the occupations.

Parental educational level plays a significant role in shaping child and youth career prospects. Ruiz-Valenzuela (2015) has given importance to the parental educational status and has combined it with the household's SES to find its impact on child development. In the current study, I categorise parental education into two categories: where the highest academic qualification achieved is "A/AS diploma/GCSE" and is coded as (0 "not higher"; 1 "higher degree"). Parental age is considered an important factor in education literature affecting children's academic achievement, and hence is essential to find out the impact on the youth's career prospects due to the variations due to parental age. Also, the "still not working" spell to date (maternal and paternal) of the interview is asked of both mothers and fathers, helping to capture the period of unemployment, which affects youth views on the preferences of working and financial anxiety; also, it helps to address the gaps between those parents who did not work to focus on their children's education at an early age (Lim & Sng, 2006; Mooi-Reci et al., 2019). During COVID, many households received furlough which impacted their SES; it led to differences in people's attitude regarding work and the prospects of working, such as reduced motivation at work, a reduction in working hours. In the MCS, the question on "whether receiving any form of benefits during pandemic" was asked of the parents or parental figures. The current analysis has considered parents who have received furlough (1 if responded as received furlough and 0 if not received furlough). For analysis, I have considered individual parental controls.

The household-specific controls used are IMD (1 deprived, 0 not deprived); IMD 1, 2 and 3 coded as "deprived", and 4-10 coded as "not deprived"; OECD income quintiles (quintiles 1, 2 and 3 are coded as lower) (quintiles 4 and 5 are coded as higher). The wealth index is formed by using variables such as receiving a pension (1 if "yes"), owning a house (1 if "yes"), owning an automobile (1 if "yes"), having investments (1 if "yes") and not receiving any state benefits (1 if "yes"). The principal components analysis forms the wealth index, and the resulting household wealth variable is coded as wealthy (1) and not wealthy (0). Not wealthy is considered as the base category. Therefore, by utilising cohort-specific, parental and household controls, the current study aims to address the likelihood of the preferences of the youth to continue in education.

4.4 The impact of change in parental working and SES on the youth's preferences to continue education during COVID.

The section highlights the econometric framework and presents results on the youth's educational impact during COVID.

4.4.1 Method

The current analysis uses the regression model where the outcome variable is binary, ranging from 0 to 1. Anders et al. (2021) used this approach to determine the risk of transitioning from "likely" to "unlikely" to apply to university. This model is widely used to evaluate the probabilities of the decisions made based on the choices available to the individual.

In the current study context, the cohort's decision on whether to continue a course ranges from 0 ("stop the course") to 1 ("start the course"). The transitions in these factors played an important role in the youth career prospects. For the initial analysis of the impact of inputs on the cohort's decision the present study uses the probit model. Since the analysis is based on the likelihood, the conditional probability of the cohort member choosing to continue the course or not could be shown below:

$$P(Course_i = 1|X_i) = \phi(X^{'}\beta) \tag{4.1}$$

where X_i is a set of all exogenous variables such as parental change in working hours, parental inputs and cohort-specific variables. ϕ is the cumulative distribution function of the standard normal distribution. The following equation (4.2) shows the binary distribution of the variable "Course".,

$$Course = \begin{cases} 1, Y^* > 0, \\ 0, \text{otherwise}, \end{cases} \tag{4.2}$$

The probability of whether the cohort wants to continue the course reported in equation (4.1) is not unbiased and consistent since the equation (4.1) has estimation issues as the $E(X_i|\mu_i + \varepsilon_i) \neq 0$. where,

$E(X_i) = E((\sum_{i=1}^{c} C_i, \sum_{i=1}^{p} P_i)), where, \sum_{i=1}^{c} C_i$ is the set of cohort inputs and $\sum_{i=1}^{p} P_i$ is the set of parental inputs.

Hence, the error term is not independent of the explanatory variables and violates the estimates' zero conditional mean assumption. So, the estimation result from equation (4.1) would be biased and inconsistent as the covariance between error and independent terms will be indifferent from zero.

The present analysis finds the probability of the youth to continue education by analysing the mother only and father only separately for single and two parent households. The equations (4.3) to (4.5) capture the maternal only impact on the probability to continue education. Similarly, equations (4.6) to (4.8) capture paternal only impact on the probability of the youth to continue education. And the final sets of equations (4.9) to (4.11) capture the combined maternal and paternal impact incase of two parent only households. The motivation of individual analysis of single and two parent household is to separate out the impact of two-parent households from the single parents. Hence, the equations (4.9) to (4.11) separate out the impact of two parents on the youth preferences to continue education from the single parent households as shown in equations (4.3) to (4.8). The coefficients captured from two-parent equations, certainly include only two parent households impact on the youth preferences to continue education, and the outcomes could be compared from the coefficients obtained through single parent equations. As pointed out by previous studies such as (Hilger, 2016; Huttunen & Riukula, 2019) giving importance to the maternal and paternal roles, especially in the events encompassing loss to the economy, such as the occurrence of Great Depression, and parental job loss in the past. These are characterised by a negative impact on the children educational achievements as evident from the above studies' findings. In the present context, analysing the impact of maternal and paternal job transition for single and two parent households, during the COVID-19 pandemic highlights the importance of individual maternal and paternal roles in influencing the educational outcomes of the youth.

The present study uses the probit model to find out the educational outcomes of the youth. The choice of whether to continue a course depends upon the following probit approach analysed as maternal-only, paternal-only for single and two-parent households and two-parent only households. By addressing all the possible cohorts, parental and household-specific controls, the probit model approach is shown below in equation (4.3) to (4.11), capture the likelihood of a youth's choice to continue education. The choice of whether to continue a course can be analysed as maternal-only, paternal-only for single and two-parent households and two-parent only households.

$$Y_i^* = \beta_0 + \beta_1 Precov_i^M + \beta_2 Cov_i^M + \beta_\tau Precov_i^M \times {Cov_i}^M + \mathbf{Z}\lambda + e_i \tag{4.3}$$

$$Y_i^* = \beta_0 + \beta_1 Fin_i^M + \mathbf{Z}\lambda + e_i \tag{4.4}$$

$$Y_i^* = \beta_0 + \beta_1 SOCGOR_i^M + \mathbf{Z}\lambda + e_i \tag{4.5}$$

In equations (4.3) to (4.5), ${Y^*}_i$ takes on only two values ranging from 0 to 1. "1" indicates young people choosing to continue a course during COVID, and "0" means that young people want to end a course. The variables of interest are maternal pre-COVID and COVID working hours,denoted by ${Precov^M}_i$ and ${Cov^M}_i$, consisting of categories such as working (1) and not working (0), financial status change, denoted by ${Fin^M}_i$, and median income ranking of SOC in a government office region(GOR), denoted by ${SOCGOR^M}_i$. $\mathbf{Z}\lambda$ stands for the set of all maternal, household and cohort inputs respectively for each observation. The maternal inputs consist of mothers having higher education, unemployment history, which is coded as 1 if unemployed at least once and 0 if not unemployed. The wealth index indicates whether the cohort belongs to a wealthy household. The wealth index comprises 0, 1 values where 1 stands for wealthy and 0 otherwise, whether mothers received furlough, whether working zero hours, the maternal age, and age squared.

Similar to equations (4.3) to (4.5), the variables are defined in (4.6) to (4.8), but for paternal values rather than maternal, as shown below:

$$Y_i^* = \beta_0 + \beta_1 Precov^F{}_i + \beta_2 Cov^F{}_i + \beta_\tau Precov^F{}_i x Cov^F{}_i + \mathbf{Z}\lambda + e_i. \tag{4.6}$$

$$Y_i^* = \beta_0 + \beta_1 Fin_i^F + \mathbf{Z}\lambda + e_i \tag{4.7}$$

$$Y_i^* = \beta_0 + \beta_1 SOCGOR_i^F + \mathbf{Z}\lambda + e_i \tag{4.8}$$

In two-parent-only households the combined maternal and paternal analysis is carried out in equations (4.9) to (4.11) as shown below:

$$Y^*{}_i = \beta_0 + \beta_1 Precov^M{}_i + \beta_2 Cov^M{}_i + \beta_{\tau,M} Precov^M{}_i x Cov^M{}_i + \beta_3 Precov^F{}_i + \beta_4 Cov^F{}_i + \beta_{\tau,F} Precov^F{}_i x Cov^F{}_i + \mathbf{Z}\lambda + e_i \tag{4.9}$$

$$Y_i^* = \beta_0 + \beta_1 Fin_i^M + \beta_2 Fin_i^F + \mathbf{Z}\lambda + e_i \tag{4.10}$$

$$Y_i^* = \beta_0 + \beta_1 SOCGOR_i^M + \beta_1 SOCGOR_i^F + \mathbf{Z}\lambda + e_i \tag{4.11}$$

In equations (4.3) to (4.5), $Y^*{}_i$ takes on only two values ranging from 0 to 1. "1" indicates young people choosing to continue a course during COVID, and "0" means that young people want to end a course. The variables of interest are maternal pre-COVID and COVID working hours,denoted by $Precov^M{}_i$ and $Cov^M{}_i$, consisting of categories such as working (1) and not working (0), financial status change, denoted by $Fin^M{}_i$, and median income ranking of SOC in a government office region(GOR), denoted by $SOCGOR^M{}_i$. $\mathbf{Z}\lambda$ stands for the set of all maternal, household and cohort inputs respectively for each observation. The maternal inputs consist of mothers having higher education, unemployment history, which is coded as 1 if unemployed at least once and 0 if not unemployed. The wealth index indicates whether the cohort belongs to a wealthy household. The wealth index comprises 0, 1 values where 1 stands for wealthy and 0 otherwise, whether mothers received furlough, whether working zero hours, the maternal age, and age squared.

Here, $Y^*{}_i$ takes on only two values ranging from 0 to 1. 1 to indicate young people choosing to continue from a course during COVID, and 0 meaning young people want to end a course. As per equation (4.9) the variables of interest are maternal and paternal pre-COVID-19 and COVID-19 working hours denoted by $Precov^M{}_i$ and $Cov^M{}_i$ and $Precov^F{}_i$ and $Cov^F{}_i$, respectively. The categories consist of working (1) and not working (0). Also, $\sum_{i=1}^{p,M} P_i^M$ and $\sum_{i=1}^{p,F} P_i^F$ stands for the set of all maternal and paternal inputs. Similarly, $Fin^M{}_i$ and $Fin^F{}_i$, show the change in maternal and paternal financial status in two-parent households, and in equation (4.11), $SOCGOR^M{}_i$ and $SOCGOR^F{}_i$ show the maternal and paternal median income SOC ranking for respective GOR. In all the equations (4.9) to (4.11), $\mathbf{Z}\lambda$ stands for the set of all household and cohort inputs respectively for each observation. The parental inputs consist of mothers and fathers having higher education, unemployment history, coded as 1 if unemployed at least once and 0 if not unemployed, whether mother, father or both received furlough, whether working zero hours, and the parental age, and age squared. Similar to previous equations (4.3) to (4.8), the wealth index indicates whether the cohort belongs to a wealthy household in a two parent setting, which is shown in equations (4.9) to (4.11). The wealth index comprises 0, 1 values where 1 stands for wealthy and 0 otherwise.

4.4.2 Results

This subsection provides a detailed presentation of the study's findings. The primary insight suggests that parental job insecurities, influenced by pandemic-induced shifts in work hours and sectors, play a pivotal role in shaping youth's educational preferences.

Single and two-parent households impact on the likelihood of continuing education.

The single and two-parent households consist of both mother and father respondents belonging to single and two-parent households. It is not necessary for parents live in the same household as the cohort. If both parents responded to the survey, it is considered as two parents and in case only one parent responded it is considered a single-parent household. Tables 4.2 & 4.3 show the marginal impact of change in parental work status and the SES impact on the youth's educational preferences respectively during COVID by controlling for COVID waves utilised in the current chapter.

Table 4.4 delineates the influence of parental employment changes from the pre-COVID period to during COVID pandemic on youth educational choices. Mothers continuing employment during COVID is shown to positively influences youth decision to continue education. This effect is attenuated slightly when controlling for confounding variables. The marginal probabilities in panel B, column (2) suggest that there is a 2.9 (28.9-26.0=2.9) ppt decrease for mothers transitioning from work to non-work and a 14.5 (28.9-14.4=14.5) ppt decrease for mothers transitioning into employment during COVID.

The sharp decline in educational propensity for youths whose mothers began working during the pandemic might be attributed to a need to supplement family income due to paternal job losses. This shows the impact of increased demand in the female-dominated sectors such as health and social care during COVID. Also, during COVID the demand for health care assistants increased. The minimal qualification for the health care assistants is GCSEs, therefore, the mothers who joined the health care sector may have given youth a reason to not continue courses and encouraged them to pursue jobs which do not require degrees. This trend resonates with the findings of Ruiz-Valenzuela (2015) on the limited impact of maternal job loss during Spain's great depression.

TABLE 4.2: Individual maternal and paternal marginal impact of change in work status and financial situation on youth educational preferences controlling for COVID waves.

	Maternal				Paternal			
Controls(Waves)	All waves (base=1st wave)	Wave1	Wave2	Wave3	All waves (base=1st wave)	Wave1	Wave2	Wave3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Predictive margins (Parental change in work status)								
PreCOVIDxCOVID work status (0=not working, 1=working)								
0x0	0.218	0.217	0.218	0.274	0.223	0.219	0.222	0.198
0x1	0.138	0.152	0.139	0.139	0.437	0.438	0.433	0.332
1x0	0.230	0.239	0.230	0.256	0.378	0.363	0.378	0.344
1x1	0.281	0.279	0.281	0.288	0.277	0.278	0.277	0.304
Observations	1,826	1,826	1,826	1,826	959	959	959	959
Panel B: Predictive margins (Change in financial situation)								
About same	0.259	.254	.259	.274	.254	.257	.253	.284
Much worse	0.276	.278	.276	.323	.304	.301	.304	.329
Little worse	0.260	.266	.259	.275	.288	.285	.287	.273
Little better	0.295	.297	.295	.296	.327	.322	.327	.321
Much better	0.220	.219	.220	.260	.218	.210	.220	.314
Observations	1,826	1,826	1,826	1,826	959	959	959	959

Notes:- The author uses the sample of MCS COVID data to collect information on maternal and paternal shifts in working and financial status. Predictive margins are shown for respective categories of change in parental working status and change in financial situation during COVID.

The data shows a contrasting trend for fathers as shown in panel B, column (4) of table 4.4 youth whose fathers transitioned out of employment during COVID have a 3.7 (34.2-30.5=3.7) ppt higher likelihood to continue their studies compared to those fathers who remain working. This possibly points to the long-term value families place on education when faced with economic uncertainties. Children of fathers who started working during the pandemic exhibited a 3.1 (33.6-30.5=3.1) ppt increase in continuing education with respect to fathers who remain working, contrasting with a significant decline (14.5(28.9-14.4=14.5) ppt) when mothers entered the workforce. Previous studies have found a positive significant impact of father's employment status on children's academic achievement. Mothers have a significant impact in the early years but less so in later life. The current findings corroborate in this regard that the positive paternal impact of working has long-term impact on children's academic achievements. There is also a role of family importance given to education that positively inclines youth preferences towards education.

TABLE 4.3: Combined maternal and paternal impact of change in work status and financial situation on youth educational preferences controlling for COVID waves.

Controls(Waves)	Maternal and Paternal			
	All waves (base=1st wave) (1)	Wave1 (2)	Wave2 (3)	Wave3 (4)
Panel A: Predictive margins(Maternal)				
PreCOVIDxCOVID work status(0=not working, 1=working)				
0x0	.295	.290	.296	.294
0x1	.238	.229	.237	.138
1x0	.321	.330	.318	.315
1x1	.320	.315	.320	.316
Observations	784	784	784	784
Panel B: Predictive margins(Paternal)				
PreCOVIDxCOVID work status(0=not working, 1=working)				
0x0	0.230	.234	.226	.177
0x1	0.635	.627	.628	.338
1x0	0.337	.322	.338	.322
1x1	0.267	.268	.267	.282
Observations	784	784	784	784
Panel C: Predictive margins				
Change in financial situation(Maternal)				
About same	.252	.248	.252	.278
Much worse	.300	.317	.298	.315
Little worse	.313	.311	.313	.315
Little better	.315	.315	.314	.301
Much better	.144	.160	.143	.260
Observations	784	784	784	784
Panel D: Predictive margins				
Change in financial situation(Paternal)				
About same	.244	.253	.243	.285
Much worse	.363	.359	.362	.375
Little worse	.289	.282	.288	.271
Little better	.297	.285	.298	.290
Much better	.223	.216	.225	.299
Observations	784	784	784	784

Notes:- The author uses the sample of MCS COVID data to collect information on maternal and paternal shifts in working and financial status. Predictive margins are shown for respective categories of change in parental working status and change in financial situation during COVID.

Table 4.5 extends the analysis by examining the financial situation of households. As shown in panel B column (2) of table 4.5 a worsening financial condition for mothers relates to a 32% increased likelihood of youths continuing education, with a similar but slightly lower trend (approximately 28%) for those in a slightly worsened financial situation. Mothers experiencing an improved financial status, likely due to joining the workforce, show a lesser inclination (approximately 26%) for their offspring to continue education. At the same time, for fathers, a deteriorating financial situation reflects approximately 33% rise in the likelihood of their children continuing

TABLE 4.4: Individual maternal and paternal impact of change in work status on youth educational preferences.

	Maternal		Paternal	
Panel A:	Baseline (1)	With controls (2)	Baseline (3)	With controls (4)
Pre-COVID (Base=Not working)				
Working	-0.152	-0.038	0.193	0.437
	(0.141)	(0.179)	(0.216)	(0.298)
COVID(Base=Not working)				
Working	-0.451	-0.456	0.527	0.421
	(0.343)	(0.348)	(0.587)	(0.576)
Pre-COVID x COVID				
(Base=not working x not working)				
Working x Working	0.550	0.543	-0.607	-0.525
	(0.362)	(0.377)	(0.610)	(0.609)
Constant	-0.443***	20.207	-0.585***	21.357
	(0.092)	(37.700)	(0.150)	(52.383)
Panel: B Predictive margins				
PreCOVID x COVID work status(0=not working, 1=working)				
0x0	0.329	0.273	0.279	0.199
0x1	0.186	0.144	0.477	0.336
1x0	0.276	0.260	0.348	0.342
1x1	0.310	0.289	0.318	0.305
Observations	1,826	1,826	959	959

Notes.- The analysis consists of probit predictions for the probability of continuing the course. The robust standard errors are in parenthesis, clustered at the individual level. The model controls for the parental highest education level, wealth index, IMD of the cohort, cohort's ethnicity, whether the household received furlough, whether parents ever experienced unemployment, any siblings aged 6 or under, and whether parents are working zero hours contracts. The outcome is a dummy ranging from 0 to 1, where 1 represents the cohort's decision to continue the course, and 0 otherwise. The robust standard errors are reported in parentheses.
*** Significant at the 1 per cent level ** Significant at the 5 per cent level. * Significant at the 10 per cent level.

education. However, a slightly worsened financial condition relates to a 27% decrease as shown in panel B, column (4) of table 4.5.

The results underscore the variable influence of economic stressors on youth education decisions. While financial adversities triggered by COVID seem to emphasise the significance of education for many, parental employment changes, particularly for mothers, might divert youths from academia.

In contrast to the extant literature (Anders (2017)) suggests that socioeconomic factors and historical family views on education do play a crucial role in influencing these outcomes. The paper's findings on "two and single-parent households" also challenge

TABLE 4.5: Individual maternal and paternal change in financial impact on youth educational preferences.

	Maternal		Paternal	
	Baseline (1)	With controls (2)	Baseline (3)	With controls (4)
Financial satisfaction (Base=about the same)				
much worse off	0.098	0.147	0.249	0.126
	(0.129)	(0.133)	(0.171)	(0.174)
little worse off	0.041	0.015	0.073	-0.034
	(0.091)	(0.097)	(0.126)	(0.133)
little better off	0.095	0.075	0.150	0.101
	(0.097)	(0.101)	(0.125)	(0.134)
much better off	-0.062	-0.043	0.096	0.080
	(0.192)	(0.216)	(0.262)	(0.264)
Pre-COVID(Base=Not working)				
Working		0.064		0.296
		(0.125)		(0.230)
Constant	-0.534***	19.201	-0.548***	23.835
	(0.049)	(37.712)	(0.071)	(52.486)
Panel: B Predictive margins				
Change in financial situation				
About same	0.297	0.273	0.292	0.285
Much worse	0.331	0.324	0.382	0.329
Little worse	0.311	0.278	0.317	0.274
Little better	0.330	0.299	0.345	0.320
Much better	0.275	0.259	0.326	0.313
Observations	1,826	1,826	959	959

Notes.- The analysis consists of probit predictions for the probability of continuing the course. The robust standard errors are in parenthesis, clustered at the individual level. The model controls for the parental highest education level, wealth index, IMD of the cohort, cohort ethnicity, whether the household received furlough, whether parents ever experienced unemployment, any siblings aged 6 or under, and whether parents are working zero hours contracts. The outcome is a dummy ranging from 0 to 1, where 1 represents the cohort's decision to continue the course, and 0 otherwise. The robust standard errors are reported in parentheses.
*** Significant at the 1 per cent level ** Significant at the 5 per cent level. * Significant at the 10 per cent level.

Lim & Sng's (2006) argument, suggesting that heightened financial anxiety due to COVID might intensify youths' value for education rather than deterring them.

The two-parent-only households impact on the likelihood of continuing education.

The two-parent-only households consist of mothers and fathers who responded to the survey. It is not necessary for parents to live together or in the same household as the child or young person.

In the current analysis, the likelihood of continuing education is placed against the backdrop of parental employment changes during the COVID-19 pandemic in the case of two-parent-only households, table 4.6 shows that the youths' educational continuity is positively influenced when mothers remain in the labour force during the pandemic compared to those who were never employed. However, a significant negative deviation of 18(31.6-13.8=17.8) ppt is observed when mothers initiate employment during this period, from a baseline of 31.6% for those consistently employed to 13.8% as shown in panel B, column (2) of table 4.6 for those entering the labour force. Conversely, the working status of fathers presents an inverse relationship. Youth with fathers who persisted in their employment during the pandemic exhibited a lower likelihood of continuing education. Remarkably, when fathers commenced work during this period, the likelihood of the youth's continuation in education surged by 12.4(34.2-18=12.4) ppt.

TABLE 4.6: Combined maternal and paternal impact on youth educational preferences.

	Maternal and Paternal	
	Baseline (1)	With controls (2)
Maternal: Pre-COVID (Base=Not working)		
Working	-0.117	0.069
	(0.209)	(0.261)
COVID(Base=Not working)		
Working	-0.501	-0.544
	(0.442)	(0.452)
Pre-COVID x COVID		
(Base=not working x not working)		
Working x Working	0.486	0.540
	(0.473)	(0.495)
Paternal: Pre-COVID (Base=Not working)		
Working	0.226	0.449
	(0.237)	(0.320)
COVID(Base=Not working)		
Working	0.387	0.509
	(0.798)	(0.783)
Pre-COVID x COVID		
(Base=not working x not working)		
Working x Working	-0.486	-0.615
	(0.817)	(0.806)
Constant	-0.495**	19.975
	(0.196)	(68.384)
Predictive margins:		
*PreCOVID x COVID work status(**Maternal**)*(0=not working, 1=working)		
0x0	0.371	0.293
0x1	0.204	0.138
1x0	0.328	0.318
1x1	0.323	0.316
*PreCOVID x COVID work status(**Paternal**)*(0=not working, 1=working)		
0x0	0.255	0.180
0x1	0.393	0.342
1x0	0.332	0.320
1x1	0.298	0.283
Observations	784	784

Notes.- The analysis consists of probit predictions for the probability of continuing a course. The robust standard errors are in parenthesis, clustered at the individual level. The model controls for the parental highest education level, wealth index, IMD of the cohort, cohort ethnicity, whether the household received furlough, whether parents ever experienced unemployment, any siblings aged 6 or under, and whether parents working a zero hours contract. The outcome is a dummy ranging from 0 to 1, where 1 represents the cohort's decision to continue a course, and 0 otherwise. The robust standard errors are reported in parentheses.
*** Significant at the 1 per cent level ** Significant at the 5 per cent level. * Significant at the 10 per cent level.

TABLE 4.7: Combined impact of maternal and paternal change in the financial situation on youth educational preferences.

	Maternal and Paternal	
	Baseline	With controls
	(1)	(2)
Maternal: Change in financial situation		
(Base=about the same)		
much worse off	0.083	0.121
	(0.232)	(0.229)
little worse off	0.042	0.107
	(0.139)	(0.152)
little better off	0.064	0.071
	(0.149)	(0.158)
much better off	-0.147	-0.045
	(0.315)	(0.344)
Pre-COVID(Base=Not working)		
Working		0.102
		(0.192)
Paternal: Change in financial situation		
(Base=about the same)		
much worse off	0.300	0.245
	(0.204)	(0.211)
little worse off	0.043	-0.045
	(0.142)	(0.152)
little better off	0.060	0.005
	(0.144)	(0.151)
much better off	-0.027	0.034
	(0.282)	(0.280)
Pre-COVID(Base=Not working)		
Working		0.275
		(0.258)
Constant	-0.557***	24.37
	(0.089)	(69.038)
Predictive margins:		
*Change in financial situation(**Maternal**)*		
About same	0.305	0.277
Much worse	0.335	0.319
Little worse	0.320	0.315
Little better	0.328	0.302
Much better	0.256	0.263
*Change in financial situation(**Paternal**)*		
About same	0.296	0.286
Much worse	0.407	0.375
Little worse	0.311	0.271
Little better	0.317	0.288
Much better	0.287	0.298
Observations	784	784

Notes.- The analysis consists of probit predictions for the probability of continuing the course. The robust standard errors are in parenthesis, clustered at the individual level. The model controls for the parental highest education level, wealth index, IMD of the cohort, cohort ethnicity, whether the household received furlough, whether parents ever experienced unemployment, any siblings aged 6 or under, and whether parents working a zero hours contract. The outcome is a dummy ranging from 0 to 1, where 1 represents the cohort's decision to continue the course, and 0 otherwise. The robust standard errors are reported in parentheses.
*** Significant at the 1 per cent level ** Significant at the 5 per cent level. * Significant at the 10 per cent level.

Further, panel B column (2) of table 4.6 demonstrates that mothers exiting the workforce during the pandemic leads to a marginal boost in educational continuation by approximately 3(31.8-29.3=2.5) ppt. In stark contrast, the youth's likelihood of persisting in education jumped by 14(32-18=14) ppt. when fathers ceased employment during the same period.

Table 4.7 delves deeper into the intricate relationship between familial financial circumstances and the propensity for educational continuation. Importantly, as shown in column (2) of panel B of table 4.5, both mothers' and fathers' deteriorating financial conditions during the pandemic were associated with an increased likelihood (32% and approximately 33% respectively) of the youth continuing their education. This seems to underline the intrinsic value placed on education as a stabilising factor during financially challenging times. These observations align with the findings of Ruiz-Valenzuela (2015), reinforcing the notion that familial perceptions towards the attainment of specific educational levels are often decoupled from the immediate socioeconomic realities. In essence, the findings from table 4.7 elucidate that youths who witnessed financial upheavals, either for the better or worse, placed an elevated emphasis on education, leading them to persist in their academic pursuits. This heightened receptivity to remain in education in the face of paternal financial shifts is evident across two-parent-only households. Furthermore, the financial contributions of mothers, indicative of their heightened economic participation, also lead to a positive correlation with youths' inclination to continue with education.

Hence, in case of two-parent-only households, the intertwined dynamics of parental employment and financial status during the COVID pandemic have displayed a profound influence on the educational trajectories of the youth. The subtle interplay of these factors underscores the criticality of socio-economic considerations in shaping academic outcomes during unprecedented times.

4.4.3 Testing the analysis-Margins of significance.

TABLE 4.8: Individual maternal and paternal impact of change in work status on youth educational preferences.

Panel A:	Single & Two parent	
	Chi2 (1)	Hypothesis (2)
Maternal:		
PreCOVIDxCOVID-PreCOVIDxCOVID		
WorkxNot work-WorkxWork		
1x0-1x1	2.08	moving out job sig. differ from those who remain working.
Not WorkxWork-WorkxWork		
0x1-1x1	2.08	entering the job sig. differ from those who remain working.
Paternal:		
PreCOVIDxCOVID-PreCOVIDxCOVID		
WorkxNot work-WorkxWork		
1x0-1x1	0.74	moving out job sig. differ from those who remain working.
Not WorkxWork-WorkxWork		
0x1-1x1	0.74	entering the job sig. differ from those who remain working.
Panel B:	Two parent	
	Chi2 (1)	Hypothesis (2)
Maternal:		
PreCOVIDXCOVID-PreCOVIDXCOVID		
WorkxNot work-WorkxWork		
1x0-1x1	1.19	moving out job sig. differ from those who remain working.
Not WorkxWork-WorkxWork		
0x1-1x1	1.19	entering the job sig. differ from those who remain working.
Paternal:		
PreCOVIDXCOVID-PreCOVIDXCOVID		
WorkxNot work-WorkxWork		
1x0-1x1	0.58	moving out job sig. differ from those who remain working.
Not WorkxWork-WorkxWork		
0x1-1x1	0.58	entering the job sig. differ from those who remain working.

Notes.- The analysis consists of testing for the significance of margins obtained using probit predictions.

Table 4.8 shows pivotal insights regarding the consequences of parental shifts in work status during COVID. An analysis of the shifts in mothers' working statuses—whether transitioning to non-working during COVID, commencing work during the pandemic, or maintaining their employment from pre-COVID times—reveals a pronounced effect. Specifically, the associated impact on their children's educational continuity ranges between 10% and 25%. These findings underscore the significant role maternal employment dynamics play in shaping educational outcomes for the youth. Fathers exhibit a strong influence on educational trajectories, but in a somewhat larger effect size compared to mothers. Transitions from employment to non-employment during the pandemic, and vice versa, as well as sustained employment throughout the period, correlate with a substantial 25% to 50% change in the likelihood of their children's continued education. Panel B of table 4.8 delves deeper into the dynamics of two-parent households. For the two-parent households, the marginal probabilities pertaining to the continuity of education are even more pronounced, ranging between 25% and 50%. Interestingly, for fathers within these households who transitioned into employment during the pandemic, the associated probability also lies within this same 25% and 50% range. This consistency suggests a heightened influence of paternal employment dynamics in two-parent-only households on educational decisions.

4.5 Impact of parental pre-COVID SOC occupations on the youth's likelihood to continue education.

The current section highlights the impact of pre-COVID SOC occupations on the educational preferences of the youth by using the econometric framework utilised in section 4.4.

4.5.1 Impact of SOC occupations on the likelihood to continue education.

Table 4.9 delves into the geographical distinctions, examining educational continuance across the UK GOR. Here, an upward trend is observed with rising SOC occupation rankings. Specifically, for mothers, there is a marked 1.2% spike in educational continuity likelihood when adjusting for exogenous factors. Conversely, a slight 0.3% upward trend is noted with increasing SOC median income.

TABLE 4.9: Marginal impact of occupations by Government office region

SOC occupation within GOR						
	Single & two parent				Two parent only	
	Matrnl. (1)	Patrnl. (2)	Matrnl. (3)	Patrnl. (4)	(5)	(6)
SOC occup.	0.016***	0.009	0.012**	0.003		
	(0.005)	(0.008)	(0.006)	(0.009)		
Mtrl. SOC occup.					0.012	0.011
					(0.008)	(0.009)
Ptrl. SOC occup.					0.004	-0.004
					(0.009)	(0.011)
GOR	0.006	-0.005	0.010	-0.003		
	(0.012)	(0.017)	(0.013)	(0.018)		
Mtrl. GOR					4.140***	4.251***
					(0.243)	(0.395)
Ptrl. GOR					-4.142***	-4.246***
					(0.241)	(0.393)
Controls	No	No	Yes	Yes	No	Yes
Constant	-0.751***	-0.611***	26.109	14.953	-0.707***	20.551
	(0.105)	(0.167)	(38.732)	(55.452)	(0.193)	(70.810)
Observations	1,668	891	1,668	891	721	721

Notes.- The analysis consists of GOR SOC rankings for mothers and fathers to predict the likelihood of continuing education. Standard errors are shown in parenthesis.
*** Significant at the 1 per cent level ** Significant at the 5 per cent level. * Significant at the 10 per cent level.

TABLE 4.10: Marginal impact of maternal occupations by Government office region

Maternal SOC rank by GOR	NE (1)	NW (2)	Y& H (3)	EM (4)	WM (5)	E (6)	L (7)	SE (8)	SW (9)	WL (10)	SC (11)
Corp. Mng. & Dirctr.											**.381 [24]**
Othr Mng. and Prptr.											**.344 [15]**
Sci., resch., eng. & tech. profnl.											**.377 [23]**
Hlth. Profnl.										**.364 [19]**	
Teach. & Edutnl. Profnl.										**.381 [23]**	
Busi., Media & Pub. Serv. Prof										**.368 [20]**	
Sci., Eng. & Tech. Assoc.										**.352 [16]**	
Hlth. & Soc. Care Assoc. Profnl.										**.328 [10]**	
Prot. Serv. Occ.										**.385 [24]**	
Cultr., Media & Sprt. Occ.										**.344 [14]**	
Busi. & Pub. Serv. Assoc. Profnl.										**.356 [17]**	
Admtr. Occ.									**.320 [8]**	**.320 [9]**	
Secrt. & Reltd. Occ.										**.305 [4]**	
Skld. Agri. & Reltd. Trad.											**.328 [11]**
Skld. Mtl, Elctrl. & Elect. trad.										**.360 [18]**	
Skld. Const. & Bldg Trad.	**.312 [13]**										
Txtls, Prntg & Othr. Skld. Trad.									**.313 [6]**		
Crng. Prsnl. Serv. Occ.									**.309 [5]**		
Leis., Trvl. & Rltd. Prsnl. Serv. occ.									**.301 [3]**		
Sales Occup.									**.297 [2]**		
Custmr. Serv. Occ.									**.316 [7]**		
Prcs, Plnt. & Mach. Oprtvs.									**.336 [12]**	**.336 [13]**	
Trnspt. & Mob. Mach. Drvrs. & oprtvs.										**.340 [14]**	
Elmtry. Trds. & Rltd. Occ.									**.324 [9]**		
Elmtry. Admn. & Serv. Occ.									**.294 [1]**		
N(Obs. for each GOR)	38	129	106	130	99	155	128	280	149	232	222

Notes.- The analysis consists of overall SOC rankings for mothers in single and two-parent households to predict the likelihood of continuing education. Regional rankings are shown in square brackets. The highlighted margins correspond to the highest likelihood of continuing education for the respective occupation. Detailed margins are provided in Table C.1
(NE-North East; NW-North West; Y & H-Yorkshire & Humber; EM-East Midlands; WM-West Midlands; E-East of England; L-London; SE-South East; WL-Wales; SC-Scotland)

TABLE 4.11: Marginal impact of paternal occupations by Government office region

Paternal SOC rank by GOR	NE (1)	NW (2)	Y& H (3)	EM (4)	WM (5)	E (6)	L (7)	SE (8)	SW (9)	WL (10)	SC (11)
Corp. Mng. & Dirctr.	**.356** **[22]**										
Othr Mng. & Prptr.	**.351** **[16]**										
Sci., resch., eng. & tech. profnl.	**.357** **[23]**										
Hlth. Profnl.	**.354** **[19]**										
Teach. & Edutnl. Profnl.										**.348** **[23]**	
Busi., Media & Pub. Serv. Prof	**.354** **[20]**										
Sci., Eng. & Tech. Assoc.	**.348** **[13]**										
Hlth. & Soc. Care Assoc. Profnl.	**.344** **[9]**										
Prot. Serv. Occ.										**.347** **[22]**	
Cultr., Media & Sprt. Occ.	**.349** **[14]**										
Busi. & Pub. Serv. Assoc. Profnl.	**.350** **[15]**										
Admtr. Occ.										**.335** **[9]**	
Secrt. & Reltd. Occ.										**.331** **[4]**	
Skld. Agri. & Reltd. Trad.	**.345** **[10]**										
Skld. Mtl, Elctrl. & Elect. trad.	**.353** **[18]**										
Skld. Const. & Bldg Trad.	**.352** **[17]**										
Txtls, Prntg & Othr. Skld. Trad.		**.350** **[5]**									
Crng. Prsnl. Serv. Occ.	**.340** **[5]**										
Leis., Trvl. & Rltd. Prsnl. Serv. occ.			**.313** **[2]**								
Sales Occup.		**.316** **[1]**									
Custmr. Serv. Occ.	**.341** **[6]**										
Prcs, Plnt. & Mach. Oprtvs.	**.347** **[12]**										
Trnspt. & Mob. Mach. Drvrs. & oprtvs.	**.346** **[11]**										
Elmtry. Trds. & Reltd. Occ.	**.380** **[8]**										
Elmtry. Admtr. & Serv. Occ.	**.337** **[1]**										
N(Obs. for each GOR)	25	62	45	75	64	83	73	146	74	123	121

Notes.- The analysis consists of overall SOC rankings for fathers in single and two-parent households to predict the likelihood of continuing education. Regional rankings are shown in square brackets. The highlighted margins correspond to the highest likelihood of continuing education for the respective occupation. Detailed margins are provided in Table C.2
(NE-North East; NW-North West; Y & H-Yorkshire & Humber; EM-East Midlands; WM-West Midlands; E-East of England; L-London; SE-South East; WL-Wales; SC-Scotland)

Table 4.10 and table 4.11 map the regional disparities in educational outcomes against occupational sectors. In the North East, the highest probability of educational persistence is seen for children of mothers employed as corporate managers and directors (33%). In stark contrast, fathers working in manual labour — roles such as farm labourers and construction workers — boost the likelihood of their children's educational continuity by a significant 38%. The propensity for these youth to prioritise education could be driven by a desire for greater job flexibility in the future, recognising the rigidity and potential instability associated with manual labour roles. In the North West, a distinct pattern emerges as outlined in table 4.11. Mothers in science, research, and technical professions boost the probability of their children's educational continuity by 34%. Mothers holding positions like corporate managers, scientific and research engineers, health professionals, teaching professionals, and those in protective service occupations tend to have the most substantial positive influence, with impacts ranging between 30% to 38%.

Table 4.12 and table 4.13 summarise the dynamics in two-parent households. Fathers in roles requiring limited formal education but demanding more hands-on labour or short formal training seem to foster a more substantial inclination in their children towards education, even after accounting for youth in paid apprenticeships. Conversely, mothers in roles demanding higher educational qualifications and permitting work flexibility, such as work-from-home arrangements, maintain their strong positive influence on their children's educational trajectories.

In summary, the above findings highlight a complex interplay between parental occupations, median income brackets, and regional disparities in shaping the educational aspirations and continuities among the youth. The distinction between the impacts of maternal and paternal occupations, particularly in the context of the flexibility and nature of their roles during the COVID era, underscores the multifaceted socio-economic considerations underpinning educational decisions.

TABLE 4.12: Marginal impact of maternal occupation in two-parent household by Government office region

Govt. Office Regions:	NE (1)	NW (2)	Y& H (3)	EM (4)	WM (5)	E (6)	L (7)	SE (8)	SW (9)	WL (10)	SC (11)
Matrnl. SOC rank by GOR											
Corp. Mng. & Dirctr.										**.379**[24]	
Othr Mng. & Prptr.	**.346**[12]										**.346**[15]
Sci., resch., eng. and tech. profnl.										**.376**[23]	
Hlth. Profnl.									**.365**[19]		
Teach. & Edutnl. Profnl.									**.380**[23]		
Busi., Media & Pub. Serv. Prof									**.369**[20]		
Sci., Eng. & Tech. Assoc.									**.354**[16]		
Hlth. & Soc. Care Assoc. Profnl.	**.332**[8]									**.332**[10]	
Prot. Serv. Occ.									**.383**[24]		
Cultr., Media & Sprt. Occ.					**.311**[11]						
Busi. & Pub. Serv. Assoc. Profnl.									**.358**[17]		
Admtr. Occ.									**.325**[8]	**.325**[9]	
Secrt. & Reltd. Occ.									**.311**[4]		
Skld. Agri. & Reltd. Trad										**.332**[11]	
Skld. Mtl, Elctrl. & Elect. trad									**.361**[18]		
Skld. Const. & Bldg Trad.	**.350**[13]										
Txtls, Prntg & Othr. Skld. Trad.									**.318**[6]		
Crng. Prsnl. Serv. Occ.	**.321**[5]										
Leis., Trvl. & Reltd. Prsnl. Serv. occ.									**.308**[3]		
Sales Occup.									**.304**[2]		
Custmr. Serv. Occ.										**.318**[7]	
Prcs, Plnt. & Mach. Oprtvs.									**.339**[12]	**.339**[13]	
Trnspt. & Mob. Mach. Drvrs. & oprtvs.								**.316**[14]			
Elmtry. Trds. & Reltd. Occ.								**.306**[11]			
Elmtry. Admtr. & Serv. Occ.	**.307**[1]										
N(Obs. for each GOR)	22	50	36	58	53	70	50	127	64	96	95

Notes.- The analysis consists of overall SOC rankings for mothers in two-parent-only households to obtain the likelihood of continuing education. Regional rankings are shown in square brackets. The highlighted margins correspond to the highest likelihood of continuing education for the respective occupation. Detailed margins are provided in Table C.3 (NE-North East; NW-North West; Y & H-Yorkshire & Humber; EM-East Midlands; WM-West Midlands; E-East of England; L-London; SE-South East; WL-Wales; SC-Scotland)

TABLE 4.13: Marginal impact of paternal occupation in two-parent household by Government office region

Govt. Office Regions:	NE (1)	NW (2)	Y& H (3)	EM (4)	WM (5)	E (6)	L (7)	SE (8)	SW (9)	WL (10)	SC (11)
Patrnl. SOC rank by GOR											
Corp. Mng. & Dirctr.									**.341**[23]		
Othr Mng. & Prptr.									**.352**[14]		
Sci., resch., eng. & tech. profnl.									**.342**[22]		
Hlth. Profnl.									**.345**[20]	**.345**[20]	
Teach. & Edutnl. Profnl.										**.341**[23]	
Busi., Media & Pub. Serv. Prof									**.343**[21]	**.343**[21]	
Sci., Eng. & Tech. Assoc.									**.348**[17]	**.348**[17]	
Hlth. & Soc. Care Assoc. Profnl.										**.357**[10]	
Prot. Serv. Occ.										**.342**[22]	
Cultr., Media & Sprt. Occ.										**.360**[8]	
Busi. & Pub. Serv. Assoc. Profnl.									**.347**[18]	**.347**[18]	
Admtr. Occ.									**.360**[8]		
Secrt. & Reltd. Occ.										**.365**[4]	
Skld. Agri. & Reltd. Trad.									**.359**[9]		
Skld. Mtl, Elctrl. & Elect. trad.									**.346**[19]	**.346**[19]	
Skld. Const. and Bldg Trad.										**.351**[15]	
Txtls, Prntg & Othr. Skld. Trad.										**.362**[6]	
Crng. Prsnl. Serv. Occ.										**.363**[5]	
Leis., Trvl. & Reltd. Prsnl. Serv. occ.									**.327**[3]		
Sales Occup.			**.334**[1]								
Custmr. Serv. Occ.									**.361**[7]	**.361**[7]	
Prcs, Plnt. & Mach. Oprtvs.									**.353**[13]	**.353**[13]	
Trnspt. & Mob. Mach. Drvrs. & oprtvs.									**.355**[12]		
Elmtry. Trds. & Reltd. Occ.			**.325**[9]								
Elmtry. Admtr. & Serv. Occ.									**.369**[1]		
N(Obs. for each GOR)	22	50	38	60	53	69	53	126	64	96	95

Notes.- The analysis consists of overall SOC rankings for fathers in two-parent households to obtain the likelihood of continuing education. Regional rankings are shown in square brackets. The highlighted margins correspond to the highest likelihood of continuing education for the respective occupation. Detailed margins are provided in Table C.4
(NE-North East; NW-North West; Y & H-Yorkshire & Humber; EM-East Midlands; WM-West Midlands; E-East of England; L-London; SE-South East; WL-Wales; SC-Scotland)

4.6 Discussion

The continuing effect of the COVID pandemic have highlighted a shift in education and employment. The current study re-evaluates existing patterns surrounding youth educational decisions and career aspirations. Mooi-Reci et al. (2019) had previously underlined the relationship between parental employment status and youth motivation toward education. Current findings are in line with this, but with a fine distinction—despite the significant socio-economic shifts experienced during the pandemic, the overarching impact of parental financial situations on youth continuing in educational appeared to be insignificant. This, to some extent, can be attributed to the 'shock' that the youth experienced during COVID.

Parental education has been found to be a significant driver of child academic performance up to teenage years, although its influence appears less significant by age 20. The present study confirms this trajectory, suggesting that youth decisions around university or college enrollment are less dependent on parental influence and more guided by their own preferences. However, a notable departure from previous literature is observed in parental employment shifts during COVID. The current results indicate that such shifts, especially among mothers who maintained consistent employment throughout the pandemic, display a profound influence on the likelihood of their children's continued education. The distinct roles of mothers and fathers in this context offer rich insights. Specifically, while a persistent work status among mothers during COVID tends to support continuing in educational among youth, the influence of fathers rests largely on their SES and nature of occupation. Hence, it is important to highlight the bias caused by the endogeneity in the estimation of the impact of parental work transition on a child's educational preferences. There are inherited abilities by the child which are not measurable to evaluate the causal impact. Given the household heterogeneity, parental decision taken during children's early years, shapes youth educational preferences.

In line with Anders et al. (2021), the current study underscores the diminishing aspirations among youth belonging to households where parents face an improvement in their financial situation due to the parents facing a slight improvement in their financial condition, which could be attributed to the parents becoming employed as essential workers during the pandemic.

Hence, there is a household heterogeneity in the present study in terms of changes in the work status of the parents during COVID. The sample size of fathers in the current analysis is half the percentage of mothers. There is a gender differential in terms of the maternal and paternal impact of SOC occupations during the pandemic on the youth's career preferences. The small sample size of fathers relative to mothers in the presence of household heterogeneity can lead to non-valid causal results. Another source of household heterogeneity is seen in parental decision taken when children are young, which is prominent in the present MCS data given the ethnic and cultural diversity in the UK. Hence, a larger sample size could provide a better estimate of the causal impact of the pandemic on the youth's education preferences. The current analysis has bias that is caused by the endogeneity in the estimation of the impact of parental work transition on youth educational preferences. The endogeneity is caused by inherited abilities not measurable to evaluate the causal impact.

Some variables in the current analysis lack responses and could not be considered in the analysis. This has reduced statistical power in the analysis. The reduction in the statistical power has led to type 2 errors. This could be confirmed from table 4.8 showing the margins of significance. The margins test shows that, firstly, there is a significant marginal impact of parents moving away from jobs than those who remain in the work during COVID, and, secondly, those entering work compared to those who remain working during COVID. Hence, by increasing the statistical power, the analysis could provide causal impact on youth preference to continue education.

The model specification employed in this research has allowed for an examination of the various factors shaping youth educational preferences, particularly in the 18-20 age bracket. Regional variations have pointed towards distinct patterns based on the nature of parental occupations and their adaptability to remote working conditions.

Drawing on prior studies, current research utilised a probit model to explain youth educational preference, placing particular emphasis on the differential impacts of maternal and paternal influences. The distinct outcomes for youths based on the nature of their parents' professions, especially in the context of contact intensity, provided a distinct perspective on the factors guiding educational decisions.

4.7 Conclusion

This investigation into the educational trajectories of youth aged 18 to 23 underscores the multifaceted influences shaping their decision to pursue further education during COVID. These youth faced a pandemic-induced shock to their educational pathways. The study illuminates the distinct roles of mothers and fathers, highlighting that while maternal economic transitions generally supported educational aspirations, paternal shifts often diminished them. An exceptional element, "change in financial situation", emerged as a crucial determinant in shaping these decisions. Additionally, regional maternal occupation distinction further clarified these choices. Despite the profound influences of the COVID pandemic and parental socioeconomic transitions, it appears that foundational views cultivated during youth's formative years held strongly, often overriding immediate external pressures.

Regarding future educational trajectory, it is important to address the widening inequality, both in learning environments and socio-economic contexts. A promising avenue for subsequent investigations could be an examination of the nature of youth employment and its bearing on their future career choices. Integrating additional familial parameters and cohort-specific variables would not only offer a more holistic understanding but also help policymakers with various insights to create strategies targeting youth entering the job market.

In sum, the current findings show that there is no evidence of the impact of parental work transition and change in socioeconomic conditions on the youth educational preferences during the COVID. This is accounted for by the fact that there is low statistical power that could provide a valid causal impact on youth educational preferences. Hence, this study attempts to highlight the flexibility and adaptability of the youth, navigating their educational and career paths amidst unprecedented challenges, drawing from their individual experiences and familial influences. Therefore, future studies can explore this analysis by increasing the sample size, which can lead to valid causal results in terms of socioeconomic and parental work transitions.

Chapter 5

Conclusion

The current thesis has focused on the child development starting from the early years when the child aged 3 years till the youth aged 18-23 years. By using the child production function approach by Todd & Wolpin (2007) throughout the thesis, the chapters have highlighted the importance of parental and household inputs into the child development outcomes at various stages of child development. The chapter 1, highlights the household heterogeneity in terms of language spoken to the child on the child's cognitive and behavioural outcomes. The focus is on the parental identity speaking in non-English to the child. The results corroborated with the previous findings of bilingual children and showed a converging cognitive development gap with the monolingual counterparts. The impact of parental mobility descisions between different deciles of the IMD levels in shaping child development is shown in chapter 2 of the current thesis. The outcomes are captured by the changes in the vocabulary and behavioural scores for the child aged between 7 to 14 years as a result of the movements to areas having different deciles of the IMD. The findings of the second chapter highlighted that as the child reaches teenage, there exists a negative impact of the households moving to lower decile of the IMD on the behavioural conduct and to a little extent on the vocabulary development. Hence, highlighting that movements along the deciles of IMD has a pertinent role in the determination of the behavioural conduct during teenage. It is interesting to find the impact of the early and teenage developmental outcomes and inputs on later ages i.e., during youth. The chapter 3, in the current thesis uses the pre-COVID and COVID MCS datasets to find out the impact of shocks such as the COVID in shaping educational preferences once the child reaches youth. Alongwith the parental changes in the work status and SES the chapter 3, finds out the maternal and paternal impact on the youth's likelihood to continue education and highlights that overall the youth's descisions are indifferent of their parental inputs. This is explained by the fact that youth has acquired a set of inputs from their early age that plays an important role in the youth alongwith their parental inputs.

There exists some limitations while using the MCS dataset in terms of household heterogeneity and the nature of the questions asked of the respondents. The MCS dataset used throughout consists of a broad collection of responses asked of parents about their children's development from early to teenage years. The subsequent questions are asked of the young people themselves as they reach their teenage years. The responses asked of both parents and young people are both objective and subjective. The cohort outcome is evaluated in terms of cognitive and behavioural scores. For the purpose of analysis, the cognitive and non-cognitive test scores are Z standardised with mean 0 and standard deviation 1 to compare the ability test scores across various age-groups. One such use of the subjective response is seen in the third chapter by using the qualitative responses of the youth on "whether to continue in education or not?" to analyse the impact of COVID-19 shocks on youth educational choices based on the employment status of the parents i.e., whether parents are working or not. The financial satisfaction variable also captures the change in the financial condition of the parents compared with pre-COVID by asking "how well they define their present financial situation compared to the pre-COVID", the responses being coded as "much worse off", "a little worse off", "about the same", "a little better off" and "much better off". The presence of objective responses gives the best possible judgement of the respondents but are lacking in subjective context. Therefore, there is not a common judgement level through which the parents could decide as to which choices best suites their current financial situation, resulting in bias in the analysis of outcome variables while estimating such responses. Since parents are asked to discuss their financial and employment situations during the pre- and during-COVID times, the interviews rely on the audience being genuine in their responses. However, the MCS dataset lacks in capturing the quantitative response such as the parental pre- and during-COVID income, or the threshold amount that gives a quantitative estimate of the respondents' financial situation. The outcome variable in the third chapter is subjective and asks youth opinion on their career choices during COVID. The career path variable captures the changes in youth who suffered a shock in their career path, and the response gives the estimate of youth motivation for their career path. This provides a picture of how youth have been affected by COVID. For a precise analysis, the quantitative variable capturing youth attitudes could have been the change in the academic performance scores obtained in the previous assessment. The MCS dataset lacks the information on this quantitative assessment of youth during this time period.

Although the MCS dataset is a rich collection of the parental and young people responses, where the interviewers ask respondents detailed questions about family, education, and employment status, there exists a household heterogeneity in terms of language spoken by the parents to the child and ethnicity in the MCS dataset. As some questions are targeted at specific demographics, not all questions will be answered by all participants. By considering the fact that there is a household heterogeneity in terms of language spoken by the parents to the child and also the heterogeneity in parental identity, this unequal representation of the ethnicity has led to smaller sample sizes regarding some of the questions asked of the parents such as "whether language spoken to child is other than English", which is also of interest in chapter 1. Those who speak in non-English are termed as bilingual households. Chapter 1 focuses on the bilingual respondents, categorised as 'mother only', 'father only', 'both mother and father' speaking non-English to the child. The small sample sizes of the respondents within the bilingual households has limited the scope of finding the true impact of the bilingual respondents speaking non-English to the cohort in terms of their child outcomes, though the analysis has presented a significant response to the outcome variables. To move forward with the imprecisions of the study, future studies should aim at finding the bilingual impact from a dataset consisting of a larger observations that could validate the present findings.

Chapter 3 consists of individual maternal, paternal and both maternal and paternal analysis of the impact of the change in the work status on the youth preferences to continue education. There is a small sample of the youth who responded to the question on "whether to continue in education" during the pandemic and thus a small sample size of parents who responded to questions on their employment and financial status. Apart from the insignificant parental impact on child development, the geographic regional analysis showed a significant impact on child outcomes for individual maternal and paternal analysis, but not in the case of two-parent households. This allows for future studies to focus on the parental impact that could validate the impact on child outcomes by using a larger sample of parental observations.

By virtue of the panel nature of the MCS dataset, the study accounts for the contemporaneous and previous period parental and child inputs shaping the child development path. Several studies (Bono et al., 2016; Clifton-Sprigg, 2016; Cobb-Clark et al., 2018) have also used the child production function approach which has proved to be useful to study the importance of various inputs in child development. However, the present research lacks in one important perspective of child development path: finding the causal impact of government policies on child development outcomes. Although the MCS dataset focuses on disadvantaged children, the current study fails to find the impact of the role of government policies aimed at disadvantaged children. Apart from the shortcomings of not addressing the government policies, the inclusion of the role of siblings influencing child outcomes could be an important criteria to elaborate on the influence of household dynamics in child development through their youth. Hence, in the present research, there would have been potential for evaluating the intervention of the government policies to address the impact on child development in the UK. Some of the government policies, for example Sure Start, would be a useful focus to evaluate the impact on the child cognitive and non-cognitive outcomes and contribute to the field of child development literature. At the same time, the present research provides potential for the future research in the field of education and sociology that could focus on the potential mediating role of government policies to influence parental decisions, and so highlighting the concurrent impact of the household heterogeneity shaping child outcomes.

Appendix A

The Impact of Individual Parental Bilingualism on Children's Vocabulary and Behavioural Development in the UK.

Appendix A.1 Additional description of data and analysis.

TABLE A.1: Descriptive statistics of Non-standardised outcome variables

Panel A:	Vocabulary Score			SDQ Score		
	Age 3 years (1)	Age 5 years (2)	Age 7 years (3)	Age 3 years (4)	Age 5 years (5)	Age 7 years (6)
Mean	51.46	56.45	114.42	8.78	6.37	6.49
Median	51.00	57.00	115.00	8.00	6.00	5.00
St. dev.	11.05	10.57	17.06	4.88	4.42	4.84
Min.	20.00	20.00	55.00	0.00	0.00	0.00
Max.	80.00	80.00	145.00	31.00	32.00	33.00
N	7,764	7,242	6,784	7,764	7,242	6,784
Panel B:	Vocabulary Score-(observed throughout)			SDQ Score (observed throughout)		
	Age 3 years (1)	Age 5 years (2)	Age 7 years (3)	Age 3 years (4)	Age 5 years (5)	Age 7 years (6)
Mean	52.22	57.00	114.73	8.57	6.22	6.31
Median	51.00	57.00	116.00	8.00	5.00	5.00
St. dev.	10.74	10.29	16.92	4.74	4.32	4.71
Min.	20.00	20.00	56.00	0.00	0.00	0.00
Max.	80.00	80.00	145.00	30.00	32.00	33.00
N	6,073	6,073	6,073	6,073	6,073	6,073

Notes.- The author uses the sample of pooled data to calculate descriptive statistics for child's average ages of 3, 5 and 7 years. The average age corresponds to sweeps 2, 3 and 4, respectively. The non-standardised test scores are used for the analysis. Panel A corresponds to the observations which are present for specific age group, whereas, Panel B corresponds to the observations which are present throughout each age. The descriptives shown in Panel B are approximately the same as in Panel A. Hence, there is no impact of attrition on the analysis. High mean vocabulary and total difficulty scores correspond to more vocabulary knowledge and behavioural issues respectively.

TABLE A.2: Summary statistics of outcome and control variables used in analysis, by bilingual status of parents

Mean:	Mother (1)	Father (2)	English hhld. (3)
Outcome variable			
Vocabulary score	65.76	65.57	73.47
	[34.83]	[35.11]	[30.53]
SDQ score	8.52	8.68	7.14
	[5.45]	[5.49]	[4.77]
Time-varying controls			
Child's age in months	59.92	60.07	61.23
	[20.35]	[20.47]	[20.29]
Mother's age (years)	34.00	33.79	35.37
	[5.50]	[5.54]	[5.41]
Father's age (years)	37.48	37.34	37.75
	[6.64]	[6.63]	[6.02]
OECD weighted income	0.21	0.23	0.07
(Lowest quintile)	[0.41]	[0.42]	[0.25]
		Parental time investment	
Freq. of reading to child	0.26	0.26	0.18
	[0.44]	[0.44]	[0.39]
Freq. of taking child to library	0.42	0.42	0.45
	[0.49]	[0.49]	[0.50]
Singing music with child	0.68	0.66	0.76
	[0.47]	[0.47]	[0.43]
Painting with child	0.51	0.52	0.46
	[0.50]	[0.50]	[0.50]
Time invariant controls			
Female	0.53	0.53	0.50
	[0.50]	[0.50]	[0.50]
Non-white	0.70	0.76	0.04
	[0.46]	[0.43]	[0.19]
Birth weight	3.08	3.06	3.25
	[0.63]	[0.60]	[0.58]
Mother's age at birth of CM	29.00	28.77	30.27
	[5.27]	[5.27]	[5.12]
Mother born outside UK	0.58	0.57	0.05
	[0.49]	[0.50]	[0.22]
Father born outside UK	0.57	0.65	0.05
	[0.50]	[0.48]	[0.21]
Young age arrival	0.39	0.41	0.03
	[0.49]	[0.49]	[0.16]
Natural Father resident at home	1.00	1.00	0.99
	[0.05]	[0.00]	[0.10]
Mother with Higher degree	0.45	0.42	0.47
	[0.50]	[0.49]	[0.50]
Father with Higher degree	0.45	0.43	0.42
	[0.50]	[0.50]	[0.49]
N	1,980	1,775	19,688

Notes.- The author uses the sample of pooled data to calculate descriptive statistics for individual parental language input. The standard deviations are shown in square brackets. N shows the number of observations under each category of person speaking non-English (mother, father, and both) or English only. N is same under each parental/hhld. speaking non-English/English only descriptives.

TABLE A.3: Summary statistics of lag period controls used in analysis by bilingual status of parents

Mean:(at Age 3)	Mother	Father	English hhld.
	(1)	(2)	(3)
Outcome variable			
		Cognitive score	
Grossmotor delay	0.13	0.12	0.10
	[0.33]	[0.32]	[0.29]
Finemotor delay	0.08	0.09	0.05
	[0.28]	[0.28]	[0.22]
Communication delay	0.01	0.01	0.00
	[0.08]	[0.09]	[0.04]
		Non-cognitive score	
Adaptability difficulty	0.38	0.37	0.47
	[0.49]	[0.48]	[0.50]
Approach difficulty	0.43	0.41	0.53
	[0.50]	[0.49]	[0.50]
Mood difficulty	0.35	0.36	0.30
	[0.48]	[0.48]	[0.46]
Regularity difficulty	0.37	0.39	0.19
	[0.48]	[0.49]	[0.39]
Parental time investment			
Importance of stimulating the baby	0.99	0.98	0.99
	[0.12]	[0.13]	[0.08]
Importance of cuddling the baby	1.00	1.00	1.00
	[0.06]	[0.07]	[0.04]
Importance of talking to the baby	1.00	1.00	1.00
	[0.05]	[0.05]	[0.02]
N	773	696	6,943

Notes.- The mean and standard deviations for the child development outcomes and parental inputs for 9 months are shown. Standard deviations are shown in the square bracket. N is same under each parental/hhld. speaking non-English/English only descriptives.
Following the approach of (Dickerson & Popli, 2016) and also used by (Harmon et al., 2017), the current study considers a delay in a cognitive measure if the child is unable to perform the task which 90% of the children in the same age group can achieve. Delay in more than one item is considered as a delay in that particular cognitive measure. The items are taken from Denver Developmental Screening Test (DDST) (Frankenburg & Dodds, 1967; Frankenburg et al., 1992). Similarly, following the approach of (Dickerson & Popli, 2016), the current study considers a child to be ’difficult,’ i.e., having behavioural issues if they have lower than the average score for the cohort in a non-cognitive measure. The scores are derived from the subscales of the respective non-cognitive measure. The subscales are taken from the Carey Infant Temperament Scale (Carey, 1972; Carey & McDevitt, 1978). In the current study, the average age of the child is 9 months.

Dynamics of household bilingualism on the child development outcome

Table A.4 in the Appendix represents the unweighted descriptives of non-standardised vocabulary and Total Difficulty scores in a pooled data of sweeps 2, 3 and 4. The Table A.4 shows that households which speak non-English have lower mean vocabulary scores than English only families throughout the child's age 3 to 7 years. Moreover, children whose mother speaks non-English score a higher average score of 40.60 than those from households where father speak non-English, which is 40.03 throughout age 3 to 7 years. At the same time, the Total difficulty test scores are higher for the children from hhld.s where mother speak non-English throughout the child's age 3 to 7 years, which is slightly higher from the hhld.s where father speak non-English throughout age 3 to 7 years and have children with more behavioural issues than families where father speak non-English.

TABLE A.4: Descriptives of home language spoken.

	Stayed monolingual		Mother speaks non-Eng.		Father speaks non-Eng.	
	Mean (1)	St.Dev. (2)	Mean (3)	St.Dev. (4)	Mean (5)	St.Dev. (6)
Vocabulary score	52.96	[10.38]	40.60	[10.93]	40.03	[11.10]
Total Diff. score	8.48	[4.64]	10.48	[5.61]	10.84	[5.57]
Female	0.50	[0.50]	0.56	[0.50]	0.54	[0.50]
Nonwhite	0.03	[0.16]	0.77	[0.42]	0.85	[0.36]
Higher deg.(Mother)	0.49	[0.50]	0.43	[0.50]	0.39	[0.49]
Higher deg.(Father)	0.43	[0.50]	0.42	[0.49]	0.40	[0.49]
Young migrant	0.02	[0.14]	0.42	[0.49]	0.43	[0.50]
Birthweight of CM	3.26	[0.57]	3.00	[0.60]	2.98	[0.61]
Mother's age at birth of CM	30.45	[5.00]	28.90	[5.35]	28.63	[5.32]
N	5,504		284		246	

Notes.- The descriptives of individual parental shift in language spoken is shown. The standard deviations are shown in square brackets. Stayed monolingual hhld.s are those that speak English only through ages 3 to 7.

Table A.5 shows the unweighted descriptives of dynamics of non-English hhld.s. Interestingly, at age 5, hhld.s switching to monolingual(speaking English only) first time where mother spoke non-English at age 3, has a higher mean vocabulary test score of 54.21 compared to the mean vocabulary score of 40.60(shown in Table A.4) for hhld.s where mother speak non-English throughout ages 3 to 7 years. Also, at age 7 years, the mean vocabulary test score increases to approximately 121 for hhld.s where mother spoke non-English at previous ages 3 and 5 years, which is higher than that of hhld.s where mothers speak non-English throughout and families remaining monolingual(speaking English only) throughout. Similarly, hhld.s, where fathers switch to English only first time, correspond to an increase in the child's mean vocabulary score at ages 5 and 7 years to approximately 53 and 122 than speaking non-English throughout. The mean vocabulary score increases from 53 at age 5 to 122 at age 7, respectively, where the two period increases are higher from the monolingual hhld.s where fathers speak non-English throughout ages 3 to 7 years.[1]

As per Table A.5 the children's Total difficulty scores decreases for hhld.s where the mother switches to monolingual at ages 5 and 7 years compared to hhld.s where the mother speaks non-English and English only hhld.s throughout ages 3 to 7 years. On a similar note, the Total Difficulty score of children of hhld.s where fathers switch to monolingual at age 5 and 7 years show a decrease in the Total Difficulty score as compared(column 5 of Table A.4) to the 'fathers' hhld.s, which speak non-English throughout ages 3 to 7 years.

Overall, non-English speaking hhld.s switching to monolingual at ages 5 and 7 years correspond to the higher mean vocabulary score at age 7 and lower behavioural issues at ages 5 and 7 than the families that speak non-English throughout ages 3 to 7 years.

[1]The sample includes hhld.s which are present throughout sweeps 2, 3 and 4. The sample consists of those hhld.s who switched to monolingual for the first time only. The effect is seen for switch at age 5 and age 7 years. The sample size for hhld.s changing to monolingual is around 51 for mother speaking non-English at previous age and 20 for fathers at age 5 years but further decreases to only 9 for mother speaking non-English at previous age and to 3 for father speaking non-English at age 7 years.

TABLE A.5: Dynamics of household bilingualism on the child development outcome.

Mother changing to monolingual at	Age 5		Age 7	
	Mean (1)	St.Dev. (2)	Mean (3)	St.Dev. (4)
Vocabulary score	54.21	[10.91]	121.42	[14.97]
Total Diff. score	5.81	[4.79]	6.10	[4.46]
Female	0.49	[0.50]	0.44	[0.50]
Nonwhite	0.53	[0.50]	0.75	[0.44]
Higher deg.(Mother)	0.65	[0.48]	0.65	[0.48]
Higher deg.(Father)	0.60	[0.49]	0.69	[0.47]
Young migrant	0.29	[0.46]	0.27	[0.45]
Birthweight of CM	3.15	[0.59]	3.20	[0.58]
Mother’s weight at birth of CM	30.17	[5.16]	29.85	[4.79]
N	150		48	
Father changing to monolingual at	Age 5		Age 7	
	Mean (1)	St.Dev. (2)	Mean (3)	St.Dev. (4)
Vocabulary score	53.43	[10.44]	121.55	[16.49]
Total Diff. score	5.88	[5.04]	6.60	[4.81]
Female	0.47	[0.50]	0.40	[0.50]
Nonwhite	0.66	[0.48]	0.85	[0.36]
Higher deg.(Mother)	0.62	[0.49]	0.60	[0.50]
Higher deg.(Father)	0.60	[0.49]	0.66	[0.47]
Young migrant	0.31	[0.47]	0.25	[0.44]
Birthweight of CM	3.10	[0.58]	3.09	[0.57]
Mother’s age at birth of CM	29.63	[5.23]	30.23	[5.25]
N	119		40	

Notes.- The shift in the parental language input patterns are shown for the period of analysis. As mother and father change to monolingual there is an increase in the mean vocabulary and behavioural test scores. The increase in mean vocabulary scores are almost same for mother as compared to "fathers" for age 5. Mother changing to monolingual have children with almost same magnitude of behavioural problems as compared to hhld.s where "fathers" change to monolingual.

TABLE A.6: Summary of lag Cognitive development input variables used at sweep 2 (age 3 years)

Sweep, (Year)	Age (average)	Ability Test	Test Stats (Mean, Median and Standard deviation	Description
1,(2001-2002)	9 months	Grossmotor delay	0.10,0 and 0.30	The variable is a dummy and ranges from 0 to 1, where 0 stands for 'No Delay' and 1 stands for 'Delay'. The Grossmotor delay variable is a cognitive measure that the main respondent (parent or parent figure) is asked only in sweep1 of the MCS study. The gross motor variable consists of a set of child-specific grossmotor coordination (sitting, standing, walking) questions (item) taken from age-specific Denver Developmental Screening Test (DDST) (Frankenburg & Dodds, 1967; Frankenburg et al., 1992).
1,(2001-2002)	9 months	Finemotor delay	0.05,0 and 0.22	The variable is a dummy and ranges from 0 to 1, where 0 stands for 'No Delay' and 1 stands for 'Delay'. The Finemotor delay variable is a cognitive measure that the main respondent (parent or parent figure) is asked only in sweep1 of the MCS study. The Finemotor variable consists of a set of child-specific finemotor coordination (eye and hand movements) questions (item) taken from age-specific Denver Developmental Screening Test (DDST) (Frankenburg & Dodds, 1967; Frankenburg et al., 1992).
1,(2001-2002)	9 months	Communication delay	0.00,0 and 0.04	The variable is a dummy and ranges from 0 to 1, where 0 stands for 'No Delay' and 1 stands for 'Delay'. The Communication delay variable is a cognitive measure that the main respondent (parent or parent figure) is asked only in sweep1 of the MCS study. The communication delay variable consists of a set of child-specific communication development (smiles, nods, waves) questions (item) taken from MacArthur Communicative Development Inventories (CDI). The CDI measures child receptive and productive vocabulary (Fenson et al., 1998).

TABLE A.7: Summary of lag Non-Cognitive development input variables used at sweep 2 (age 3 years)

Sweep, (Year)	Age (average)	Test Stats Ability Test	(Mean, Median and Standard deviation)	Description
1,(2001-2002)	9 months	Adaptability difficulty	0.46,0 and 0.50	The variable is a dummy and ranges from 0 to 1, where 0 stands for 'No Difficulty' and 1 stands for 'Difficulty'. The Adaptability difficulty variable is a non-cognitive measure that the main respondent (parent or parent figure) is asked only in sweep1 of the MCS study. It is measured using a set of child-specific adaptability subscales such as child's behaviour in a new place, whether child cries when put down in a different sleeping place. The subscales are taken from the Carey Infant Temperament Scale (Carey, 1972; Carey & McDevitt, 1978).
1,(2001-2002)	9 months	Approach difficulty	0.52,1 and 0.50	The variable is a dummy and ranges from 0 to 1, where 0 stands for 'No Difficulty' and 1 stands for 'Difficulty'. The Approach difficulty variable is a non-cognitive measure that the main respondent (parent or parent figure) is asked only in sweep1 of the MCS study. It is measured using a set of child-specific approach subscales such as child objects bathed in a new place, wary of strangers, shy on the first meeting with another child. The subscales are taken from the Carey Infant Temperament Scale (Carey, 1972; Carey & McDevitt, 1978).
1,(2001-2002)	9 months	Mood difficulty	0.29,0 and 0.45	The variable is a dummy and ranges from 0 to 1, where 0 stands for 'No Difficulty' and 1 stands for 'Difficulty'. It is asked of the main respondent (parent or parent figure) only in sweep1. Measurement components(Carey Infant Temperament Scale(Carey, 1972; Carey & McDevitt, 1978))- child smiles, laughs, content, and pleasant in different situations.
1,(2001-2002)	9 months	Regularity difficulty	0.99,1 and 0.12	A dummy non-cog. meas., where 0 -'No Difficulty' and 1 -'Difficulty'. It is asked of the main respondent (parent or parent figure) only in sweep1. Measurement components(Carey Infant Temperament Scale(Carey, 1972; Carey & McDevitt, 1978))- child is sleepy at about the same time, same length nap, solid food at about the same time.

TABLE A.8: Summary of Cognitive outcome variables for Sweeps 2 to 4 (age 3 to 7 years)

Sweep, (Year)	Age (average)	Test Stats Ability Test	(Mean, Median and Standard deviation	Description
2,(2003-2005)	3 years	BAS Naming vocabulary (BAS-NV)	52.4, 51.0 and 10.5	(a) The BAS-NV is a Numerical variable that ranges from 20-80. It is a verbal scale for children aged 2 years 6 months to 7 years 11 months(child's average age is 3 yr.). The current study uses age-standardised BAS Naming Vocabulary T scores. The T-scores are obtained by converting BAS ability scores using the BAS manual's conversion tables. The conversion tables consist of child's 3 month age groupings of the BAS norming sample. Connelly (2013) defines the norming selection to represent the UK 1995 national population of children. The children selected-based on the type of school attended, region of residence, free school meal entitlement, parental education, gender and ethnicity. (b) The BAS-NV assessment at age 3, consists of the interviewer showing the child a booklet of coloured pictures of objects and asks to name them. (c) The interviewer codes the answer as correct, partially correct or incorrect; if partially correct-the interviewer asks the child for clarification. The subsequent response is coded as the final answer. Each child completes a set of 16 initial questions. The interviewer proceeds or stops depending on the number of correct and incorrect answers to determine ability and proceeds to more complex items if the child finds the current set of items easy or moving with simple items if the current setting of items are too tricky (Connelly, 2013).
3,(2006-2007)	5 years	BAS Naming vocabulary (BAS-NV)	57.07, 57.0 and 10.21	(a) same as BAS-NV at sweep 2 above. The child's average age is 5 years.
4,(2008-2009)	7 years	BAS Word Reading (BAS-WR)	114.31, 115.0 and 17.04	The BAS-WR ranges from 55-145, is numerical. The test assesses Eng. Reading Ability for the children aged 5 yr. to 17 yr. 11 months(child's avg. age id 7 yr.). In the assessment, the child reads each word aloud in a block to the interviewer. There are 90 words in total where each block consists of 10 words arranged in ascending order of difficulty. If errors is greater than 7 then interviewer stops the assessment.

TABLE A.9: Summary of Non-Cognitive outcome variables for Sweeps 2 to 4 (age 3 to 7 years)

Sweep, (Year)	Age (average)	Test Stats Ability Test	(Mean, Median and Standard deviation	Description
2,(2003-2004)	3 years	Strengths & Difficulty questionnaire (SDQ)	8.7,8 and 4.8	The SDQ ranges from 0-31 and is numerical. The SDQ is a child-specific behavioural questionnaire of 3 to 16-year-old children (Goodman, 1997). It consists of 25 items (5 scales), but the current study uses the first 20 items (4 scales) of the SDQ questionnaire. The combination of the first four scales (measures) of SDQ is known as the Total Difficulty Score. The scales of the SDQ are: (a) First measure is Emotional symptoms consisting of sub-items such as- child complains headaches, seems worried, often unhappy, nervous in new situations, easily scared. (b) The second measure is Conduct problems consisting of sub-items such as child shows tantrums, generally obedient, fights/bullies with other children, argumentative with adults. (c) The third measure is the Hyperactivity scale consisting of child's specific sub-items such as the child is restless, is fidgeting, can easily be distracted, thinks before acting, completes task till the end. Also, the child shows (d) peer problems consisting of sub-items such as the child playing alone, having at least one good friend, being liked by other children, being bullied by other children, and getting on better with adults. The sub-items corresponding to a child's positive behaviour is reversely coded while generating sub-scales. The main respondent, a parent or parent figure, was asked to categorise the above sub-items as 'Not true', 'Somewhat true' or 'Certainly true'. A higher Total Difficulty score corresponds to more behavioural issues. The child's average age is 3 years.
3,(2006-2007)	5 years	Strengths & Difficulty questionnaire (SDQ)	6.3,5.0 and 4.4	The description is same as sweep 2 above The child's average age is 5 years.
4,(2008-2009)	7 years	Strengths & Difficulty questionnaire (SDQ)	6.4,5.0 and 4.8	The description is same as sweep 2 above The child's average age is 7 years.

TABLE A.10: Variable of interest for sweep 2 to 4 (age 3 to 7 years).

Sweep	Age (average)	Variable of interest	Type	Description
Waves 2, 3 and 4	3, 5 & 7 years	Who speaks non-English at home	Categorical.	The variable of interest consists of the categories such as mother, and father. The final categories are derived from a wide range of categories initially given in the dataset consisting of a mother, father, and other hhld. members. The remaining missing observations for the variable are categorised as 'English only' hhld.s based on the language spoken at home variable. The current study considers 'English only' hhld.s as the base category. The categories are coded as 0 to 2 where 0 stands for "English only hhld.", 1 for "Mother spk. non-Eng.", and 2 for "Father spk. non-Eng.".

TABLE A.11: Description of selected control variables used in analyses by bilingual status of parents

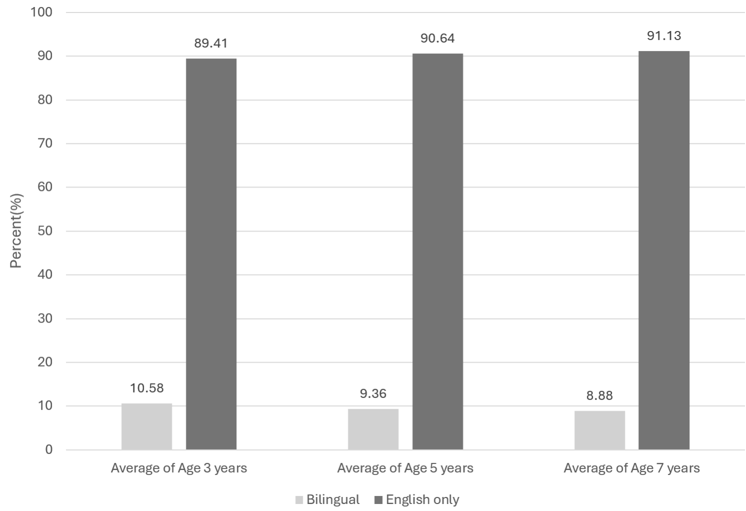
Sweep	Age (average)	Control Variable	Type	Description
2,3 and 4	3,5 & 7 years	Freq. of reading to child.	Categorical	The control variable 'Freq. of reading to child' measures parent time invst. on the child. The 'control variable' was asked only the main resp. in sweep 2 and the main and part. resp. in sweep 3 and 4. In sweeps 3 and 4, where both main and part. responded, the highest freq. of reading to a child is considered. The freq. consists of two main types: (a) Once or twice a week, the second is (b) Every day. The mult. categ. are merged as desc. in Data section footnote 24.
2,3 and 4	3,5 & 7 years	Freq. of taking child to library	Categorical	The control variable 'Freq. of taking the child to lib.' measures parent time invst. on the child. The 'control variable' was asked to main or part. resp. in sweeps 2, 3 and 4. The freq. consists of two main types: (a) Not at all, the second is (b) Once a week or more. The mult. categ. are merged as desc. in Data section footnote 24.
2,3 and 4	3,5 & 7 years	Freq. of singing music with child	Categorical	The control variable 'Freq. of singing music with child' measures parent time invst. on the child. The 'control variable' was asked to only the main resp. in sweep 2 and both main and part. respondents in sweeps 3 and 4. In sweeps 3 and 4, where both main and partner responded, the highest frequency of singing with a child is considered. The freq. consists of two main types: (a) Not at all, the second is (b) Every day. The multiple categories are merged as described in Data section footnote 24.
2, 3 and 4	3,5 & 7 years	Freq. of painting with child	Categorical	The control variable 'Freq. of painting with child' measures parent time invest. on the child. The freq. is measured same as "freq. of singing". The freq. consists of two main types: (a) Not at all, the second is (b) Every day. The multiple categories are merged as described in Data section footnote 24.

TABLE A.12: Description of selected control variables used in analyses by bilingual status of parents (continued)

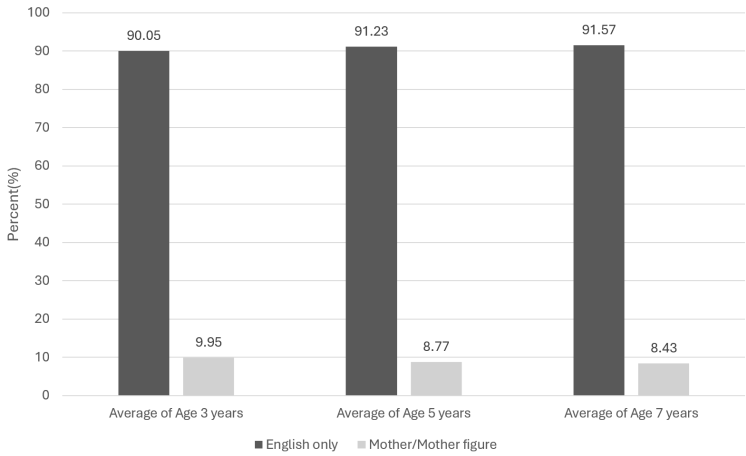
Sweep	Age (average)	Control Variable	Type	Description
2, 3 and 4	3,5 & 7 years	OECD weighted income	Categ.	In the current study, the control variable ’OECD weighted income quintile.’ is a weighted income measure for the whole UK analysis. The variable is derived by collapsing weighted equivalent weekly family income into quintile based upon UK income distribution. The OECD equivalent weekly family income is the net disposable weekly family income. It is derived by dividing the total net family income by hhld. members according to their weights on the OECD equivalent income scale. The OECD equivalent scale assigns the hhld.’s needs relative to those of a couple with no children whose scale equals 1. As per the OECD equivalent hhld. income scale, each adult and child is assigned the weight 0.5, whereas 0.3 is assigned to a child aged under 14. The UK income distribution quintiles are Lowest, First, Second, Third, Fourth and Fifth. In the current study, I categorised the First to Fifth quintiles as ’Others’, and the ’Lowest quintile’ is the base category.
1,2	9 months & 3 years	Child Ethnicity	Dummy	A time-invar. cntrl. var., asked from the main resp. only for sweeps 1 and 2 where sweep 2 consists of new families. The current study merged 8 catg. eth. to ’White’ and ’Non-white" where the ’White’ categ. is base & remaining are Non-white.
1,2,3 and 4	9 months,3,5 & 7 years	Young age arrival	Categ.	The ’Young age arrival’ is derived by combining the year of arrival in the UK and the respdnt’s yr of birth. As Cortina & Taran (2014) the respndt. falling in the age group of arrival 15-24 years is classified as a young migrant. The remaing. age grps. are categ. as "Others".
1, 2	9 months & 3 years	Mother(Father) with Higher deg.	Dummy	The var. ’Mother(Father) with Higher deg.’ - edu. qual. of the mother(father) at the birth of a child. Mult. catg. are of two types: "Higher Deg." - Higher deg., First deg., Dipl. & A/AS level, the sec. catg. "Others" is the base catg. - O level, GCSE and Other Overseas qual. The var. is derived by combining the gender (female)/(male) of the respndt.

FIGURE A.1: Types of households and the bilingualism heterogeneity over households

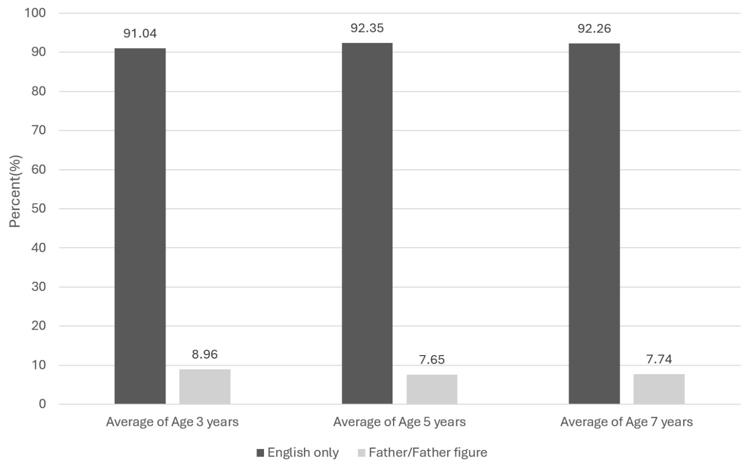
(a) Percentage distribution of types of households



(b) Percentage distribution of maternal bilingual households by child's age of 3 to 7 years.



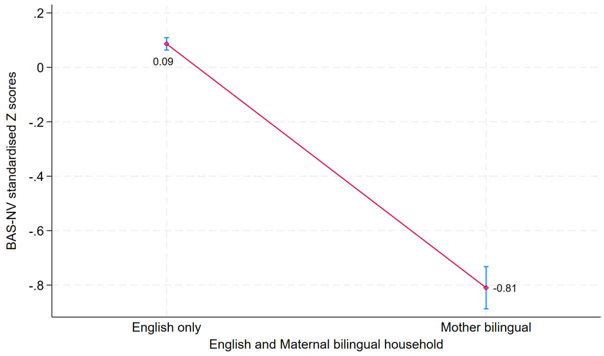
(c) Percentage distribution of paternal bilingual households by child's age of 3 to 7 years.



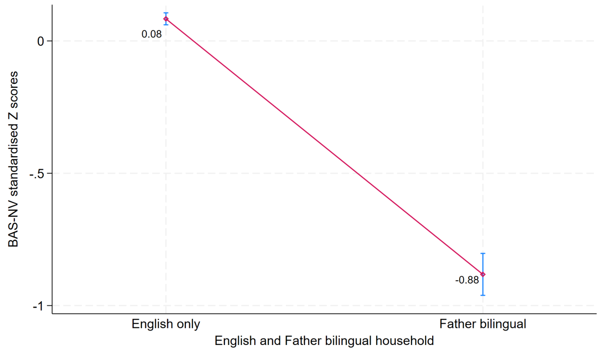
Notes.- The author's calculations are based on the English and non-English samples from the MCS dataset. N represents the number of observations for the types of households selected for the analysis. N is small for individual parental speaking non-English in the sample. The coefficients are significant in the analysis for parental analysis and standard errors are not affected by the small sample size.

FIGURE A.2: Measures of the components of Cognitive development for ages 3, 5, and 7 years.

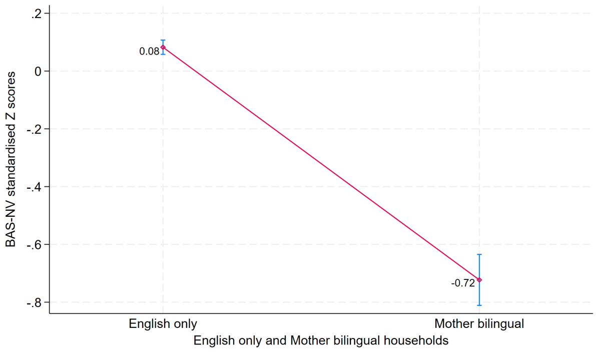
(a) Point estimates of English only and maternal bilingual cognitive scores with 95% CIs at age 3 years.



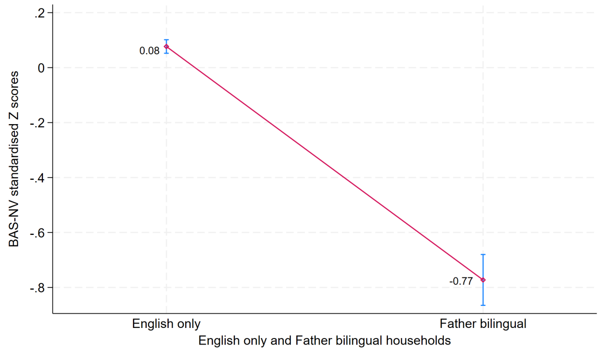
(b) Point estimates of English only and paternal bilingual cognitive scores with 95% CIs at age 3 years.



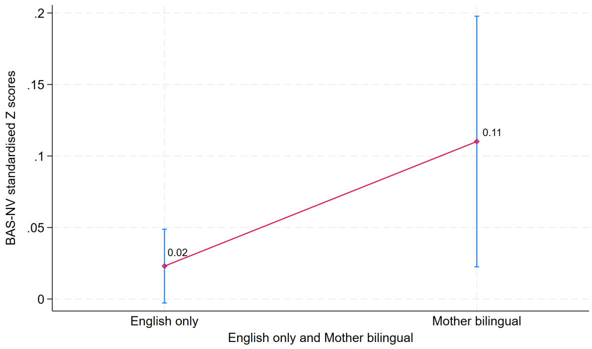
(c) Point estimates of English only and maternal bilingual cognitive scores with 95% CIs at age 5 years.



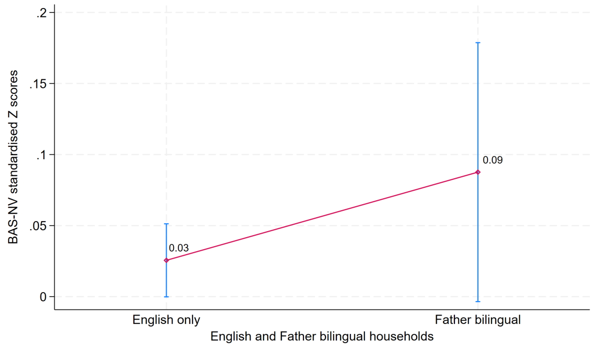
(d) Point estimates of English only and paternal bilingual cognitive scores with 95% CIs at age 5 years.



(e) Point estimates of English only and maternal bilingual cognitive scores with 95% CIs at age 7 years.



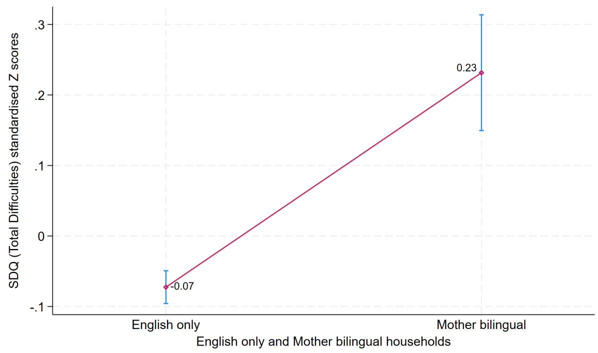
(f) Point estimates of English only and paternal bilingual cognitive scores with 95% CIs at age 7 years.



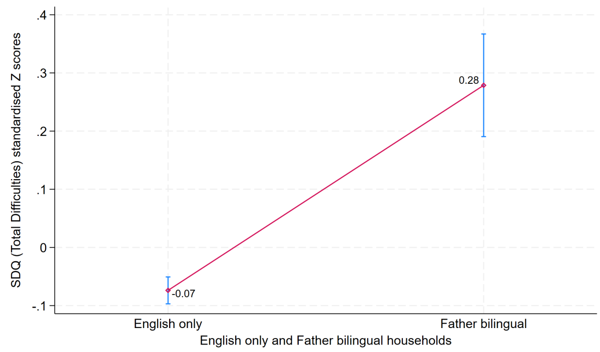
Notes.- The panels (a) to (f) denote age-specific point estimate of Z standardised Vocabulary scores with mean 0 and standard deviation 1. 'Mother speaks non-English', and 'Father speaks non-English'. At ages 3 to 5 years, English-households children vocabulary performance score is above bilingual households Vocabulary scores. But at age 7 years, children from bilingual hhld.s score higher than that of English-only counterparts.

FIGURE A.3: Measures of the components of Non-cognitive development for ages 3, 5, and 7 years.

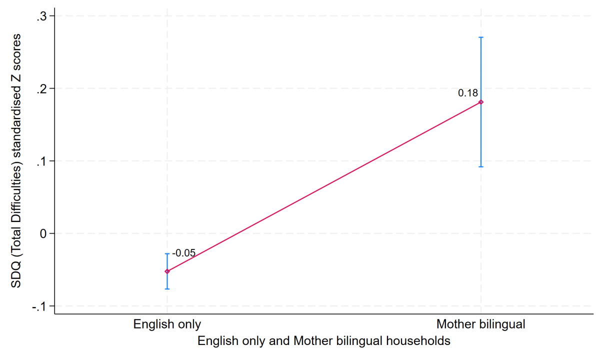
(a) Point estimates of English only and maternal bilingual non-cognitive scores with 95% CIs at age 3 years.



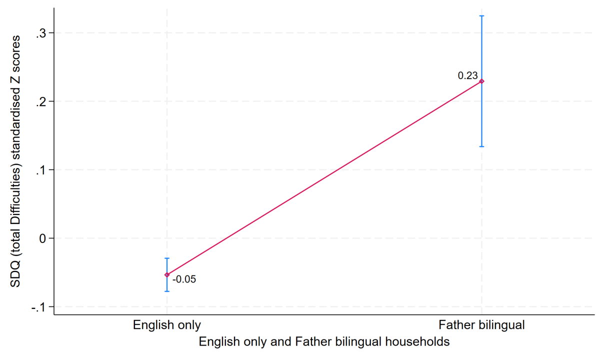
(b) Point estimates of English only and paternal bilingual non-cognitive scores with 95% CIs at age 3 years.



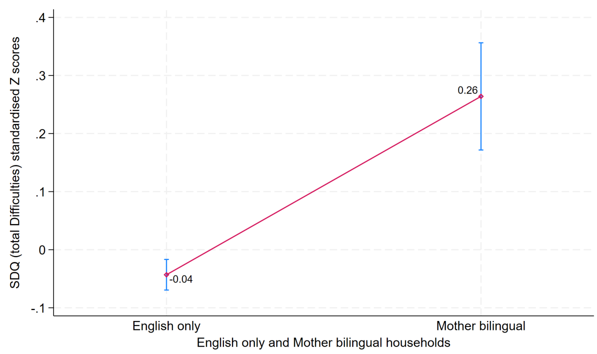
(c) Point estimates of English only and maternal bilingual non-cognitive scores with 95% CIs at age 5 years.



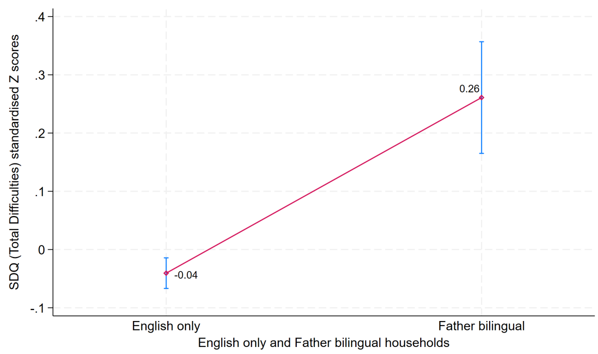
(d) Point estimates of English only and paternal bilingual non-cognitive scores with 95% CIs at age 5 years.



(e) Point estimates of English only and maternal bilingual non-cognitive scores with 95% CIs at age 7 years.



(f) Point estimates of English only and paternal bilingual non-cognitive scores with 95% CIs at age 7 years.



Notes.- The panels (a) to (f) denote age-specific point estimate of Z standardised Total Difficulty scores with mean 0 and sd 1. 'Mother speaks non-English', and 'Father speaks non-English'. The higher point estimates represent higher behavioural issues. At ages 3 to 7, father bilingual hhld.s have higher point estimates than english speaking families followed by mother bilingual households. Overall, the English-only homes have lower point estimates than the non-English hhld.s. Hence, children from non-English households have more behavioural issues than their English-only counterparts at ages 3 to 7 years.

Appendix B

Further detailed analysis of the Estimation of Child Cognitive and Non-cognitive outcomes.

Appendix B.1 Further detailed analysis of the Estimation of Child Cognitive and Non-cognitive outcomes.

Tables B.1, B.2 and Tables B.3, B.4 show the set of all contemporaneous, lagged and time-variant and time-invariant child and parental specific controls used in the age-specific, fixed effects and cumulative analysis for vocabulary and behavioural outcomes respectively.

TABLE B.1: The Effect of a person speaking in non-English on Vocabulary scores (all coefficients).

	Value added models			Indv.FE	Cum.val.add (CVA)
	Age 3	Age 5	Age 7		
	(1)	(2)	(3)	(4)	(5)
Z Vocabulary Score					
Person Speaking (n) (Base=English only)					
mother	-0.375*	-0.193***	0.205**	-0.094	-0.029
	(0.136)	(0.029)	(0.085)	(0.128)	(0.073)
father	-0.130	0.056	-0.032	0.011	0.003
	(0.103)	(0.098)	(0.081)	(0.127)	(0.082)
Person Speaking (n-1) /((n+1)-for FE estimates) (Base=English only)					
mother		0.011	0.058	0.013	-0.091
		(0.045)	(0.113)	(0.152)	(0.073)
father		-0.160	-0.006	-0.009	-0.127
		(0.104)	(0.116)	(0.154)	(0.081)
Child's cognitive development(n-1)					
Z Vocabulary score (n-1)		0.419***	0.266***		0.283***
		(0.011)	(0.020)		(0.009)
Grossmotor delay(Base=No)	-0.253***				
	(0.027)				
Finemotor delay(Base=No)	-0.103*				
	(0.048)				
Communication delay(Base=No)	0.347				
	(0.301)				
Child's noncognitive development(n-1)					
Z SDQ score (n-1)		-0.035***	-0.129***		-0.079***
		(0.010)	(0.010)		(0.008)
Adaptability difficulty(Base=No)	-0.016**				
	(0.006)				
Regularity difficulty(Base=No)	-0.088***				
	(0.020)				
Mood difficulty(Base=No)	0.020**				
	(0.008)				
Approach difficulty(Base=No)	0.047*				
	(0.022)				
Time-invariant controls					
Nonwhite (Base=White)	-0.083**	0.017	0.123*	-	0.146***
	(0.033)	(0.068)	(0.058)	-	(0.041)
Cohort sex (Base=Male)					
Female	0.244***	-0.038**	0.132***	-	0.048***
	(0.017)	(0.014)	(0.019)	-	(0.016)
Cohort birthweight(kgs)	0.084***	0.029	0.040***	-	0.027*
	(0.022)	(0.016)	(0.012)	-	(0.014)
Mother's age at birth of CM	0.014	-0.063***	-0.024*	-	-0.042***
	(0.013)	(0.012)	(0.012)	-	(0.014)
Time-variant controls					
Child's age in months	0.016***	-0.032***	-0.026***	-0.013***	-0.021***
	(0.002)	(0.004)	(0.005)	(0.005)	(0.0030)
Mother caring for child (n-1)(Base=Father)		0.057	-0.007		0.034
		(0.034)	(0.061)		(0.045)
Mother caring for child (n)(Base=Father)	0.002	0.165*	0.001	0.029	0.049
	(0.021)	(0.076)	(0.037)	(0.058)	(0.062)
Mother's Higher degree (Base=Other)	0.175***	0.194***	0.154***	-	0.174***
	(0.015)	(0.016)	(0.015)	-	(0.017)
Father's Higher degree (Base=Other)	0.179***	0.159***	0.224***	-	0.200***
	(0.009)	(0.015)	(0.023)	-	(0.018)
Mother's age	-0.006	0.069***	0.029*	0.014	0.044***
	(0.014)	(0.010)	(0.013)	(0.023)	(0.014)
Father's age	0.004***	0.004	-0.001	0.021	0.003
	(0.001)	(0.003)	(0.001)	(0.019)	(0.002)
No. of hhld. members(n-1)	-0.127***	-0.051**	-0.073***		-0.072***
	(0.022)	(0.018)	(0.020)		(0.017)
No. of hhld. members(n)	-0.031	-0.042**	-0.000	-0.016	-0.016
	(0.023)	(0.015)	(0.018)	(0.024)	(0.017)
OECD weighted income(n-1)	-0.215***	-0.123***	-0.090		-0.103***
	(0.040)	(0.022)	(0.076)		(0.036)
OECD weighted income(n)	-0.142***	-0.048	-0.021	0.044	-0.051
	(0.023)	(0.033)	(0.048)	(0.046)	(0.037)
Natural father resident(n-1)	-0.012	0.127***	0.408**		0.201
	(0.186)	(0.027)	(0.157)		(0.214)
Natural father resident(n)	0.228*	-0.151*	-0.147***	-0.493	-0.108
	(0.107)	(0.079)	(0.031)	(0.363)	(0.207)
Main carer is Young migrant(Base=Others)	-0.177**	0.005	-0.004	-	-0.005
	(0.074)	(0.021)	(0.023)	-	(0.048)
Mother born in the UK(Base=No)	0.004	-0.054	0.083**	-	0.033
	(0.065)	(0.039)	(0.026)	-	(0.037)
Father born in the UK(Base=No)	-0.111*	-0.065***	0.104**	-	0.012
	(0.053)	(0.018)	(0.034)	-	(0.040)

*** Significant at the 1 percent level.** Significant at the 5 percent level.* Significant at the 10 percent level.

TABLE B.2: The Effect of a person speaking in non-English on Vocabulary scores (all coefficients)(continued).

	Value added models			Indv.FE	Cum.val.add (CVA)
	Age 3 (1)	Age 5 (2)	Age 7 (3)	(4)	(5)
Z Vocabulary Score(continued)					
Parental time investment(n-1)					
Importance of stimulation(Base=No)	0.259**				
	(0.088)				
Importance of cuddling(Base=No)	0.178				
	(0.334)				
Importance of talking(Base=No)	-0.418				
	(0.654)				
Importance of regularity(Base=No)	-0.008				
	(0.026)				
Everyday reading to child(Base=once or twice a week)		-0.031*	-0.056		-0.042*
		(0.014)	(0.039)		(0.022)
Once a week or more taking child to library(Base=Not at all)		0.053	0.055**		0.057***
		(0.035)	(0.018)		(0.019)
Everyday listening to music with child(Base=Not at all)		0.089***	-0.005		0.038*
		(0.013)	(0.030)		(0.022)
Everyday painting with child(Base=Not at all)		0.005	-0.075**		-0.036*
		(0.027)	(0.024)		(0.020)
Parental time investment(n)					
Everyday reading to child(Base=once or twice a week)				-0.025	
				(0.026)	
Once a week or more taking child to library(Base=Not at all)				0.050*	
				(0.026)	
Everyday listening to music with child(Base=Not at all)				-0.030	
				(0.027)	
Everyday painting with child(Base=Not at all)				0.033	-0.023
				(0.023)	(0.019)
Government Office Region(Base=North East)					
North West	-0.052	0.196***	-0.110***	0.030	0.060
	(0.074)	(0.008)	(0.010)	(0.520)	(0.054)
Yorkshire & the Humber	-0.119	0.182***	-0.126***	0.185	0.026
	(0.112)	(0.016)	(0.035)	(0.468)	(0.054)
East Midlands	0.042	0.200***	-0.193***	0.448	0.016
	(0.110)	(0.023)	(0.014)	(0.421)	(0.054)
West Midlands	0.031	0.052	-0.078	-0.011	-0.024
	(0.047)	(0.035)	(0.045)	(0.381)	(0.055)
East of England	0.034	0.229***	-0.139***	0.071	0.012
	(0.074)	(0.062)	(0.036)	(0.457)	(0.054)
London	-0.028	0.100	0.093	-0.083	0.085
	(0.088)	(0.055)	(0.086)	(0.449)	(0.055)
South East	0.005	0.111***	-0.101***	-0.254	-0.000
	(0.076)	(0.025)	(0.012)	(0.454)	(0.051)
South West	0.229***	0.122***	-0.148***	0.131	-0.007
	(0.054)	(0.017)	(0.032)	(0.493)	(0.054)
Wales	0.060	0.005	-0.315***	0.137	-0.161***
	(0.079)	(0.020)	(0.016)	(0.484)	(0.050)
Scotland	0.145*	0.128**	-0.229***	-0.223	-0.058
	(0.073)	(0.043)	(0.009)	(0.504)	(0.051)
Northern Ireland	0.089	0.136***	-0.369***	0.241	-0.116**
	(0.074)	(0.031)	(0.021)	(1.157)	(0.053)
Sweep2(d2)				-0.234**	-
				(0.095)	-
Sweep3(d3)				-	-0.391***
				-	(0.069)
Sweep4(d4)				-	-
				-	-
Constant	-1.015	1.127***	1.784***	0.026	1.308***
	(0.622)	(0.147)	(0.234)	(1.136)	(0.274)
R-squared	0.173	0.288	0.211	0.005	0.136
N	7,765	7,243	6,784		
Observations				13,260	13,260

Notes.- Robust s.e in parenthesis are clustered at family level (stratum within each countries of the UK) for VA and at individual level for cumulative value added (CVA). The value added and CVA regressions are weighted at overall population level to avoid a non-response rate. Outcome is standardised Z vocabulary score with mean 0 and standard deviation 1 for age specific and CVA models. The individual fixed effects (FE) model is estimated using unbalanced panel of child wave observations. The CVA model includes sweep dummies. Lags are added for all time variant controls in case of value added model and of selected time variant controls for CVA. Individual FE also includes lead of main variable of interest to check for exogeneity. "-" shows omitted because of collinearity.*** Significant at the 1 percent level.** Significant at the 5 percent level.* Significant at the 10 percent level.

TABLE B.3: The Effect of a person speaking in non-English on SDQ scores (all coefficients).

	Value added models			Indv.FE	Cum.val.add (CVA)
	Age 3	Age 5	Age 7		
	(1)	(2)	(3)	(4)	(5)
Z SDQ Score					
Person Speaking (n) (Base=English only)					
mother	0.050*	-0.011	0.263***	0.144	0.092
	(0.026)	(0.139)	(0.073)	(0.103)	(0.062)
father	0.129**	0.152	-0.111***	0.074	0.046
	(0.049)	(0.163)	(0.024)	(0.104)	(0.069)
Person Speaking (n-1) /((n+1)-for FE estimates) (Base=English only)					
mother		-0.079	-0.207***	0.022	-0.028
		(0.099)	(0.049)	(0.132)	(0.059)
father		-0.072	-0.030	-0.021	-0.082
		(0.097)	(0.062)	(0.142)	(0.067)
Child's cognitive development(n-1)					
Z Vocabulary score (n-1)		-0.065***	-0.025***		-0.039***
		(0.005)	(0.004)		(0.008)
Grossmotor delay(Base=No)	0.084***				
	(0.017)				
Finemotor delay(Base=No)	0.054				
	(0.034)				
Communication delay(Base=No)	-0.134				
	(0.145)				
Child's noncognitive development(n-1)					
Z SDQ score (n-1)		0.494***	0.645***		0.401***
		(0.008)	(0.020)		(0.010)
Adaptability difficulty(Base=No)	-0.088***				
	(0.011)				
Regularity difficulty(Base=No)	0.233***				
	(0.035)				
Mood difficulty(Base=No)	0.229***				
	(0.016)				
Approach difficulty(Base=No)	-0.048*				
	(0.022)				
Time-invariant controls					
Nonwhite (Base=White)	0.021	0.068	-0.053	-	0.013
	(0.037)	(0.050)	(0.030)	-	(0.041)
Cohort sex (Base=Male)					
Female	-0.200***	-0.080***	-0.143***	-	-0.144***
	(0.012)	(0.023)	(0.007)	-	(0.016)
Cohort birthweight(kgs)	-0.056***	-0.044**	-0.037*	-	-0.056***
	(0.010)	(0.015)	(0.018)	-	(0.014)
Mother's age at birth of CM	-0.020	0.002	0.002	-	0.012
	(0.052)	(0.015)	(0.026)	-	(0.018)
Time-variant controls					
Child's age in months	-0.011	-0.020**	-0.005*	-0.013***	-0.013***
	(0.006)	(0.006)	(0.002)	(0.004)	(0.003)
Mother caring for child (n-1)(Base=Father)		-0.009	-0.043		-0.049
		(0.026)	(0.194)		(0.042)
Mother caring for child (n)(Base=Father)	0.090	-0.176	0.202***		0.036
	(0.051)	(0.100)	(0.040)		(0.059)
Mother's Higher degree (Base=Other)	-0.241***	-0.082***	-0.034	-	-0.091***
	(0.008)	(0.011)	(0.027)	-	(0.018)
Father's Higher degree (Base=Other)	-0.154***	-0.047***	-0.070***	-	-0.082***
	(0.018)	(0.007)	(0.016)	-	(0.018)
Mother's age	0.005	-0.005	-0.010	-0.020	-0.019
	(0.051)	(0.015)	(0.027)	(0.024)	(0.018)
Father's age	-0.007***	-0.001	0.003	0.003	-0.001
	(0.001)	(0.001)	(0.002)	(0.019)	(0.010)
No. of hhld. members(n-1)	-0.029	-0.040***	-0.056***		-0.048***
	(0.020)	(0.009)	(0.014)		(0.018)
No. of hhld. members(n)	0.040**	0.013	0.069***	0.092***	0.042**
	(0.014)	(0.021)	(0.009)	(0.028)	(0.018)
OECD weighted income(n-1)	0.289***	0.111***	0.119		0.161***
	(0.041)	(0.022)	(0.070)		(0.036)
OECD weighted income(n)	0.161**	0.041	0.015	-0.057	0.067*
	(0.049)	(0.037)	(0.024)	(0.049)	(0.038)
Natural father resident(n-1)	-0.228	-0.929***	-0.491***		-0.373
	(0.370)	(0.206)	(0.051)		(0.408)
Natural father resident(n)	-0.110	0.878***	0.371***	-0.448	0.159
	(0.143)	(0.205)	(0.054)	(0.833)	(0.413)
Main carer is Young migrant(Base=Others)	0.023	0.003	0.088*	-	0.008
	(0.067)	(0.107)	(0.042)	-	(0.049)
Mother born in the UK(Base=No)	-0.085	0.064	-0.009	-	0.018
	(0.048)	(0.039)	(0.026)	-	(0.039)
Father born in the UK(Base=No)	0.021	0.027	-0.011	-	0.026
	(0.044)	(0.017)	(0.058)	-	(0.038)

*** Significant at the 1 percent level.** Significant at the 5 percent level.* Significant at the 10 percent level.

TABLE B.4: The Effect of a person speaking in non-English on internalising scores (all coefficients)(continued).

	Value added models			Indv.FE	Cum.val.add (CVA)
	Age 3	Age 5	Age 7		
	(1)	(2)	(3)	(4)	(5)
Z SDQ Score(continued)					
Parental time investment(n-1)					
Importance of stimulation(Base=No)	-0.337***				
	(0.061)				
Importance of cuddling(Base=No)	0.565				
	(0.501)				
Importance of talking(Base=No)	0.004				
	(0.581)				
Importance of regularity(Base=No)	-0.077				
	(0.047)				
Everyday reading to child(Base=once or twice a week)		-0.017	-0.035		-0.024
		(0.012)	(0.025)		(0.021)
Once a week or more taking child to library(Base=Not at all)		-0.027**	-0.030**		-0.016
		(0.011)	(0.010)		(0.018)
Everyday listening to music with child(Base=Not at all)		0.001	-0.041*		-0.028
		(0.037)	(0.019)		(0.020)
Everyday painting with child(Base=Not at all)		-0.072**	-0.030**		-0.042**
		(0.022)	(0.009)		(0.018)
Parental time investment(n)					
Everyday reading to child(Base=once or twice a week)				0.029	0.039*
				(0.026)	(0.023)
Once a week or more taking child to library(Base=Not at all)				0.024	-0.048***
				(0.023)	(0.017)
Everyday listening to music with child(Base=Not at all)				0.023	0.013
				(0.025)	(0.018)
Everyday painting with child(Base=Not at all)				-0.007	-0.048***
				(0.022)	(0.018)
Government Office Region(Base=North East)					
North West	-0.091**	0.051	-0.084	0.085	0.002
	(0.036)	(0.044)	(0.085)	(0.358)	(0.055)
Yorkshire & the Humber	0.062	-0.046	0.045	0.347	-0.013
	(0.083)	(0.040)	(0.056)	(0.320)	(0.056)
East Midlands	-0.064	0.011	-0.033***	-0.085	0.023
	(0.067)	(0.032)	(0.007)	(0.272)	(0.056)
West Midlands	-0.018	0.087***	0.057*	0.012	0.084
	(0.061)	(0.018)	(0.025)	(0.249)	(0.057)
East of England	-0.066	0.000	0.025	-0.215	0.031
	(0.066)	(0.003)	(0.020)	(0.270)	(0.055)
London	-0.166**	-0.045*	-0.016	0.092	-0.020
	(0.060)	(0.021)	(0.015)	(0.271)	(0.056)
South East	-0.070	0.028**	-0.043	0.007	0.001
	(0.074)	(0.011)	(0.045)	(0.275)	(0.051)
South West	-0.077	0.059***	-0.003	-0.024	0.048
	(0.079)	(0.003)	(0.042)	(0.280)	(0.055)
Wales	-0.162	0.094**	-0.071	0.032	0.006
	(0.129)	(0.032)	(0.049)	(0.259)	(0.051)
Scotland	-0.106	0.098***	-0.105**	0.175	0.017
	(0.094)	(0.007)	(0.034)	(0.357)	(0.051)
Northern Ireland	-0.178	-0.003	-0.054	-0.440	-0.026
	(0.097)	(0.040)	(0.031)	(1.158)	(0.053)
Sweep2(d2)				-0.379***	-
				(0.091)	-
Sweep3(d3)				-	-0.300***
				-	(0.069)
Sweep4(d4)				-	-
				-	-
Constant	1.522**	1.991***	0.887***	1.364	2.101***
	(0.551)	(0.395)	(0.260)	(1.350)	(0.287)
R-squared	0.126	0.326	0.440	0.010	0.123
N	7,765	7,243	6,784		
Observations				13,260	13,260

Notes.- Robust s.e in parenthesis are clustered at family level (stratum within each countries of the UK) for VA and at individual level for cumulative value added (CVA). The value added and CVA regressions are weighted at overall population level to avoid a non-response rate. Outcome is standardised Z internalising score with mean 0 and standard deviation 1 for age specific and CVA models. The individual fixed effects (FE) model is estimated using unbalanced panel of child wave observations. The CVA model includes sweep dummies. Lags are added for all time variant controls in case of value added model and of selected time variant controls for CVA. Individual FE also includes lead of main variable of interest to check for exogeneity."-" shows omitted because of collinearity. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

Table B.5 shows the individual SDQ scores for each age group. At age 3, mother being bilingual has no significant impact on the child's emotional problems compared to the children belonging to English only speaking household.

For age 5, there exists a decrease in conduct problems due to lagged paternal bilingualism whereas, an increase in peer problems due to contemporaneous maternal bilingualism. There is a significant increase in conduct problems due to contemporaneous maternal bilingualism at age 7 but an insignificant decline in peer relationship problems due to lagged maternal and paternal bilingualism at age 7. Also, there exists an increase in the emotional, conduct, hyperactivity and peer relations due to lagged outcomes from age 5 to 7.

Table B.6 shows the SDQ Prosocial score; at age 3, mother bilingual has significant negative impact on the Prosocial score, but the father bilingual have a significant positive effect on the prosocial development. But, at the successive age group of 5 and 7 years, there is no significant impact of parental bilingualism on the prosocial score. There is a significant positive impact of lagged prosocial scores for ages 5 to 7. Therefore, bilingual children's lagged prosocial development positively influences current prosocial behaviour.

TABLE B.5: The Effects of person speaking in non-English on individual SDQ scores.

	Age 3 (1)	Age 5 (2)	Age 7 (3)
Z SDQ Emotional Score			
Person Speaking (n) (Base=English only)			
mother	-0.136	0.006	0.068
	(0.137)	(0.080)	(0.090)
father	0.248*	0.224	-0.024
	(0.122)	(0.195)	(0.098)
Person Speaking (n-1) (Base=English only)			
mother		-0.099	-0.079
		(0.061)	(0.074)
father		-0.064	0.063
		(0.089)	(0.060)
Emotional(n-1)		0.415***	0.467***
		(0.009)	(0.015)
Constant	0.091	0.753**	-0.221
	(0.118)	(0.326)	(0.255)
R-squared	0.054	0.175	0.266
Z SDQ Conduct Score			
Person Speaking (n) (Base=English only)			
mother	0.048	0.033	0.369***
	(0.047)	(0.089)	(0.047)
father	-0.059	0.105	-0.235***
	(0.046)	(0.075)	(0.030)
Person Speaking (n-1) (Base=English only)			
mother		-2.18e-05	-0.197**
		(0.081)	(0.060)
father		-0.182**	0.033
		(0.063)	(0.058)
Conduct(n-1)		0.487***	0.570***
		(0.011)	(0.010)
Constant	1.927***	1.392***	-0.149
	(0.537)	(0.256)	(0.290)
R-squared	0.082	0.240	0.355
Z SDQ Hyperactivity Score			
Person Speaking (n) (Base=English only)			
mother	-0.059	-0.122	0.085
	(0.077)	(0.127)	(0.077)
father	0.195**	0.011	-0.054
	(0.061)	(0.121)	(0.063)
Person Speaking (n-1) (Base=English only)			
mother		-0.007	-0.067
		(0.054)	(0.072)
father		0.053	0.015
		(0.076)	(0.106)
Hyperactivity(n-1)		0.406***	0.630***
		(0.003)	(0.009)
Constant	1.438**	1.615***	1.761***
	(0.569)	(0.163)	(0.084)
R-squared	0.085	0.297	0.417
Z SDQ Peer relationship Score			
Person Speaking (n) (Base=English only)			
mother	0.249**	0.286**	-0.012
	(0.092)	(0.099)	(0.061)
father	-0.056	-0.023	0.099
	(0.079)	(0.084)	(0.095)
Person Speaking (n-1) (Base=English only)			
mother		-0.134	-0.125
		(0.104)	(0.086)
father		-0.013	0.059
		(0.077)	(0.049)
Peer(n-1)		0.391***	0.494***
		(0.009)	(0.027)
Constant	1.937***	1.903***	1.207***
	(0.513)	(0.205)	(0.264)
R-squared	0.065	0.157	0.296
Observations	7,765	7,243	6,784

Notes.- Robust s.e. in parenthesis are clustered at family level (stratum within each countries of the UK) for age-specific value-added model. The value added(VA) regressions are weighted at overall pop. level to avoid non-resp. rate. The outcome is standardized Z SDQ score with mean 0 and sd. 1 for age-specific models. The controls used are same as that used in Tables B.3 to B.4. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

TABLE B.6: The Effects of person speaking in non-English on Prosocial scores.

	Age 3 (1)	Age 5 (2)	Age 7 (3)
Z SDQ Prosocial Score			
Person Speaking (n) (Base=English only)			
mother	-0.156**	-0.040	-0.076
	(0.061)	(0.133)	(0.092)
father	0.237***	0.022	0.053
	(0.064)	(0.126)	(0.085)
Person Speaking (n-1) (Base=English only)			
mother		0.024	-0.054
		(0.070)	(0.117)
father		0.056	0.269**
		(0.113)	(0.100)
Prosocial(n-1)		0.435***	0.465***
		(0.008)	(0.023)
Constant	-1.214*	-1.140***	-0.269
	(0.620)	(0.208)	(0.333)
R-squared	0.045	0.185	0.267
Observations	7,732	7,209	6,780

Notes.- Robust standard errors in parenthesis are clustered at family level (stratum within each countries of the UK) for age-specific value added model. The value added regressions are weighted at overall population level to avoid non-response rate. Outcome is standardized Z Prosocial score with mean 0 and standard deviation 1 for age specific model. The age specific controls used in the analysis include child's age in months, ethnicity, gender, birth-weight, mother's age at the birth of child, young migrant, natural father resident, weighted household income, mother's and father's highest educational qualification, mother father born in the UK, government office regions, parent caring most of the time, parental time investments in the form of frequency of reading to child, taking child to library, playing music and painting with child. Lags are added for all time variant controls in case of VA model.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE B.7: The change in child vocabulary scores between sweeps 2 (age 3) and 4 (age 7) for families moving/staying at sweep 3 (age 5).

			Overall & Indv. IMDs				
	All reason (1)	Inc. (2)	Hlth. (3)	Edu. (4)	House (5)	Crime (6)	Env. (7)
OLS with controls							
Δ(IMD) (Base=Stayed at same IMD or different IMD)							
Change in IMD	-0.107**	0.011	-0.076*	0.012	-0.035	-0.030	0.003
	(0.048)	(0.041)	(0.046)	(0.020)	(0.039)	(0.031)	(0.022)
$\Delta\ IMD^2$	0.017**	-0.004	0.014*	-0.007**	0.002	0.002	-0.006**
	(0.008)	(0.008)	(0.008)	(0.003)	(0.008)	(0.005)	(0.003)
-ve Δ IMD	-0.046	-0.013	0.016	-0.120	0.003	0.072	-0.019
movement (Base=+ve change)	(0.234)	(0.108)	(0.105)	(0.095)	(0.091)	(0.091)	(0.102)
Δ IMD	0.093	-0.048	0.052	-0.076***	0.053	0.051	0.004
for lower IMD (M#low IMD)	(0.172)	(0.066)	(0.087)	(0.025)	(0.062)	(0.045)	(0.048)
$\Delta\ IMD^2$	-0.015	-0.003	-0.025*	0.009**	0.001	-0.010	0.005
for lower IMD (M#M#low IMD)	(0.025)	(0.013)	(0.014)	(0.004)	(0.012)	(0.008)	(0.010)
Lagged z test scores (sweep2)	-0.291***	-0.292***	-0.290***	-0.289***	-0.292***	-0.289***	-0.289***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Age 3 IMD (sweep2)	-0.016	-0.014	-0.016	0.007	-0.023	-0.013	-0.007
	(0.015)	(0.014)	(0.018)	(0.011)	(0.015)	(0.013)	(0.012)
Current IMD (sweep4)	0.024	0.027*	0.022	-0.006	0.033**	0.007	0.003
	(0.015)	(0.014)	(0.018)	(0.011)	(0.015)	(0.013)	(0.012)
Nonwhite child	-0.194***	-0.202***	-0.190***	-0.192***	-0.193***	-0.184***	-0.184***
(Base=White)	(0.062)	(0.062)	(0.061)	(0.062)	(0.061)	(0.062)	(0.062)
Child suffering from	-0.004	-0.001	-0.002	-0.004	-0.001	-0.000	-0.001
long-term illness	(0.036)	(0.036)	(0.037)	(0.037)	(0.036)	(0.037)	(0.037)
Having sibling at age 3	-0.030	-0.030	-0.031	-0.028	-0.031	-0.029	-0.027
	(0.033)	(0.033)	(0.033)	(0.032)	(0.033)	(0.032)	(0.033)
Child's age(years)	-0.211	-0.201	-0.212	-0.204	-0.207	-0.199	-0.192
	(0.220)	(0.220)	(0.220)	(0.218)	(0.218)	(0.219)	(0.219)
Child's age(squared)	0.001	0.001	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Birth weight of the child(Kgs)	0.066***	0.066***	0.067***	0.068***	0.066***	0.066***	0.068***
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
Mother suffering from illness	-0.038	-0.037	-0.039	-0.038	-0.038	-0.040	-0.039
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
English spoken to child at age3	0.300***	0.298***	0.301***	0.300***	0.298***	0.296***	0.297***
	(0.064)	(0.064)	(0.064)	(0.064)	(0.064)	(0.064)	(0.064)
Two-parent household at age 3	0.017	0.017	0.016	0.015	0.018	0.028	0.022
	(0.044)	(0.044)	(0.044)	(0.044)	(0.044)	(0.044)	(0.044)
OECD inc. of hhld at age 3	0.001***	0.001***	0.001***	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mother age at the birth of child	0.006**	0.006**	0.006**	0.006**	0.006**	0.007**	0.006**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Mother born in the UK	0.235***	0.231***	0.234***	0.233***	0.227***	0.232***	0.234***
	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)
Nonwhite mother	-0.023	-0.027	-0.024	-0.019	-0.022	-0.022	-0.018
(Base=White)	(0.054)	(0.054)	(0.054)	(0.055)	(0.055)	(0.055)	(0.055)
Mother having	-0.141***	-0.141***	-0.143***	-0.149***	-0.136***	-0.153***	-0.149***
A/AS/GCSE/other qual.	(0.031)	(0.031)	(0.031)	(0.030)	(0.031)	(0.031)	(0.030)
Mother having	-0.305***	-0.305***	-0.311***	-0.321***	-0.301***	-0.331***	-0.326***
no academic qual.	(0.055)	(0.055)	(0.055)	(0.053)	(0.055)	(0.054)	(0.054)
Constant	9.835	9.384	9.864	9.566	9.654	9.327	9.019
	(9.529)	(9.553)	(9.554)	(9.474)	(9.475)	(9.463)	(9.519)
R^2	0.165	0.165	0.165	0.167	0.165	0.165	0.164
N	6,058	6,058	6,058	6,058	6,058	6,058	6,058

Notes.- The robust standard errors are in parenthesis. Child's age in months is controlled. Outcome is difference in standardised Z vocabulary scores between sweeps 2 (age 3) to 4 (age 7) with mean 0 and standard deviation 1. The variable of interest is the movement of families between IMDs for sweeps 2 (age 3) and 3 (age 5). The analysis also controls for child's age in months at sweeps 4(age 7), current period IMD level. The lagged period controls used consist of respective lagged period (sweep 2) outcomes, sweep 2 IMD. Child and household level controls consist of child's and mother's ethnicity, number of siblings, OECD weighted household income, mother's highest educational qualification, language spoken in household, mother having long term illness, child having long term illness, whether mother is born in the UK, child birth-weight, whether single or two parent household, age of mother at birth of child.
*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

TABLE B.8: The change in child vocabulary scores between sweeps 2 (age 3) and 5 (age 11) for families moving/staying at sweep 3 (age 5).

			Overall & Indv. IMDs				
	All reason (1)	Inc. (2)	Hlth (3)	Edu. (4)	House (5)	Crime (6)	Env. (7)
OLS with controls							
Δ(IMD) (Base=Stayed at same IMD or different IMD)							
Change in IMD	-0.050	0.041	-0.044	0.012	0.005	-0.041*	-0.049*
	(0.050)	(0.044)	(0.062)	(0.033)	(0.035)	(0.025)	(0.027)
Δ IMD^2	0.000	-0.016	0.002	-0.009	-0.014*	0.003	0.007*
	(0.008)	(0.010)	(0.011)	(0.008)	(0.006)	(0.004)	(0.004)
-ve Δ IMD	0.128	0.032	0.033	-0.014	0.067	0.051	0.047
movement (Base=+ve change)	(0.171)	(0.077)	(0.072)	(0.075)	(0.069)	(0.071)	(0.077)
Δ IMD	0.133	-0.013	0.025	-0.027	0.008	0.030	0.084**
for lower IMD (M#low IMD)	(0.132)	(0.052)	(0.078)	(0.045)	(0.035)	(0.033)	(0.035)
Δ IMD^2	0.013	0.022**	-0.006	0.013**	0.013	-0.005	-0.001
for lower IMD (M#M#low IMD)	(0.018)	(0.011)	(0.015)	(0.006)	(0.010)	(0.006)	(0.006)
Lagged z test scores (sweep2)	-0.620***	-0.621***	-0.619***	-0.619***	-0.621***	-0.619***	-0.618***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Age 3 IMD (sweep2)	0.006	0.012	0.001	-0.001	0.003	-0.005	-0.002
	(0.009)	(0.009)	(0.009)	(0.007)	(0.008)	(0.008)	(0.008)
Current IMD (sweep5)	0.001	-0.000	-0.002	0.016**	0.007	-0.004	-0.006
	(0.008)	(0.009)	(0.009)	(0.007)	(0.008)	(0.007)	(0.007)
Nonwhite child	-0.127**	-0.135**	-0.121**	-0.133**	-0.126**	-0.115*	-0.115*
(Base=White)	(0.061)	(0.061)	(0.061)	(0.060)	(0.061)	(0.061)	(0.061)
Child suffering from	-0.018	-0.017	-0.017	-0.024	-0.018	-0.016	-0.017
long-term illness	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)
Having sibling at age 3	0.031	0.031	0.031	0.030	0.028	0.030	0.033
	(0.031)	(0.031)	(0.031)	(0.030)	(0.030)	(0.031)	(0.031)
Child's age(years)	0.035	0.037	0.039	0.039	0.043	0.039	0.031
	(0.115)	(0.115)	(0.117)	(0.118)	(0.116)	(0.117)	(0.116)
Child's age(squared)	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Birth weight of the child(Kgs)	0.011	0.011	0.012	0.015	0.011	0.011	0.011
	(0.019)	(0.019)	(0.020)	(0.020)	(0.020)	(0.020)	(0.019)
Mother suffering from illness	0.018	0.020	0.016	0.020	0.017	0.014	0.017
	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)
English spoken to child at age3	0.131**	0.135**	0.134**	0.133**	0.127**	0.130**	0.129**
	(0.055)	(0.055)	(0.055)	(0.055)	(0.055)	(0.055)	(0.055)
Two-parent household at age 3	-0.015	-0.019	-0.011	-0.021	-0.012	-0.006	-0.005
	(0.041)	(0.040)	(0.041)	(0.039)	(0.040)	(0.041)	(0.041)
OECD inc. of hhld at age 3	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mother age at the birth of child	0.005**	0.005**	0.006**	0.006**	0.005**	0.006**	0.006**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Mother born in the UK	0.122***	0.121***	0.117**	0.121***	0.118**	0.117**	0.116**
	(0.047)	(0.047)	(0.047)	(0.046)	(0.047)	(0.047)	(0.047)
Nonwhite mother	-0.039	-0.039	-0.038	-0.041	-0.037	-0.034	-0.036
(Base=White)	(0.052)	(0.051)	(0.052)	(0.051)	(0.051)	(0.052)	(0.052)
Mother having	-0.124***	-0.122***	-0.129***	-0.130***	-0.119***	-0.133***	-0.128***
A/AS/GCSE/other qual.	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
Mother having	-0.232***	-0.224***	-0.246***	-0.244***	-0.223***	-0.257***	-0.253***
no academic qual.	(0.052)	(0.050)	(0.053)	(0.049)	(0.051)	(0.053)	(0.051)
Constant	-1.622	-1.733	-1.838	-1.979	-2.127	-1.827	-1.366
	(7.266)	(7.247)	(7.416)	(7.429)	(7.325)	(7.385)	(7.357)
R^2	0.420	0.421	0.419	0.422	0.422	0.419	0.420
N	5,274	5,274	5,274	5,274	5,274	5,274	5,274

Notes.- The robust standard errors are in parenthesis. Child's age in months is controlled. Outcome is difference in standardised Z vocabulary scores between sweeps 2 (age 3) to 5 (age 11) with mean 0 and standard deviation 1. The variable of interest is the movement of families between IMDs for sweeps 2 (age 3) and 3 (age 5). The analysis also controls for child's age in months at sweeps 5(age 11), current period IMD level. The lagged period controls used consist of respective lagged period (sweep 2) outcomes, sweep 2 IMD. Child and household level controls consist of child's and mother's ethnicity, number of siblings, OECD weighted household income, mother's highest educational qualification, language spoken in household, mother having long term illness, child having long term illness, whether mother is born in the UK, child birth-weight, whether single or two parent household, age of mother at birth of child.
*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

TABLE B.9: The change in child vocabulary scores between sweeps 2 (age 3) and 6 (age 14) for families moving/staying at sweep 3 (age 5).

	Overall & Indv. IMDs						
	All reason (1)	Inc. (2)	Hlth (3)	Edu. (4)	House (5)	Crime (6)	Env. (7)
OLS with controls							
Δ(IMD) (Base=Stayed at same IMD or different IMD)							
Change in IMD	0.006	0.018	-0.001	-0.004	0.000	0.015	0.004
	(0.021)	(0.016)	(0.022)	(0.010)	(0.020)	(0.015)	(0.018)
Δ IMD^2	0.001	-0.002	0.002	-0.001	0.001	-0.002	-0.001
	(0.004)	(0.003)	(0.004)	(0.002)	(0.003)	(0.002)	(0.003)
-ve Δ IMD	-0.033	-0.016	-0.055	0.007	0.001	0.034	-0.009
movement (Base=+ve change)	(0.120)	(0.048)	(0.049)	(0.042)	(0.045)	(0.042)	(0.045)
Δ IMD	-0.010	-0.057*	-0.065	-0.006	0.012	0.002	0.003
for lower IMD (M#low IMD)	(0.088)	(0.032)	(0.047)	(0.015)	(0.032)	(0.020)	(0.024)
Δ IMD^2	0.003	-0.004	-0.010	0.002	0.002	0.002	0.004
for lower IMD (M#M#low IMD)	(0.013)	(0.006)	(0.010)	(0.003)	(0.005)	(0.004)	(0.004)
Lagged z test scores (sweep2)	-0.882***	-0.878***	-0.878***	-0.878***	-0.878***	-0.878***	-0.878***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Age 3 IMD (sweep2)	0.004	0.001	0.004	-0.005	0.006	0.004	0.004
	(0.004)	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.004)
Current IMD (sweep6)	-0.004	-0.001	-0.003	0.001	-0.003	-0.006*	-0.007**
	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
Ethnicity of child	-0.005	-0.007	-0.007	0.002	-0.006	-0.001	0.002
	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)
Child suffering from	-0.016	-0.016	-0.017	-0.016	-0.016	-0.016	-0.016
long-term illness	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Having sibling at age 3	0.000	0.002	0.002	0.001	0.001	0.002	0.002
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Child's age(years)	0.041	0.042	0.042	0.042	0.041	0.038	0.038
	(0.053)	(0.053)	(0.053)	(0.053)	(0.053)	(0.053)	(0.053)
Child's age(squared)	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Birth weight of the child(Kgs)	0.008	0.006	0.006	0.005	0.006	0.006	0.005
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Mother suffering from illness	0.039***	0.039***	0.039***	0.038***	0.038***	0.039***	0.037***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
English spoken to child at age3	0.059**	0.048*	0.048*	0.049*	0.049*	0.049*	0.049*
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
Two-parent household at age 3	0.014	0.015	0.015	0.015	0.014	0.014	0.014
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
OECD inc. of hhld at age 3	0.000***	0.000***	0.000***	0.000***	0.000**	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mother age at the birth of child	0.003*	0.003*	0.002*	0.002*	0.002*	0.003**	0.003**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Mother born in the UK	0.065***	0.065***	0.065***	0.064**	0.066***	0.067***	0.067***
	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)
Mother's ethnicity	-0.013	-0.016	-0.016	-0.013	-0.013	-0.015	-0.012
	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)
Mother having	-0.049***	-0.048***	-0.048***	-0.048***	-0.045***	-0.052***	-0.050***
A/AS/GCSE/other qual.	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Mother having	-0.046*	-0.045*	-0.044*	-0.047*	-0.042	-0.053**	-0.051**
no academic qual.	(0.025)	(0.025)	(0.025)	(0.025)	(0.026)	(0.025)	(0.025)
Constant	-3.782	-3.868	-3.847	-3.880	-3.776	-3.539	-3.513
	(4.321)	(4.342)	(4.337)	(4.338)	(4.332)	(4.332)	(4.332)
R^2	0.842	0.842	0.842	0.842	0.842	0.842	0.842
N	4,433	4,433	4,433	4,433	4,433	4,433	4,433

Notes.- The robust standard errors are in parenthesis. Child's age in months is controlled. Outcome is difference in standardised Z vocabulary scores between sweeps 2 (age 3) to 6 (age 14) with mean 0 and standard deviation 1. The variable of interest is the movement of families between IMDs for sweeps 2 (age 3) and 3 (age 5). The analysis also controls for child's age in months at sweeps 6(age 14), current period IMD level. The lagged period controls used consist of respective lagged period (sweep 2) outcomes, sweep 5 IMD. Child and household level controls consist of child's and mother's ethnicity, number of siblings, OECD weighted household income, mother's highest educational qualification, language spoken in household, mother having long term illness, child having long term illness, whether mother is born in the UK, child birth-weight, whether single or two parent household, age of mother at birth of child.
*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

TABLE B.10: The change in child internalising(total difficulty) scores between sweeps 2 (age 3) and 4 (age 7) for families moving/staying at sweep 3 (age 5).

	Overall & Indv. IMDs						
	All reason (1)	Inc. (2)	Hlth. (3)	Edu. (4)	House (5)	Crime (6)	Env. (7)
OLS with controls							
Δ(IMD) (Base=Stayed at same IMD or different IMD)							
Change in IMD	0.062	0.062	-0.085*	0.012	-0.035	-0.014	0.005
	(0.047)	(0.044)	(0.049)	(0.022)	(0.041)	(0.033)	(0.024)
$\Delta\ IMD^2$	-0.008	-0.009	0.015*	-0.007**	0.002	-0.000	-0.006*
	(0.008)	(0.008)	(0.009)	(0.003)	(0.008)	(0.005)	(0.003)
-ve Δ IMD	0.201	0.050	0.018	-0.135	0.026	0.079	-0.016
movement (Base=+ve change)	(0.221)	(0.108)	(0.096)	(0.095)	(0.090)	(0.090)	(0.105)
Δ IMD	0.119	0.065	0.054	-0.082***	0.081	0.057	0.007
for lower IMD (M#low IMD)	(0.148)	(0.064)	(0.084)	(0.027)	(0.062)	(0.044)	(0.042)
$\Delta\ IMD^2$	0.025	0.023**	-0.032**	0.010**	0.002	-0.008	0.004
for lower IMD (M#M#low IMD)	(0.020)	(0.011)	(0.013)	(0.004)	(0.012)	(0.007)	(0.009)
Lagged z test scores (sweep2)	-0.467***	-0.469***	-0.068***	-0.068***	-0.068***	-0.070***	-0.069***
	(0.016)	(0.016)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Age 3 IMD (sweep2)	0.024	0.016	-0.009	0.004	-0.020	-0.003	-0.001
	(0.015)	(0.015)	(0.018)	(0.012)	(0.015)	(0.014)	(0.013)
Current IMD (sweep4)	-0.031**	-0.028*	0.008	-0.004	0.020	-0.008	-0.008
	(0.015)	(0.014)	(0.018)	(0.011)	(0.015)	(0.014)	(0.013)
Nonwhite child	0.053	0.065	-0.223***	-0.229***	-0.225***	-0.218***	-0.214***
(Base=White)	(0.073)	(0.073)	(0.064)	(0.064)	(0.064)	(0.064)	(0.064)
Child suffering from	0.060	0.061	0.019	0.018	0.019	0.021	0.021
long-term illness	(0.039)	(0.039)	(0.037)	(0.037)	(0.037)	(0.037)	(0.037)
Having sibling at age 3	-0.119***	-0.119***	0.041	0.043	0.043	0.042	0.044
	(0.033)	(0.033)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)
Child's age(years)	-0.230	-0.234	-0.249	-0.252	-0.252	-0.250	-0.234
	(0.210)	(0.210)	(0.224)	(0.222)	(0.222)	(0.222)	(0.223)
Child's age(squared)	0.001	0.001	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Birth weight of the child(Kgs)	-0.030	-0.029	0.032	0.034	0.033	0.033	0.033
	(0.020)	(0.020)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Mother suffering from illness	0.105***	0.105***	-0.004	-0.001	-0.003	-0.003	-0.002
	(0.028)	(0.028)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)
English spoken to child at age3	-0.062	-0.056	0.495***	0.495***	0.495***	0.491***	0.491***
	(0.066)	(0.066)	(0.067)	(0.067)	(0.068)	(0.067)	(0.067)
Two-parent household at age 3	-0.120***	-0.118**	-0.047	-0.049	-0.047	-0.035	-0.041
	(0.046)	(0.046)	(0.045)	(0.045)	(0.045)	(0.045)	(0.045)
OECD inc. of hhld at age 3	-0.000***	-0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mother age at the birth of child	-0.004	-0.004	0.000	-0.001	0.000	0.001	-0.000
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Mother born in the UK	0.038	0.036	0.252***	0.249***	0.248***	0.248***	0.250***
	(0.054)	(0.054)	(0.057)	(0.057)	(0.057)	(0.057)	(0.057)
Nonwhite mother	-0.019	-0.015	-0.063	-0.059	-0.063	-0.060	-0.056
(Base=White)	(0.062)	(0.062)	(0.055)	(0.055)	(0.055)	(0.055)	(0.055)
Having sibling at age 3	-0.119***	-0.119***	0.041	0.043	0.043	0.042	0.044
	(0.033)	(0.033)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)
Mother having	0.022	0.018	-0.062*	-0.064**	-0.063*	-0.070**	-0.064**
A/AS/GCSE/other qual.	(0.029)	(0.029)	(0.033)	(0.032)	(0.033)	(0.032)	(0.032)
Mother having	0.124**	0.114**	-0.119**	-0.118**	-0.120**	-0.134**	-0.127**
no academic qual.	(0.055)	(0.055)	(0.057)	(0.055)	(0.057)	(0.056)	(0.056)
Constant	10.955	11.129	11.789	11.948	11.945	11.876	11.194
	(9.101)	(9.096)	(9.706)	(9.639)	(9.647)	(9.638)	(9.666)
R^2	0.186	0.187	0.085	0.088	0.085	0.086	0.085
N	6,058	6,058	6,058	6,058	6,058	6,058	6,058

Notes.- The robust standard errors are in parenthesis. Child's age in months is controlled. Outcome is difference in standardised Z internalising scores between sweeps 2 (age 3) to 4 (age 7) with mean 0 and standard deviation 1. The variable of interest is the movement of families between IMDs for sweeps 2 (age 3) and 3 (age 5). The analysis also controls for child's age in months at sweeps 4(age 7), current period IMD level. The lagged period controls used consist of respective lagged period (sweep 2) outcomes, sweep 6 IMD. Child and household level controls consist of child's and mother's ethnicity, number of siblings, OECD weighted household income, mother's highest educational qualification, language spoken in household, mother having long term illness, child having long term illness, whether mother is born in the UK, child birth-weight, whether single or two parent household, age of mother at birth of child.
*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

TABLE B.11: The change in child internalising(total difficulty) scores between sweeps 2 (age 3) and 5 (age 11) for families moving/staying at sweep 3 (age 5).

	Overall & Indv. IMDs						
	All reason (1)	Inc. (2)	Hlth (3)	Edu. (4)	House (5)	Crime (6)	Env. (7)
OLS with controls							
Δ(IMD) (Base=Stayed at same IMD or different IMD)							
Change in IMD	0.016	0.054	-0.070	0.003	0.002	-0.019	-0.076**
	(0.068)	(0.052)	(0.071)	(0.038)	(0.042)	(0.031)	(0.030)
Δ IMD^2	0.000	-0.020*	0.005	-0.009	-0.013	-0.000	0.011**
	(0.009)	(0.011)	(0.013)	(0.007)	(0.009)	(0.004)	(0.005)
-ve Δ IMD	0.238	0.041	0.036	-0.039	0.074	0.047	0.088
movement(Base=+ve change)	(0.265)	(0.095)	(0.094)	(0.091)	(0.083)	(0.090)	(0.097)
Δ IMD	0.130	-0.020	0.012	-0.015	0.014	0.027	0.099**
for lower IMD (M#low IMD)	(0.189)	(0.066)	(0.098)	(0.043)	(0.057)	(0.039)	(0.040)
Δ IMD^2	0.016	0.025*	-0.024	0.015*	0.009	-0.001	-0.013
for lower IMD (M#M#low IMD)	(0.025)	(0.013)	(0.019)	(0.008)	(0.013)	(0.007)	(0.008)
Lagged z test scores (sweep2)	-0.491***	0.050***	0.049***	0.050***	0.048***	0.049***	0.047***
	(0.018)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Age 3 IMD (sweep2)	-0.018	0.004	0.006	-0.006	0.001	0.006	-0.004
	(0.013)	(0.011)	(0.012)	(0.009)	(0.010)	(0.010)	(0.010)
Current IMD (sweep5)	0.005	-0.011	-0.016	0.019**	-0.012	-0.021**	-0.015
	(0.012)	(0.011)	(0.012)	(0.009)	(0.010)	(0.009)	(0.009)
Ethnicity of child	0.137	-0.168**	-0.167**	-0.182**	-0.171**	-0.161**	-0.153**
	(0.086)	(0.078)	(0.077)	(0.077)	(0.077)	(0.078)	(0.077)
Child suffering from	0.173***	0.003	0.000	-0.004	-0.002	0.001	0.003
long-term illness	(0.046]	(0.043]	(0.043]	(0.042]	(0.042]	(0.043]	(0.042]
Having sibling at age 3	-0.136***	0.167***	0.167***	0.167***	0.166***	0.165***	0.168***
	(0.039)	(0.038)	(0.038)	(0.037)	(0.037)	(0.038)	(0.038)
Child's age(years)	0.013	0.072	0.067	0.069	0.078	0.063	0.054
	(0.188)	(0.170)	(0.172)	(0.172)	(0.173)	(0.171)	(0.171)
Child's age(squared)	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Birth weight of the child(Kgs)	-0.005	-0.043*	-0.044*	-0.041*	-0.044*	-0.044*	-0.045**
	(0.024)	(0.022)	(0.022)	(0.023)	(0.022)	(0.022)	(0.022)
Mother suffering from illness	0.182***	0.044	0.040	0.047	0.043	0.041	0.043
	(0.033)	(0.031)	(0.032)	(0.031)	(0.031)	(0.032)	(0.032)
English spoken to child at age3	-0.129*	0.490***	0.486***	0.487***	0.487***	0.483***	0.479***
	(0.077)	(0.072)	(0.071)	(0.071)	(0.072)	(0.071)	(0.071)
Two-parent household at age 3	-0.145**	-0.134***	-0.134***	-0.144***	-0.138***	-0.128**	-0.126**
	(0.058)	(0.049)	(0.049)	(0.048)	(0.048)	(0.050)	(0.049)
OECD inc. of hhld at age 3	-0.000***	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mother age at the birth of child	0.002	0.000	0.000	0.000	0.001	0.001	0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Mother born in the UK	-0.019	0.199***	0.198***	0.200***	0.202***	0.198***	0.197***
	(0.056)	(0.060)	(0.060)	(0.060)	(0.060)	(0.060)	(0.061)
Mother's ethnicity	-0.060	-0.127**	-0.131**	-0.134**	-0.131**	-0.124*	-0.125**
	(0.076)	(0.063)	(0.064)	(0.064)	(0.064)	(0.064)	(0.064)
Mother having	0.026	-0.006	-0.008	-0.003	-0.015	-0.012	-0.002
A/AS/GCSE/other qual.	(0.032)	(0.035)	(0.035)	(0.035)	(0.036)	(0.035)	(0.035)
Mother having	0.094	0.026	0.020	0.037	0.017	0.012	0.019
no academic qual.	(0.070)	(0.063)	(0.065)	(0.062)	(0.064)	(0.065)	(0.064)
Constant	0.045	-3.704	-3.320	-3.611	-4.049	-3.077	-2.534
	(11.835)	(10.712)	(10.885)	(10.855)	(10.906)	(10.798)	(10.765)
R^2	0.200	0.076	0.075	0.077	0.077	0.075	0.077
N	5,274	5,274	5,274	5,274	5,274	5,274	5,274

Notes.- The robust standard errors are in parenthesis. Child's age in months is controlled. Outcome is difference in standardised Z vocabulary scores between sweeps 2 (age 3) to 5 (age 11) with mean 0 and standard deviation 1. The variable of interest is the movement of families between IMDs for sweeps 2 (age 3) and 3 (age 5). The analysis also controls for child's age in months at sweeps 5(age 11), current period IMD level. The lagged period controls used consist of respective lagged period (sweep 2) outcomes, sweep 5 IMD. Child and household level controls consist of child's and mother's ethnicity, number of siblings, OECD weighted household income, mother's highest educational qualification, language spoken in household, mother having long term illness, child having long term illness, whether mother is born in the UK, child birth-weight, whether single or two parent household, age of mother at birth of child.
*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

TABLE B.12: The change in child internalising(total difficulty) scores between sweeps 2 (age 3) and 6 (age 14) for families moving/staying at sweep 3 (age 5).

	Overall & Indv. IMDs						
	All reason (1)	Inc. (2)	Hlth (3)	Edu. (4)	House (5)	Crime (6)	Env. (7)
OLS with controls							
Δ(IMD) (Base=Stayed at same IMD or different IMD)							
Change in IMD	0.066	0.026	-0.026	-0.015	-0.045	-0.013	-0.019
	(0.054)	(0.041)	(0.050)	(0.022)	(0.042)	(0.036)	(0.035)
$\Delta\ IMD^2$	-0.013	-0.004	0.003	-0.000	0.006	0.001	0.004
	(0.009)	(0.007)	(0.009)	(0.004)	(0.008)	(0.006)	(0.007)
-ve Δ IMD	0.127	-0.095	-0.101	-0.098	-0.010	0.060	-0.044
movement (Base=+ve change)	(0.234)	(0.241)	(0.099)	(0.094)	(0.091)	(0.090)	(0.093)
Δ IMD	0.015	-0.064	-0.133*	0.010	0.124*	0.055	0.007
for lower IMD (M#low IMD)	(0.176)	(0.055)	(0.077)	(0.031)	(0.064)	(0.045)	(0.048)
$\Delta\ IMD^2$	0.029	0.002	-0.044***	0.007	0.004	-0.006	-0.006
for lower IMD (M#M#low IMD)	(0.025)	(0.010)	(0.015)	(0.006)	(0.011)	(0.008)	(0.010)
Lagged z test scores (sweep2)	-0.480***	0.117***	0.118***	0.120***	0.116***	0.118***	0.118***
	(0.017)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Age 3 IMD (sweep2)	-0.004	0.001	0.019*	-0.004	0.007	0.015	0.007
	(0.010)	(0.009)	(0.011)	(0.009)	(0.010)	(0.009)	(0.009)
Current IMD (sweep6)	-0.015	-0.017*	-0.027**	-0.001	-0.023**	-0.022**	-0.020**
	(0.010)	(0.009)	(0.011)	(0.009)	(0.009)	(0.009)	(0.008)
Ethnicity of child	0.074	-0.133*	-0.147**	-0.143**	-0.138**	-0.138**	-0.130*
	(0.080)	(0.070)	(0.070)	(0.070)	(0.070)	(0.070)	(0.070)
Child suffering from	0.152***	-0.066	-0.068*	-0.065	-0.068*	-0.066	-0.065
long-term illness	(0.043)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)
Having sibling at age 3	-0.045	0.222***	0.223***	0.220***	0.224***	0.218***	0.220***
	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)
Child's age(years)	-0.154	-0.006	0.002	0.005	-0.004	-0.006	-0.012
	(0.193)	(0.181)	(0.181)	(0.181)	(0.181)	(0.181)	(0.181)
Child's age(squared)	0.000	0.000	0.000	-0.000	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Birth weight of the child(Kgs)	-0.052**	-0.069***	-0.070***	-0.071***	-0.071***	-0.073***	-0.072***
	(0.024)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)
Mother suffering from illness	0.105***	0.088***	0.088***	0.092***	0.089***	0.092***	0.087***
	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)
English spoken to child at age3	-0.070	0.534***	0.533***	0.537***	0.544***	0.535***	0.536***
	(0.067)	(0.063)	(0.062)	(0.063)	(0.062)	(0.062)	(0.062)
Two-parent household at age 3	-0.133**	-0.134***	-0.142***	-0.141***	-0.148***	-0.137***	-0.142***
	(0.053)	(0.046)	(0.046)	(0.046)	(0.046)	(0.046)	(0.046)
OECD inc. of hhld at age 3	-0.000***	-0.000*	-0.000**	-0.000***	-0.000	-0.000**	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mother age at the birth of child	-0.001	-0.005*	-0.006**	-0.006**	-0.005*	-0.006*	-0.005*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Mother born in the UK	0.005	0.232***	0.238***	0.232***	0.242***	0.237***	0.236***
	(0.055)	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)
Mother's ethnicity	0.007	-0.133**	-0.138**	-0.139**	-0.136**	-0.141**	-0.133**
	(0.067)	(0.061)	(0.061)	(0.061)	(0.061)	(0.061)	(0.061)
Mother having	0.034	0.127***	0.133***	0.140***	0.120***	0.131***	0.136***
A/AS/GCSE/other qual.	(0.032)	(0.035)	(0.035)	(0.035)	(0.036)	(0.035)	(0.035)
Mother having	0.132**	0.378***	0.391***	0.408***	0.371***	0.390***	0.395***
no academic qual.	(0.059)	(0.058)	(0.058)	(0.057)	(0.059)	(0.058)	(0.057)
Constant	13.375	0.533	-0.100	-0.354	0.454	0.543	1.075
	(15.647)	(14.692)	(14.730)	(14.666)	(14.745)	(14.759)	(14.667)
R^2	0.203	0.189	0.190	0.188	0.190	0.189	0.189
N	4,433	4,433	4,433	4,433	4,433	4,433	4,433

Notes.- The robust standard errors are in parenthesis. Child's age in months is controlled. Outcome is difference in standardised Z vocabulary scores between sweeps 2 (age 3) to 6 (age 14) with mean 0 and standard deviation 1. The variable of interest is the movement of families between IMDs for sweeps 2 (age 3) and 3 (age 5). The analysis also controls for child's age in months at sweeps 6(age 14), current period IMD level. The lagged period controls used consist of respective lagged period (sweep 2) outcomes, sweep 6 IMD. Child and household level controls consist of child's and mother's ethnicity, number of siblings, OECD weighted household income, mother's highest educational qualification, language spoken in household, mother having long term illness, child having long term illness, whether mother is born in the UK, child birth-weight, whether single or two parent household, age of mother at birth of child.
*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

TABLE B.13: The change in child externalising(prosocial) scores between sweeps 2 (age 3) and 4 (age 7) for families moving/staying at sweep 3 (age 5).

	Overall & Indv. IMDs						
	All reason (1)	Inc. (2)	Hlth. (3)	Edu. (4)	House (5)	Crime (6)	Env. (7)
OLS with controls							
Δ(IMD) (Base=Stayed at same IMD or different IMD)							
Change in IMD	-0.050	-0.066*	-0.089*	0.012	-0.038	-0.015	0.005
	(0.039)	(0.035)	(0.049)	(0.021)	(0.041)	(0.034)	(0.024)
$\Delta\ IMD^2$	0.007	0.009	0.016*	-0.007**	0.003	-0.000	-0.007**
	(0.006)	(0.006)	(0.009)	(0.003)	(0.008)	(0.005)	(0.003)
-ve Δ IMD	0.090	0.036	0.016	-0.143	0.022	0.076	-0.017
movement (Base=+ve change)	(0.202)	(0.088)	(0.098)	(0.095)	(0.091)	(0.090)	(0.106)
Δ IMD	0.083	0.042	0.064	-0.081***	0.085	0.057	0.004
for lower IMD (M#low IMD)	(0.141)	(0.047)	(0.086)	(0.027)	(0.062)	(0.045)	(0.041)
$\Delta\ IMD^2$	-0.002	-0.013	-0.032**	0.010**	0.002	-0.008	0.004
for lower IMD (M#M#low IMD)	(0.021)	(0.009)	(0.013)	(0.004)	(0.012)	(0.008)	(0.009)
Lagged z test scores (sweep2)	-0.656***	-0.656***	0.001	-0.001	0.000	0.001	0.000
	(0.014)	(0.014)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Age 3 IMD (sweep2)	-0.004	-0.007	-0.008	0.006	-0.020	-0.003	-0.002
	(0.012)	(0.012)	(0.019)	(0.012)	(0.015)	(0.014)	(0.013)
Current IMD (sweep4)	0.009	0.011	0.008	-0.006	0.022	-0.006	-0.006
	(0.012)	(0.012)	(0.019)	(0.011)	(0.015)	(0.014)	(0.013)
Nonwhite child	0.023	0.021	-0.226***	-0.232***	-0.230***	-0.222***	-0.219***
(Base=White)	(0.062)	(0.062)	(0.063)	(0.063)	(0.063)	(0.063)	(0.063)
Child suffering from	0.006	0.006	0.009	0.009	0.010	0.012	0.011
long-term illness	(0.032)	(0.032)	(0.037)	(0.037)	(0.037)	(0.037)	(0.037)
Having sibling at age 3	0.050*	0.051*	0.041	0.043	0.042	0.042	0.044
	(0.029)	(0.028)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)
Child's age(years)	0.140	0.143	-0.235	-0.235	-0.238	-0.235	-0.222
	(0.179)	(0.178)	(0.225)	(0.223)	(0.223)	(0.223)	(0.224)
Child's age(squared)	-0.001	-0.001	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Birth weight of the child(Kgs)	0.009	0.009	0.038	0.039*	0.038	0.038	0.039*
	(0.018)	(0.018)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Mother suffering from illness	-0.009	-0.009	-0.015	-0.012	-0.014	-0.015	-0.014
	(0.024)	(0.024)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)
English spoken to child at age3	0.048	0.046	0.491***	0.490***	0.490***	0.486***	0.486***
	(0.056)	(0.056)	(0.068)	(0.067)	(0.068)	(0.067)	(0.068)
Two-parent household at age 3	0.102***	0.103***	-0.040	-0.041	-0.039	-0.027	-0.034
	(0.037)	(0.037)	(0.045)	(0.045)	(0.045)	(0.045)	(0.045)
OECD inc. of hhld at age 3	0.000	0.000	0.000***	0.000***	0.000***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mother age at the birth of child	0.000	0.000	0.001	0.001	0.002	0.002	0.001
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Mother born in the UK	0.013	0.013	0.256***	0.253***	0.252***	0.253***	0.254***
	(0.046)	(0.046)	(0.057)	(0.057)	(0.057)	(0.057)	(0.057)
Nonwhite mother	0.022	0.022	-0.053	-0.049	-0.052	-0.050	-0.046
(Base=White)	(0.050)	(0.050)	(0.055)	(0.055)	(0.055)	(0.055)	(0.055)
Mother having	0.021	0.021	-0.076**	-0.078**	-0.075**	-0.084***	-0.079**
A/AS/GCSE/other qual.	(0.027)	(0.027)	(0.032)	(0.032)	(0.033)	(0.032)	(0.032)
Mother having	-0.074	-0.075	-0.156***	-0.158***	-0.156***	-0.173***	-0.167***
no academic qual.	(0.046)	(0.046)	(0.056)	(0.055)	(0.057)	(0.056)	(0.055)
Constant	-6.933	-7.055	11.119	11.164	11.254	11.120	10.585
	(7.762)	(7.746)	(9.755)	(9.681)	(9.690)	(9.688)	(9.718)
R^2	0.363	0.363	0.082	0.084	0.081	0.082	0.082
N	6,058	6,058	6,058	6,058	6,058	6,058	6,058

Notes.- The robust standard errors are in parenthesis. Child's age in months is controlled. Outcome is difference in standardised Z vocabulary scores between sweeps 2 (age 3) to 4 (age 7) with mean 0 and standard deviation 1. The variable of interest is the movement of families between IMDs for sweeps 2 (age 3) and 3 (age 5). The analysis also controls for child's age in months at sweeps 4(age 7), current period IMD level. The lagged period controls used consist of respective lagged period (sweep 2) outcomes, sweep 4 IMD. Child and household level controls consist of child's and mother's ethnicity, number of siblings, OECD weighted household income, mother's highest educational qualification, language spoken in household, mother having long term illness, child having long term illness, whether mother is born in the UK, child birth-weight, whether single or two parent household, age of mother at birth of child.
*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

TABLE B.14: The change in child externalising(prosocial) scores between sweeps 2 (age 3) and 5 (age 11) for families moving/staying at sweep 3 (age 5).

	Overall & Indv. IMDs						
	All reason (1)	Inc. (2)	Hlth (3)	Edu. (4)	House (5)	Crime (6)	Env. (7)
OLS with controls							
Δ(IMD) (Base=Stayed at same IMD or different IMD)							
Change in IMD	-0.079	0.050	-0.072	0.004	0.001	-0.017	-0.078**
	(0.063)	(0.051)	(0.072)	(0.039)	(0.042)	(0.031)	(0.030)
$\Delta\ IMD^2$	0.014	-0.019*	0.006	-0.009	-0.013	-0.000	0.011**
	(0.009)	(0.011)	(0.013)	(0.007)	(0.010)	(0.004)	(0.005)
-ve Δ IMD	0.199	0.049	0.041	-0.034	0.072	0.050	0.095
movement (Base=+ve change)	(0.192)	(0.096)	(0.095)	(0.092)	(0.084)	(0.091)	(0.097)
Δ IMD	0.170	-0.013	0.015	-0.017	0.010	0.029	0.102**
for lower IMD (M#low IMD)	(0.140)	(0.067)	(0.100)	(0.045)	(0.058)	(0.039)	(0.041)
$\Delta\ IMD^2$	-0.001	0.025*	-0.025	0.015*	0.009	-0.000	-0.014*
for lower IMD (M#M#low IMD)	(0.020)	(0.014)	(0.019)	(0.008)	(0.013)	(0.007)	(0.008)
Lagged z test scores (sweep2)	-0.721***	-0.053***	-0.052***	-0.052***	-0.052***	-0.051***	-0.053***
	(0.015)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Age 3 IMD (sweep2)	0.017*	0.002	0.005	-0.006	-0.000	0.005	-0.004
	(0.010)	(0.011)	(0.012)	(0.009)	(0.010)	(0.010)	(0.010)
Current IMD (sweep5)	-0.009	-0.011	-0.016	0.020**	-0.012	-0.022**	-0.015
	(0.009)	(0.011)	(0.012)	(0.009)	(0.010)	(0.009)	(0.009)
Ethnicity of child	-0.016	-0.166**	-0.167**	-0.183**	-0.171**	-0.161**	-0.153**
	(0.066)	(0.078)	(0.077)	(0.077)	(0.078)	(0.078)	(0.077)
Child suffering from	-0.024	0.008	0.006	0.001	0.003	0.007	0.008
long-term illness	(0.037)	(0.043)	(0.043)	(0.042)	(0.042)	(0.043)	(0.042)
Having sibling at age 3	0.099***	0.157***	0.158***	0.158***	0.156***	0.156***	0.159***
	(0.033)	(0.038)	(0.038)	(0.037)	(0.038)	(0.038)	(0.038)
Child's age(years)	0.192	0.045	0.039	0.041	0.052	0.036	0.027
	(0.198)	(0.174)	(0.177)	(0.177)	(0.177)	(0.176)	(0.175)
Child's age(squared)	-0.001	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Birth weight of the child(Kgs)	0.002	-0.046**	-0.047**	-0.044**	-0.047**	-0.047**	-0.048**
	(0.019)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Mother suffering from illness	0.018	0.050	0.046	0.053*	0.049	0.047	0.049
	(0.027)	(0.031)	(0.032)	(0.031)	(0.031)	(0.032)	(0.031)
English spoken to child at age3	0.083	0.489***	0.485***	0.486***	0.487***	0.482***	0.478***
	(0.066)	(0.072)	(0.072)	(0.072)	(0.073)	(0.072)	(0.072)
Two-parent household at age 3	0.109**	-0.137***	-0.138***	-0.149***	-0.142***	-0.132***	-0.130***
	(0.046)	(0.049)	(0.049)	(0.048)	(0.048)	(0.050)	(0.049)
OECD inc. of hhld at age 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mother age at the birth of child	0.004	-0.001	-0.001	-0.001	-0.001	-0.000	-0.000
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Mother born in the UK	-0.059	0.207***	0.205***	0.208***	0.210***	0.205***	0.205***
	(0.052)	(0.061)	(0.060)	(0.060)	(0.061)	(0.061)	(0.061)
Mother's ethnicity	0.085	-0.136**	-0.140**	-0.144**	-0.141**	-0.133**	-0.134**
	(0.060)	(0.063)	(0.063)	(0.063)	(0.063)	(0.064)	(0.063)
Having sibling at age 3	0.099***	0.157***	0.158***	0.158***	0.156***	0.156***	0.159***
	(0.033)	(0.038)	(0.038)	(0.037)	(0.038)	(0.038)	(0.038)
Mother having	0.017	0.001	-0.001	0.005	-0.009	-0.005	0.005
A/AS/GCSE/other qual.	(0.029)	(0.035)	(0.035)	(0.035)	(0.036)	(0.035)	(0.035)
Mother having	-0.067	0.042	0.037	0.057	0.032	0.030	0.037
no academic qual.	(0.058)	(0.063)	(0.065)	(0.061)	(0.063)	(0.065)	(0.063)
Constant	-12.901	-1.891	-1.507	-1.756	-2.308	-1.310	-0.747
	(12.491)	(11.007)	(11.165)	(11.164)	(11.187)	(11.089)	(11.039)
R^2	0.417	0.076	0.076	0.077	0.077	0.076	0.077
N	5,274	5,274	5,274	5,274	5,274	5,274	5,274

Notes.- The robust standard errors are in parenthesis. Child's age in months is controlled. Outcome is difference in standardised Z vocabulary scores between sweeps 2 (age 3) to 5 (age 11) with mean 0 and standard deviation 1. The variable of interest is the movement of families between IMDs for sweeps 2 (age 3) and 3 (age 5). The analysis also controls for child's age in months at sweeps 5(age 11), current period IMD level. The lagged period controls used consist of respective lagged period (sweep 2) outcomes, sweep 5 IMD. Child and household level controls consist of child's and mother's ethnicity, number of siblings, OECD weighted household income, mother's highest educational qualification, language spoken in household, mother having long term illness, child having long term illness, whether mother is born in the UK, child birth-weight, whether single or two parent household, age of mother at birth of child.
*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

TABLE B.15: The change in child externalising(prosocial) scores between sweeps 2 (age 3) and 6 (age 14) for families moving/staying at sweep 3 (age 5).

	Overall & Indv. IMDs						
	All reason (1)	Inc. (2)	Hlth (3)	Edu. (4)	House (5)	Crime (6)	Env. (7)
OLS with controls							
Δ(IMD) (Base=Stayed at same IMD or different IMD)							
Change in IMD	-0.047	0.021	-0.023	-0.015	-0.051	-0.010	-0.017
	(0.040)	(0.040)	(0.049)	(0.022)	(0.042)	(0.035)	(0.035)
$\Delta\ IMD^2$	0.011	-0.004	0.003	-0.001	0.007	0.001	0.004
	(0.007)	(0.007)	(0.009)	(0.004)	(0.008)	(0.005)	(0.007)
-ve Δ IMD	-0.086	-0.067	-0.075	-0.080	0.001	0.080	-0.020
movement (Base=+ve change)	(0.212)	(0.098)	(0.093)	(0.092)	(0.091)	(0.092)	(0.097)
Δ IMD	-0.097	-0.047	-0.121	0.009	0.127*	0.060	0.007
for lower IMD (M#low IMD)	(0.143)	(0.053)	(0.076)	(0.031)	(0.066)	(0.044)	(0.048)
$\Delta\ IMD^2$	-0.041**	0.002	-0.042***	0.007	0.003	-0.006	-0.007
for lower IMD (M#M#low IMD)	(0.020)	(0.010)	(0.015)	(0.006)	(0.011)	(0.008)	(0.010)
Lagged z test scores (sweep2)	-0.578***	-0.098***	-0.098***	-0.098***	-0.098***	-0.098***	-0.098***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Age 3 IMD (sweep2)	0.006	-0.001	0.020*	-0.005	0.005	0.014	0.007
	(0.009)	(0.009)	(0.011)	(0.009)	(0.010)	(0.009)	(0.009)
Current IMD (sweep6)	-0.003	-0.019**	-0.030***	-0.001	-0.026***	-0.024***	-0.022**
	(0.008)	(0.009)	(0.010)	(0.009)	(0.009)	(0.009)	(0.008)
Ethnicity of child	0.053	-0.132*	-0.150**	-0.147**	-0.141**	-0.140**	-0.131*
	(0.064)	(0.070)	(0.070)	(0.070)	(0.069)	(0.070)	(0.070)
Child suffering from	0.010	-0.052	-0.054	-0.050	-0.055	-0.051	-0.051
long-term illness	(0.037)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)
Having sibling at age 3	0.102***	0.207***	0.208***	0.205***	0.209***	0.203***	0.205***
	(0.032)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)
Child's age(years)	0.280*	-0.013	-0.005	-0.001	-0.012	-0.012	-0.020
	(0.162)	(0.183)	(0.184)	(0.183)	(0.184)	(0.184)	(0.183)
Child's age(squared)	-0.001*	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Birth weight of the child(Kgs)	0.005	-0.080***	-0.081***	-0.082***	-0.082***	-0.083***	-0.083***
	(0.021)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)
Mother suffering from illness	0.016	0.100***	0.101***	0.105***	0.101***	0.104***	0.099***
	(0.027)	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)
English spoken to child at age3	0.080	0.530***	0.529***	0.533***	0.541***	0.532***	0.532***
	(0.063)	(0.063)	(0.063)	(0.063)	(0.063)	(0.063)	(0.063)
Two-parent household at age 3	0.082*	-0.143***	-0.152***	-0.150***	-0.158***	-0.146***	-0.152***
	(0.044)	(0.046)	(0.046)	(0.046)	(0.046)	(0.046)	(0.046)
OECD inc. of hhld at age 3	0.000	-0.000**	-0.000**	-0.000***	-0.000*	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mother age at the birth of child	0.007***	-0.008***	-0.008***	-0.009***	-0.008***	-0.008***	-0.008***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Mother born in the UK	0.009	0.249***	0.255***	0.249***	0.260***	0.253***	0.252***
	(0.051)	(0.055)	(0.055)	(0.055)	(0.055)	(0.055)	(0.055)
Mother's ethnicity	0.005	-0.159***	-0.165***	-0.166***	-0.162***	-0.167***	-0.159***
	(0.057)	(0.060)	(0.060)	(0.060)	(0.060)	(0.060)	(0.060)
Having sibling at age 3							
Mother having	-0.017	0.146***	0.154***	0.163***	0.137***	0.151***	0.158***
A/AS/GCSE/other qual.	(0.030)	(0.036)	(0.036)	(0.035)	(0.036)	(0.036)	(0.035)
Mother having	-0.039	0.413***	0.429***	0.451***	0.405***	0.428***	0.435***
no academic qual.	(0.051)	(0.058)	(0.057)	(0.056)	(0.058)	(0.057)	(0.056)
Constant	-23.248*	1.364	0.689	0.300	1.239	1.276	1.898
	(13.120)	(14.876)	(14.928)	(14.858)	(14.931)	(14.950)	(14.861)
R^2	0.316	0.185	0.186	0.184	0.186	0.185	0.185
N	4,433	4,433	4,433	4,433	4,433	4,433	4,433

Notes.- The robust standard errors are in parenthesis. Child's age in months is controlled. Outcome is difference in standardised Z vocabulary scores between sweeps 2 (age 3) to 6 (age 14) with mean 0 and standard deviation 1. The variable of interest is the movement of families between IMDs for sweeps 2 (age 3) and 3 (age 5). The analysis also controls for child's age in months at sweeps 6(age 14), current period IMD level. The lagged period controls used consist of respective lagged period (sweep 2) outcomes, sweep 6 IMD. Child and household level controls consist of child's and mother's ethnicity, number of siblings, OECD weighted household income, mother's highest educational qualification, language spoken in household, mother having long term illness, child having long term illness, whether mother is born in the UK, child birth-weight, whether single or two parent household, age of mother at birth of child.
*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Appendix C

The Impact of Covid on Youth's Preferences to Continue Education

TABLE C.1: Marginal impact of maternal occupations by Government office region

Maternal SOC rank by GOR	NE (1)	NW (2)	Y& H (3)	EM (4)	WM (5)	E (6)	L (7)	SE (8)	SW (9)	WL (10)	SC (11)
Corp. Mng. & Dirctr.	.327	.331	.337	.334	.354	.309	.335	.352	.334	.377	**.381**
	[17]	[20]	[23]	[24]	[22]	[19]	[22]	[23]	[23]	[22]	**[24]**
Othr Mng. and Prptr.	.308	.308	.307	.304	.330	.287	.304	.324	.308	.340	**.344**
	[12]	[14]	[15]	[16]	[16]	[13]	[14]	[16]	[16]	[13]	**[15]**
Sci., resch., eng. & tech. profnl.	-	.335	.334	.338	.350	.306	.327	.356	.338	.372	**.377**
	[-]	[21]	[22]	[25]	[21]	[18]	[20]	[24]	[24]	[21]	**[23]**
Hlth. Profnl.	.315	.315	.318	.315	.338	.276	.323	.336	.323	**.364**	.360
	[14]	[16]	[18]	[19]	[18]	[10]	[19]	[19]	[20]	**[19]**	[19]
Teach. & Edutnl. Profnl.	.323	.327	.330	.323	.346	.294	.319	.340	.326	**.381**	.373
	[16]	[19]	[21]	[21]	[20]	[15]	[18]	[20]	[21]	**[23]**	[22]
Busi., Media & Pub. Serv. Prof	.319	.323	.322	.319	.342	.298	.331	.344	.330	**.368**	.364
	[15]	[18]	[19]	[20]	[19]	[16]	[21]	[21]	[22]	**[20]**	[20]
Sci., Eng. & Tech. Assoc.	.300	.304	.299	.297	.323	-	.300	.320	-	**.352**	.348
	[10]	[13]	[13]	[14]	[14]	[12]	[13]	[15]	[-]	**[16]**	[16]
Hlth. & Soc. Care Assoc. Profnl.	.292	.282	.281	.282	.299	.261	.286	.298	.286	**.328**	.324
	[8]	[7]	[8]	[10]	[8]	[6]	[9]	[9]	[10]	**[10]**	[10]
Prot. Serv. Occ.	-	.319	.326	-	-	-	.316	.348	.319	**.385**	.369
	[18]	[17]	[20]	[23]	[-]	[17]	[17]	[22]	[19]	**[24]**	[21]
Cultr., Media & Sprt. Occ.	-	.300	.292	.300	.311	.269	.289	.287	.265	**.344**	.316
	[-]	[12]	[11]	[15]	[11]	[8]	[10]	[6]	[4]	**[14]**	[8]
Busi. & Pub. Serv. Assoc. Profnl.	.304	.311	.311	.308	.334	.290	.312	.332	.315	**.356**	.352
	[11]	[15]	[16]	[17]	[17]	[14]	[16]	[18]	[18]	**[17]**	[17]
Admtr. Occ.	.285	.285	.288	.275	.307	.269	.293	.301	.282	**.320**	**.320**
	[6]	[8]	[10]	[8]	[10]	[8]	[11]	[10]	[9]	**[8]**	**[9]**
Secrt. & Reltd. Occ.	.274	.271	.263	.258	.281	.254	.275	.276	.268	**.305**	.301
	[3]	[4]	[3]	[3]	[3]	[4]	[6]	[3]	[5]	**[4]**	[4]
Skld. Agri. & Reltd. Trad.	-	-	.277	.279	-	.265	.282	.309	.289	-	**.328**
	[-]	[-]	[7]	[9]	[-]	[7]	[8]	[12]	[11]	[-]	**[11]**
Skld. Mtl, Elctrl. & Elect. trad.	-	-	.314	.311	.326	-	-	.328	-	**.360**	-
	[-]	[-]	[17]	[18]	[15]	[-]	[15]	[17]	[17]	**[18]**	[18]
Skld. Const. & Bldg Trad.	**.312**	-	-	-	-	.287	-	-	-	-	-
	[13]	[-]	[-]	[-]	[-]	[13]	[-]	[-]	[15]	[15]	[-]
Txtls, Prntg & Othr. Skld. Trad.	-	.278	.274	.272	.292	.258	-	.290	.279	**.313**	.309
	[4]	[6]	[6]	[7]	[6]	[5]	[-]	[7]	[8]	**[6]**	[6]
Crng. Prsnl. Serv. Occ.	.281	.274	.267	.265	.288	.251	.264	.279	.272	**.309**	.305
	[5]	[5]	[4]	[5]	[5]	[3]	[3]	[4]	[6]	**[5]**	[5]
Leis., Trvl. & Rltd. Prsnl.	-	.267	.259	.255	.277	.251	.271	.283	.261	**.301**	.297
Serv. occ.	[-]	[3]	[2]	[2]	[2]	[3]	[5]	[5]	[3]	**[3]**	[3]
Sales Occup.	-	.260	.256	.251	.273	.244	.257	.272	.254	**.297**	.290
	[2]	[1]	[1]	[1]	[1]	[1]	[1]	[2]	[1]	**[2]**	[1]
Custmr. Serv. Occ.	-	-	.270	.268	.296	.258	-	.294	.275	**.316**	.313
	[-]	[-]	[5]	[6]	[7]	[5]	[7]	[8]	[7]	**[7]**	[7]
Prcs, Plnt. & Mach. Oprtvs.	-	.293	.296	-	.315	.272	.286	.313	-	**.336**	**.336**
	[-]	[10]	[12]	[12]	[12]	[9]	[9]	[13]	[13]	**[12]**	**[13]**
Trnspt. & Mob. Mach. Drvrs.	-	-	-	.293	.319	.279	.297	.317	.300	.332	**.340**
& oprtvs.	[9]	[11]	[14]	[13]	[13]	[11]	[12]	[14]	[14]	[11]	**[14]**
Elmtry. Trds. & Rltd. Occ.	.289	.289	-	.286	-	-	-	.305	-	**.324**	-
	[7]	[9]	[9]	[11]	[9]	[-]	[4]	[11]	[12]	**[9]**	[12]
Elmtry. Admn. & Serv. Occ.	.267	.264	.256	.261	.284	.247	.261	.268	.258	**.294**	.293
	[1]	[2]	[1]	[4]	[4]	[2]	[2]	[1]	[2]	**[1]**	[2]
N(Obs. for each GOR)	38	129	106	130	99	155	128	280	149	232	222

Notes.- The analysis consists of overall SOC rankings for mothers in single and two-parent-only households to obtain the likelihood of continuing education. Regional rankings are shown in square brackets. The highlighted margins correspond to the highest likelihood of continuing education for the respective occupation.
(NE-North East; NW-North West; Y & H-Yorkshire & Humber; EM-East Midlands; WM-West Midlands; E-East of England; L-London; SE-South East; WL-Wales; SC-Scotland)

TABLE C.2: Marginal impact of paternal occupations by Government office region

Paternal SOC rank by GOR	NE (1)	NW (2)	Y& H (3)	EM (4)	WM (5)	E (6)	L (7)	SE (8)	SW (9)	WL (10)	SC (11)
Corp. Mng. & Dirctr.	**.356**	.334	.331	.324	.319	.296	.269	.319	.310	.332	.350
	[22]	[21]	[22]	[23]	[11]	[21]	[22]	[24]	[23]	[23]	[25]
Othr Mng. & Prptr.	**.351**	.328	.325	.318	.313	.289	.263	.312	.304	.324	.342
	[16]	[14]	[15]	[16]	[15]	[13]	[14]	[16]	[16]	[14]	[16]
Sci., resch., eng. & tech. profnl.	**.357**	.335	.330	.325	.317	.295	.268	.320	.310	.331	.349
	[23]	[22]	[21]	[24]	[20]	[20]	[20]	[25]	[24]	[22]	[24]
Hlth. Profnl.	**.354**	.331	.327	.320	.315	.286	.267	.316	.307	.329	.346
	[19]	[17]	[18]	[19]	[17]	[10]	[19]	[20]	[20]	[20]	[20]
Teach. & Edutnl. Profnl.	-	.333	-	.322	.316	.291	.266	.316	.308	.333	**.348**
	[21]	[20]	[-]	[21]	[19]	[16]	[18]	[21]	[21]	[24]	**[23]**
Busi., Media & Pub. Serv. Prof	**.354**	.332	.328	.321	.316	.292	.269	.317	.309	.330	.346
	[20]	[19]	[19]	[20]	[18]	[17]	[21]	[22]	[22]	[21]	[21]
Sci., Eng. & Tech. Assoc.	**.348**	.327	.322	.315	.311	-	.262	.311	.302	.326	.343
	[13]	[13]	[12]	[13]	[13]	[12]	[13]	[15]	[14]	[17]	[17]
Hlth. & Soc. Care Assoc. Profnl.	**.344**	-	-	-	-	.282	.258	-	-	.321	.336
	[9]	[6]	[8]	[9]	[-]	[5]	[8]	[9]	[9]	[11]	[10]
Prot. Serv. Occ.	-	.331	.329	-	.318	.293	-	.318	.306	.334	**.347**
	[24]	[18]	[20]	[22]	[21]	[18]	[17]	[23]	[19]	[25]	**[22]**
Cultr., Media & Sprt. Occ.	**.349**	-	-	-	-	.284	.259	.303	.293	.325	.334
	[14]	[-]	[-]	[14]	[9]	[7]	[9]	[6]	[4]	[15]	[8]
Busi. & Pub. Serv. Assoc. Profnl.	**.350**	.329	.325	.318	.314	.290	.265	.315	.305	.327	.344
	[15]	[15]	[16]	[17]	[16]	[14]	[16]	[19]	[18]	[18]	[18]
Admtr. Occ.	-	.322	.320	.310	.306	.284	.260	.307	.297	.318	**.335**
	[7]	[7]	[10]	[7]	[7]	[7]	[10]	[10]	[8]	[8]	**[9]**
Secrt. & Reltd. Occ.	-	-	-	-	-	-	-	.301	-	-	**.331**
	[-]	[-]	[3]	[-]	[-]	[-]	[-]	[3]	[5]	[4]	**[4]**
Skld. Agri. & Reltd. Trad.	**.345**	.323	.317	.311	.307	.283	-	.308	.298	.319	.337
	[10]	[9]	[7]	[8]	[8]	[6]	[-]	[12]	[10]	[9]	[11]
Skld. Mtl, Elctrl. & Elect. trad.	**.353**	.330	.326	.319	.312	.291	.264	.314	.304	.328	.345
	[18]	[16]	[17]	[18]	[14]	[15]	[15]	[18]	[17]	[19]	[19]
Skld. Const. & Bldg Trad.	**.352**	.326	.324	.317	.310	.289	.261	.313	.303	.326	.341
	[17]	[12]	[14]	[15]	[12]	[13]	[11]	[17]	[15]	[16]	[15]
Txtls, Prntg & Othr. Skld. Trad.	-	**.350**	.317	.309	.304	-	.257	.304	.296	.316	.333
	[4]	**[5]**	[6]	[6]	[5]	[4]	[6]	[7]	[7]	[6]	[6]
Crng. Prsnl. Serv. Occ.	**.340**	-	.315	.307	.303	-	.254	.301	-	-	.332
	[5]	[4]	[4]	[4]	[4]	[3]	[3]	[4]	[-]	[5]	[5]
Leis., Trvl. & Rltd. Prsnl.	-	-	**.313**	.305	-	-	-	-	.292	-	-
Serv. occ.	[3]	[3]	**[2]**	[2]	[2]	[3]	[5]	[5]	[3]	[3]	[3]
Sales Occup.	-	**.316**	.312	-	-	.279	.253	.300	.291	-	-
	[2]	**[1]**	[1]	[1]	[1]	[1]	[1]	[2]	[1]	[2]	[1]
Custmr. Serv. Occ.	**.341**	.320	-	.308	-	-	.257	.305	.295	.317	.334
	[6]	[5]	[5]	[5]	[6]	[4]	[7]	[8]	[6]	[7]	[7]
Prcs, Plnt. & Mach. Oprtvs.	**.347**	.324	.321	.313	.309	.285	.258	-	.300	.323	.339
	[12]	[10]	[11]	[11]	[10]	[9]	[8]	[13]	[12]	[13]	[13]
Trnspt. & Mob. Mach. Drvrs.	**.346**	.325	.323	-	.309	.287	.261	.310	.301	.322	.340
& oprtvs.	**[11]**	[11]	[13]	[12]	[11]	[11]	[12]	[14]	[13]	[12]	[14]
Elmtry. Trds. & Reltd. Occ.	**.380**	-	.319	.312	-	-	-	.308	-	-	-
	[8]	[8]	[9]	[10]	[-]	[8]	[4]	[11]	[11]	[10]	[12]
Elmtry. Admtr. & Serv. Occ.	**.337**	.317	.312	.306	.303	-	.253	.299	.292	.312	.329
	[1]	[2]	[1]	[3]	[3]	[2]	[2]	[1]	[2]	[1]	[2]
N(Obs. for each GOR)	25	62	45	75	64	83	73	146	74	123	121

Notes.- The analysis consists of overall SOC rankings for fathers in single and two-parent-only households to obtain the likelihood of continuing education. Regional rankings are shown in square brackets. The highlighted margins correspond to the highest likelihood of continuing education for the respective occupation.
(NE-North East; NW-North West; Y & H-Yorkshire & Humber; EM-East Midlands; WM-West Midlands; E-East of England; L-London; SE-South East; WL-Wales; SC-Scotland)

TABLE C.3: Marginal impact of maternal occupation in two-parent household by Government office region

Govt. Office Regions:	NE (1)	NW (2)	Y& H (3)	EM (4)	WM (5)	E (6)	L (7)	SE (8)	SW (9)	WL (10)	SC (11)
Matrnl. SOC rank by GOR											
Corp. Mng. & Dirctr.	.364[17]	.352[20]	.354[23]	.329[24]	.349[22]	.285[19]	.316[22]	.348[23]	.339[23]	.376[22]	**.379[24]**
Othr Mng. & Prptr.	**.346[12]**	.331[14]	.326[15]	.302[16]	–[16]	.265[13]	.290[14]	.323[16]	.315[16]	.343[13]	**.346[15]**
Sci., resch., eng. and tech. profnl.	-	.356[21]	.351[22]	–[25]	.345[21]	.281[18]	.309[20]	.351[24]	.343[24]	.372[21]	**.376[23]**
Hlth. Profnl.	.354[14]	.338[16]	.337[18]	.312[19]	.335[18]	.256[10]	.306[19]	.333[19]	.329[20]	**.365[19]**	.361[19]
Teach. & Edutnl. Profnl.	.361[16]	.348[19]	.347[21]	.319[21]	.342[20]	.272[15]	.303[18]	.337[20]	.332[21]	**.380[23]**	.372[22]
Busi., Media & Pub. Serv. Prof	.357[15]	.345[18]	.340[19]	.315[20]	.338[19]	.275[16]	.313[21]	.341[21]	.336[22]	**.369[20]**	.365[20]
Sci., Eng. & Tech. Assoc.	–[10]	–[13]	.320[13]	.295[14]	–[14]	–[12]	–[13]	.320[15]	-	**.354[16]**	.350[16]
Hlth. & Soc. Care Assoc. Profnl.	**.332[8]**	.306[7]	.303[8]	.282[10]	–[8]	.244[6]	.274[9]	.299[9]	.295[10]	**.332[10]**	.329[10]
Prot. Serv. Occ.	–[18]	–[17]	–[20]	–[23]	-	–[17]	–[17]	–[22]	–[19]	**.383[24]**	.368[21]
Cultr., Media & Sprt. Occ.	-	–[12]	–[11]	.299[15]	**.311[11]**	.250[8]	.277[10]	.289[6]	.275[4]	–[14]	–[8]
Busi. & Pub. Serv. Assoc. Profnl.	.342[11]	.334[15]	.330[16]	.305[17]	.331[17]	.268[14]	.296[16]	.330[18]	.322[18]	**.358[17]**	.354[17]
Admtr. Occ.	.325[6]	.310[8]	.310[10]	.276[8]	.307[10]	.250[8]	.280[11]	.303[10]	.292[9]	**.325[8]**	**.325[9]**
Secrt. & Reltd. Occ.	–[3]	.296[4]	.287[3]	.260[3]	.284[3]	.238[4]	.264[6]	.279[3]	.278[5]	**.311[4]**	.308[4]
Skld. Agri. & Reltd. Trad	-	-	–[7]	–[9]	-	.247[7]	.270[8]	.309[12]	.298[11]	-	**.332[11]**
Skld. Mtl, Elctrl. & Elect. trad	-	-	–[17]	–[18]	.324[15]	-	–[15]	–[17]	–[17]	**.361[18]**	–[18]
Skld. Const. & Bldg Trad.	**.350[13]**	-	-	-	-	.265[13]	-	-	–[15]	–[15]	-
Txtls, Prntg & Othr. Skld. Trad.	–[4]	.303[6]	.296[6]	.273[7]	.294[6]	.241[5]	-	.293[7]	–[8]	**.318[6]**	–[6]
Crng. Prsnl. Serv. Occ.	**.321[5]**	.300[5]	.290[4]	.267[5]	.291[5]	.235[3]	.255[3]	.283[4]	.282[6]	.315[5]	.311[5]
Leis., Trvl. & Reltd. Prsnl. Serv. occ.	-	–[3]	–[2]	.257[2]	–[2]	.235[3]	.261[5]	–[5]	.272[3]	**.308[3]**	.304[3]
Sales Occup.	–[2]	.286[1]	.280[1]	–[1]	.278[1]	.229[1]	–[1]	–[2]	.266[1]	**.304[2]**	.298[1]
Custmr. Serv. Occ.	-	.303[6]	–[5]	–[6]	.297[7]	.241[5]	–[7]	.296[8]	.285[7]	–[7]	**.318[7]**
Prcs, Plnt. & Mach. Oprtvs.	-	–[10]	–[12]	–[12]	.314[12]	–[9]	.274[9]	–[13]	–[13]	**.339[12]**	**.339[13]**
Trnspt. & Mob. Mach. Drvrs. & oprtvs.	–[9]	–[11]	–[14]	–[13]	–[13]	–[11]	.283[12]	**.316[14]**	–[14]	–[11]	–[14]
Elmtry. Trds. & Reltd. Occ.	–[7]	–[9]	–[9]	.286[11]	–[9]	-	–[4]	**.306[11]**	–[12]	–[9]	–[12]
Elmtry. Admtr. & Serv. Occ.	**.307[1]**	–[2]	.280[1]	.263[4]	.287[4]	.232[2]	.252[2]	.273[1]	.269[2]	.301[1]	.301[2]
N(Obs. for each GOR)	22	50	36	58	53	70	50	127	64	96	95

Notes.- The analysis consists of overall SOC rankings for mothers in two-parent-only households to obtain the likelihood of continuing education. Regional rankings are shown in square brackets. The highlighted margins correspond to the highest likelihood of continuing education for the respective occupation.
(NE-North East; NW-North West; Y & H-Yorkshire & Humber; EM-East Midlands; WM-West Midlands; E-East of England; L-London; SE-South East; WL-Wales; SC-Scotland)

TABLE C.4: Marginal impact of paternal occupation in two-parent household by Government office region

Govt. Office Regions:	NE (1)	NW (2)	Y& H (3)	EM (4)	WM (5)	E (6)	L (7)	SE (8)	SW (9)	WL (10)	SC (11)
Patrnl. SOC rank by GOR											
Corp. Mng. & Dirctr.	.331[22]	.318[21]	.312[20]	.302[23]	.313[22]	.254[21]	.277[22]	.303[24]	.304[22]	**.341[23]**	.339[25]
Othr Mng. & Prptr.	.339[16]	.326[14]	.318[15]	.309[16]	.321[15]	.262[13]	.285[14]	.312[16]	.311[16]	**.352[14]**	.350[16]
Sci., resch., eng. & tech. profnl.	.330[23]	.317[22]	.311[21]	.301[24]	.315[20]	.255[20]	.279[20]	.301[25]	.302[24]	**.342[22]**	.340[24]
Hlth. Profnl.	–[19]	.323[17]	.314[18]	.306[19]	.319[17]	–[10]	.280[19]	.307[20]	.307[20]	**.345[20]**	**.345[20]**
Teach. & Edutnl. Profnl.	–[21]	.319[20]	-	.304[21]	.316[19]	.259[16]	.281[18]	.306[21]	.306[21]	.340[24]	**.341[23]**
Busi., Media & Pub. Serv. Prof	–[20]	.320[19]	.313[19]	.305[20]	.318[18]	.258[17]	.278[21]	.305[22]	.304[22]	**.343[21]**	**.343[21]**
Sci., Eng. & Tech. Assoc.	.342[13]	.328[13]	–[12]	.313[13]	.324[13]	–[12]	.286[13]	.313[15]	.314[14]	**.348[17]**	**.348[17]**
Hlth. & Soc. Care Assoc. Profnl.	.347[9]	–[6]	–[8]	–[9]	-	.271[5]	–[8]	–[9]	–[9]	.356[11]	**.357[10]**
Prot. Serv. Occ.	–[24]	.322[18]	.312[20]	–[22]	.314[21]	–[18]	–[17]	.304[23]	.308[19]	.338[25]	**.342[22]**
Cultr., Media & Sprt. Occ.	.341[14]	-	-	–[14]	–[9]	.269[7]	.291[9]	.324[6]	.326[4]	.351[15]	**.360[8]**
Busi. & Pub. Serv. Assoc. Profnl.	.340[15]	.325[15]	.317[16]	.308[17]	.320[16]	.261[14]	.283[16]	.309[19]	.309[18]	**.347[18]**	**.347[18]**
Admtr. Occ.	–[7]	.335[7]	–[10]	.319[7]	.331[7]	.269[7]	.290[10]	.319[10]	.321[8]	**.360[8]**	.358[9]
Secrt. & Reltd. Occ.	-	-	–[3]	-	-	-	-	.328[3]	–[5]	–[4]	**.365[4]**
Skld. Agri. & Reltd. Trad.	.346[10]	.332[9]	.327[7]	.318[8]	.330[8]	.270[6]	-	.317[12]	.318[10]	**.359[9]**	.356[11]
Skld. Mtl, Elctrl. & Elect. trad.	.336[18]	.324[16]	–[17]	.307[18]	.322[14]	.260[15]	.284[15]	.310[18]	.310[17]	**.346[19]**	**.346[19]**
Skld. Const. and Bldg Trad.	.337[17]	.329[12]	.319[14]	.310[15]	.325[12]	.262[13]	.289[11]	.311[17]	.313[15]	.350[16]	**.351[15]**
Txtls, Prntg & Othr. Skld. Trad.	–[4]	.337[5]	.328[6]	.320[6]	–[5]	–[4]	.294[6]	.323[7]	.322[7]	**.362[6]**	**.362[6]**
Crng. Prsnl. Serv. Occ.	–[5]	–[4]	.331[4]	.323[4]	.334[4]	–[3]	.298[3]	.327[4]	-	–[5]	**.363[5]**
Leis., Trvl. & Reltd. Prsnl. Serv. occ.	–[3]	–[3]	.333[2]	.325[2]	–[2]	–[3]	–[5]	–[5]	**.327[3]**	–[3]	–[3]
Sales Occup.	–[2]	–[1]	**.334[1]**	–[1]	–[1]	.276[1]	.300[1]	.329[2]	.329[1]	–[2]	–[1]
Custmr. Serv. Occ.	.351[6]	.337[5]	–[5]	.322[5]	–[6]	–[4]	–[7]	–[8]	.323[6]	**.361[7]**	**.361[7]**
Prcs, Plnt. & Mach. Oprtvs.	.344[12]	.331[10]	.322[11]	.315[11]	.327[10]	.267[9]	–[8]	–[13]	.316[12]	**.353[13]**	**.353[13]**
Trnspt. & Mob. Mach. Drvrs. & oprtvs.	.345[11]	.330[11]	.320[13]	–[12]	.326[11]	.265[11]	.288[12]	.315[14]	.315[13]	**.355[12]**	.352[14]
Elmtry. Trds. & Reltd. Occ.	–[8]	–[8]	**.325[9]**	.316[10]	-	–[8]	–[4]	.318[11]	–[11]	–[10]	–[12]
Elmtry. Admtr. & Serv. Occ.	.357[1]	.341[2]	.334[1]	.324[3]	.336[3]	–[2]	–[2]	.330[1]	.328[2]	**.369[1]**	.367[2]
N(Obs. for each GOR)	22	50	38	60	53	69	53	126	64	96	95

Notes.- The analysis consists of overall SOC rankings for fathers in two-parent-only households to obtain the likelihood of continuing education. Regional rankings are shown in square brackets. The highlighted margins correspond to the highest likelihood of continuing education for the respective occupation.

(NE-North East; NW-North West; Y & H-Yorkshire & Humber; EM-East Midlands; WM-West Midlands; E-East of England; L-London; SE-South East; WL-Wales; SC-Scotland)

Bibliography

Almond, D., & Currie, J. (2011). Human Capital Development before Age Five. In O. Ashenfelter, & D. Card (Eds.) *Handbook of Labor Economics*. Elsevier.

Anders, J. (2017). The influence of socioeconomic status on changes in young people's expectations of applying to university. *Oxford Review of Education*, *43*(4), 381–401.

Anders, J., Macmillan, L., Sturgis, P., & Wyness, G. (2021). Inequalities in young people's educational experiences and wellbeing during the Covid-19 pandemic. CEPEO Working Paper Series, UCL Centre for Education Policy and Equalising Opportunities.

ASHE (2023). Annual Survey of Hours and Earnings, 2020-2021: Weekly pay - Gross (£) - For all employee jobs$^{a} : United\ Kingdom, 2021, Online\ accessed - 9/7/2023$.

Asher, J. J., & García, R. (1969). The optimal age to learn a foreign language. *The Modern Language Journal*, *53*(5), 334–341.

Attanasio, O., Cattan, S., & Meghir, C. (2022). Early childhood development, human capital, and poverty. *Annual Review of Economics*, *14*(1), 853–892.

Bailey, N., & Livingston, M. (2008). Selective migration and neighbourhood deprivation: evidence from 2001 census migration data for England and Scotland. *Urban Studies*, *45*(4), 943–961.

Beck, B., Buttaro Jr., A., & Lennon, M. C. (2016). Home moves and child wellbeing in the first five years of life in the united states. *Longitudinal and Life Course Studies*, *7*(3), 240–264.

Bedore, L. M., Peña, E. D., Summers, C. L., Boerger, K. M., Resendiz, M. D., Greene, K., Bohman, T. M., & Gillam, R. B. (2012). The measure matters: Language dominance profiles across measures in spanish–english bilingual children. *Bilingualism: Language and Cognition*, *15*(3), 616–629.

Bialystok, E. (2001). *Bilingualism in Development: Language, Literacy, and Cognition*. Cambridge University Press.

Bialystok, E., Craik, F. I. M., & Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends in Cognitive Sciences*, *16*(4), 240–250.

Blanden, J., Doepke, M., & Stuhler, J. (2022). Education inequality. CEP Discussion Papers dp1849, Centre for Economic Performance, LSE.

Bohman, T. M., Bedore, L. M., Peña, E. D., Mendez-Perez, A., & Gillam, R. B. (2010). What you hear and what you say: Language performance in Spanish–English bilinguals. *International Journal of Bilingual Education and Bilingualism*, *13*(3), 325–344.

Bono, E. D., Francesconi, M., Kelly, Y., & Sacker, A. (2016). Early Maternal Time Investment and Early Child Outcomes. *The Economic Journal*, *126*(596), F96–F135.

Bornstein, M. H., Tamis-LeMonda, C. S., & Haynes, O. M. (1999). First words in the second year: Continuity, stability, and models of concurrent and predictive correspondence in vocabulary and verbal responsiveness across age and context. *Infant Behavior and Development*, *22*(1), 65–85.

Bowles, R. P., & Salthouse, T. A. (2008). Vocabulary test format and differential relations to age. *Psychology and Aging*, *23*(2), 366–376.

Bowles, S., & Gintis, H. (2002). Schooling in Capitalist America Revisited. *Sociology of Education*, *75*(1), 1–18.

Brotherson, S. E. (2005). *Understanding brain development in young children*. Citeseer.

Byers-Heinlein, K., & Lew-Williams, C. (2013). Bilingualism in the early years: What the science says. *Learning Landscapes*, *7*(1), 95–112.

Carey, W. B. (1972). Clinical applications of infant temperament measurements. *The Journal of Pediatrics*, *81*(2), 823–828.

Carey, W. B., & McDevitt, S. C. (1978). Revision of the Infant Temperament Questionnaire. *Pediatrics*, *61*(5), 735–739.

Carneiro, P., Cunha, F., & Heckman, J. J. (2003). Interpreting the evidence of family influence on child development. *The Economics of Early Childhood Development: Lessons for Economic Policy. The Federal Reserve Bank Minneapolis, Minnesota*.

Chevalier, A., Harmon, C., O'Sullivan, V., & Walker, I. (2010). The Impact of Parental Income and Education on the Schooling of their Children. *Lancaster University Management School*, *036*.

Chiswick, B., & Gindelsky, M. (2014). Determinants of Bilingualism among Children. IZA Discussion Papers 8488, Institute of Labor Economics (IZA).

Clair, A. (2019). Housing: an Under-Explored Influence on Children's Well-Being and Becoming. *Child Indicators Research*, *12*, 609–626.

Clark, W. (2016). Life events and moves under duress: Disruption in the life course and mobility outcomes. *Longitudinal and Life Course Studies*, *7*.

Clark, W., Van Ham, M., & Coulter, R. (2014). Spatial mobility and social outcomes. *Journal of Housing and the Built Environment*, *29*(4), 699–727.

Clark, W. A. V., & Huang, Y. (2003). The Life Course and Residential Mobility in British Housing Markets. *Environment and Planning A: Economy and Space*, *35*(2), 323–339.

Clifton-Sprigg, J. (2016). Best of both worlds? Early cognitive and non-cognitive development of bilingual children. *Unpublished working paper*.

Cobb-Clark, D. A., Harmon, C., & Staneva, A. (2018). The Bilingual Gap in Children's Language and Emotional Development. Working Papers 2018-075, Human Capital and Economic Opportunity Working Group.

Collins, B. A., Toppelberg, C. O., Suárez-Orozco, C., O'Connor, E., & Nieto-Castañon, A. (2011). Cross-sectional associations of Spanish and English competence and well-being in Latino children of immigrants in kindergarten. *International Journal of the Sociology of Language*, *2011*(208), 5–23.

Connelly, R. (2013). Millennium Cohort Study data note 2013/1: Interpreting test scores. *London: Centre for Longitudinal Studies, Institute of Education*.

Cortina, J., & Taran, P. A. (2014). *Migration and youth: Challenges and opportunities*. United Nations Children's Fund.

Cummins, J. (2005). A Proposal for Action: Strategies for Recognizing Heritage Language Competence as a Learning Resource within the Mainstream Classroom. *The Modern Language Journal*, *89*(4), 585–592.

Cunha, F., & Heckman, J. (2007). The Technology of Skill Formation. *American Economic Review*, *97*(2), 31–47.

Cunha, F., Heckman, J. J., & Navarro, S. (2005). Separating uncertainty from heterogeneity in life cycle earnings. *Oxford Economic Papers*, *57*(2), 191–261.

Cunha, F., Heckman, J. J., & Schennach, S. M. (2010). Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Econometrica*, *78*(3), 883–931.

De Houwer, A. (2007). Parental language input patterns and children's bilingual use. *Applied Psycholinguistics*, *28*(3), 411–424.

De Houwer, A. (2015). Harmonious bilingual development: Young families' well-being in language contact situations. *International Journal of Bilingualism*, *19*(2), 169–184.

Dearden, L., Sibieta, L., & Sylva, K. (2011). The socio-economic gradient in early child outcomes: Evidence from the Millennium Cohort Study. IFS Working Papers 11,03, London.

Del Boca, D., Monfardini, C., & Nicoletti, C. (2012). Self Investments of Adolescents and their Cognitive Development. IZA Discussion Papers 6868, Institute of Labor Economics (IZA).

Del Boca, D., Monfardini, C., & Nicoletti, C. (2017). Parental and Child Time Investments and the Cognitive Development of Adolescents. *Journal of Labor Economics*, *35*(2), 565–608.

Department for Communities & Local Government (2015). The English Index of Multiple Deprivation (IMD) 2015 – Guidance. Society, Department for Communities & Local Government.

Department for Education (2013). The national curriculum in England: Key Stages 1 and 2 framework document. Tech. rep. DFE-00178-2013, Department for Education.

Department for Work & Pension (2020). Households Below Average Income: An analysis of the UK income distribution: 1994/95-2018/19. Tech. rep., Department for Work Pension.

Dickerson, A., & Popli, G. (2018). The Many Dimensions of Child Poverty: Evidence from the UK Millennium Cohort Study. *Fiscal Studies*, *39*(2), 265–298.

Dickerson, A., & Popli, G. K. (2016). Persistent poverty and children's cognitive development: evidence from the UK Millennium Cohort Study. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *179*(2), 535–558.

Doepke, M., & Zilibotti, F. (2022). The role of parenting in child development. IFS Deaton Review of Inequalities, Institute for Fiscal Studies.

Duursma, E., Romero-Contreras, S., Szuber, A., Proctor, P., Snow, C., August, D., & Calderón, M. (2007). The role of home literacy and language environment on bilinguals' English and Spanish vocabulary development. *Applied Psycholinguistics*, *28*(1), 171–190.

Elliott, C. D., Smith, P., & McCulloch, K. (1996). British Ability Scales II: Administration and Scoring Manual. *Windsor, Berkshire: NFER-Nelson*, *406*.

Farkas, G. (2003). Cognitive Skills and Noncognitive Traits and Behaviors in Stratification Processes. *Annual Review of Sociology*, *29*(1), 541–562.

Fenson, L., Pethick, S., Renda, C., Dale, P., & Reznick, J. (1998). Short-form version of the MacArthur Communicative Development Inventories. *Infant Behavior Development*, *21*, 95–116.

Fiorini, M., & Keane, M. P. (2014). How the allocation of children's time affects cognitive and noncognitive development. *Journal of Labor Economics*, *32*(4), 787–836.

Flouri, E., Mavroveli, S., & Midouhas, E. (2013). Residential mobility, neighbourhood deprivation and children's behaviour in the UK. *Health Place*, *20*, 25–31.

Fradkin, A., Panier, F., & Tojerow, I. (2019). Blame the parents? how parental unemployment affects labor supply and job quality for young adults. *Journal of Labor Economics*, *37*(1), 35–100.
URL `https://doi.org/10.1086/698896`

Francesconi, M., & Heckman, J. J. (2016). Child Development and Parental Investment: Introduction. *Economic Journal*, *126*(596), 1–27.

Frankenburg, W., & Dodds, J. (1967). The Denver Developmental Screening Test. *Journal of Pediatrics*, *71*(2), 181–191.

Frankenburg, W., Dodds, J., Archer, P., Shapiro, H., & Bresnick, B. (1992). The Denver II: A major revision and restandardization of the Denver Developmental Screening Test. *Pediatrics*, *89*(1), 91–97.

Gambaro, L., & Joshi, H. (2016). Moving home in the early years: What happens to children in the UK? *Longitudinal and Life Course Studies*, *7*.

Gambaro, L., Joshi, H., & Lupton, R. (2017). Moving to a better place? Residential mobility among families with young children in the Millennium Cohort Study. *Population, Space and Place*, *23*(8), 1–14.

Goldberg, H., Paradis, J., & Crago, M. (2008). Lexical acquisition over time in minority L1 children learning English as L2. *Applied Psycholinguistics*, *29*(01), 41–65.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry*, *38*(5), 581–586.

Guiberson, M. (2013). Bilingual Myth-Busters Series Language Confusion in Bilingual Children. *Perspectives on Communication Disorders and Sciences in Culturally and Linguistically Diverse Populations*, *20*, 5.

Hakuta, K., & Diaz, R. M. (1985). The relationship between degree of bilingualism and cognitive ability: A critical discussion and some new longitudinal data. In *Children's language, Vol.5*, (pp. 319–344). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Hammer, C. S., Dunn Davison, M., Lawrence, F. R., & Miccio, A. W. (2009). The Effect of Maternal Language on Bilingual Children's Vocabulary and Emergent Literacy Development During Head Start and Kindergarten. *Scientific Studies of Reading*, *13*(2), 99–121.

Hansen, K., Jones, E., Joshi, H., & Budge, D. (2010). Millennium Cohort Study Fourth Survey: A user's guide to initial findings. London: Centre for Longitudinal Studies. *Institute of Education, University of London*.

Harmon, C., Staneva, A., & Cobb-Clark, D. (2017). Early Cognitive Development and Educational Assimilation of Bilingual Children Working Draft.

Harris, D., Sass, T., & Semykina, A. (2010). Value-Added Models and the Measurement of Teacher Productivity. CALDER Working Paper No.54. *Urban Institute*.

Heckman, J. J. (2011). The Economics of Inequality The Value of Early Childhood Education. *The American Educator*, *35*(1), 31–35.

Heckman, J. J., & Mosso, S. (2014). The Economics of Human Development and Social Mobility. *Annual Review of Economics*, *6*(1), 689–733.

Hilger, N. G. (2016). Parental job loss and children's long-term outcomes: Evidence from 7 million fathers' layoffs. *American Economic Journal: Applied Economics*, *8*(3), 247–283.
URL `http://www.jstor.org/stable/24739135`

Hoff, E., Burridge, A., Ribot, K. M., & Giguere, D. (2018). Language specificity in the relation of maternal education to bilingual children's vocabulary growth. *Developmental Psychology*, *54*(6), 1011–1019.

Huitt, W., & Hummel, J. (2003). Piaget's theory of cognitive development. *Educational psychology interactive*, *3*(2), 1–5.

Hupkau, C., Isphording, I. E., Machin, S., & Ruiz-Valenzuela, J. (2020). Labour Market Shocks during the COVID-19 Pandemic, Inequalities and Child Outcomes. IZA Discussion Papers 14000, Institute of Labor Economics (IZA).

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: relation to language input and gender. *Developmental psychology*, *27*(2), 236.

Huttunen, K., & Riukula, K. (2019). Parental Job Loss and Children's Careers. IZA Discussion Papers 12788, Institute of Labor Economics (IZA).
URL `https://ideas.repec.org/p/iza/izadps/dp12788.html`

Jelleyman, T., & Spencer, N. (2008). Residential mobility in childhood and health outcomes: A systematic review. *Journal of epidemiology and community health*, *62*, 584–92.

Johnson, J., Atkinson, M., Rosenberg, R., Schoon, I., Joshi, H., & Smith, K. (2015). Millenium Cohort Study Psychological, Developmental, and Health Inventories: User guide to the data. *Centre for Longitudinal Studies*.

Kaushanskaya, M., Gross, M., & Buac, M. (2013). Gender differences in child word learning. *Learning and Individual Differences*, *27*, 82–89.

Keatley, C. W. (1992). History of Bilingualism Research in Cognitive Psychology. In R. J. Harris (Ed.) *Cognitive Processing in Bilinguals*, vol. 83, (pp. 15–49). North-Holland.

Kelly, Y., Sacker, A., Bono, E., Francesconi, M., & Marmot, M. (2011). What role for the home learning environment and parenting in reducing the socioeconomic gradient in child development? Findings from the Millennium Cohort Study. *Archives of disease in childhood*, *96*(9), 832–7.

Kim, H., Barron, C., Sinclair, J., & Eunhee Jang, E. (2020). Change in home language environment and English literacy achievement over time: A multi-group latent growth curve modeling investigation. *Language Testing*, *37*(4), 573–599.

Kirk, E., Pine, K., Wheatley, L., Howlett, N., Schulz, J., & Fletcher, B. C. (2015). A longitudinal investigation of the relationship between maternal mind-mindedness and theory of mind. *British Journal of Developmental Psychology*, *33*(4), 434–445.

Kramarz, F., & Skans, O. N. (2014). When Strong Ties are Strong: Networks and Youth Labour Market Entry. *Review of Economic Studies*, *81*(3), 1164–1200.

Lanza, E. (1992). Can bilingual two-year-olds code-switch? *Journal of Child Language*, *19*(3), 633–658.

Lanza, E. (1997). *Language Mixing in Infant Bilingualism: A Sociolinguistic Perspective*. Oxford University Press.

Li, R., Liu, T., & Liu, X. (2022). Parental Factors Influencing Expenditure in Preschool Education: Evidence from 2018 China Family Panel Studies (CFPS). In *Proceedings of the 2022 5th International Conference on Humanities Education and Social Sciences (ICHESS 2022)*, (pp. 2929–2942). Atlantis Press.

Lim, V., & Sng, Q. (2006). Does Parental Job Insecurity Matter? Money Anxiety, Money Motives, and Work Motivation. *The Journal of Applied Psychology*, *91*, 1078–87.

Lupton, R. (2016). The Influence(s) of housing policies on the residential moves of families with young children. *Longitudinal and Life course Studies*, *7*(3), 288–301.

Macmillan, L., & Tominey, E. (2023). Parental inputs and socio-economic gaps in early child development. *Journal of Population Economics*, *36*(3), 1513–1543.

Major, L., & Machin, S. (2020). Covid-19 and social mobility. CEP Covid-19 Analysis, London: Centre for Economic Performance.

Major, L. E., Eyles, A., & Machin, S. (2021). Unequal learning and labour market losses in the crisis: consequences for social mobility. CEP Discussion Papers dp1748, Centre for Economic Performance, LSE.

Massey, D. S., Arango, J., Hugo, G., Kouaouci, P. A., Ali, & Taylor J., E. (1993). Theories of International Migration: A Review and Appraisal. *Population and Development Review*, *19*(3), 431–466.

Mattock, K., Polka, L., Rvachew, S., & Krehm, M. (2010). The first steps in word learning are easier when the shoes fit: comparing monolingual and bilingual infants. *Developmental Science*, *13*(1), 229–243.

Métraux, R. W. (1965). A Study of Bilingualism among Children of U.S - French Parents. *The French Review*, *38*(5), 650–665.

Ministry of Housing Communities & Local Government (2022). Tackling the under-supply of housing in England. Tech. rep., Ministry of Housing, Communities & Local Government.

Mooi-Reci, I., Bakker, B., Curry, M., & Wooden, M. (2019). Why Parental Unemployment Matters for Children's Educational Attainment: Empirical Evidence from The Netherlands. *European Sociological Review*, *35*(3), 394–408.

Murphy, R., & Wyness, G. (2020). Minority Report: the impact of predicted grades on university admissions of disadvantaged groups. CEPEO Working Paper Series 20-07, UCL Centre for Education Policy and Equalising Opportunities.

National Centre for Social Research (2003). Millennium Cohort Study First Survey: CAPI Questionnaire Documentation. Tech. rep., Department for Education.

Nicoletti, C., & Rabe, B. (2019). Sibling spillover effects in school achievement. *Journal of Applied Econometrics*, *34*(4), 482–501.

Noble, M., Wright, G., Dibben, C., Smith, G., McLennan, D., Anttila, C., Barnes, H., Mokhtar, C., Noble, S., Avenell, D., Gardner, J., Covizzi, I., & Lloyd, M. (2004). The English Indices of Deprivation 2004. Tech. rep., Office of the Deputy Prime Minister.

Oishi, S., & Schimmack, U. (2010). Residential mobility, well-being, and mortality. *Journal of personality and social psychology*, *98*(6), 980–94.

ONS (2023). SOC 2020 Volume 1: structure and descriptions of unit groups. Tech. rep., Office for National Statistics, UK Data Service.

ONS-Edu. (2023). Education, England and Wales:Census 2021. Tech. rep., Office for National Statistics, ONS website, Statistical bulletin.

Özerk, M., Handorff, J. A., & Özerk, K. (2011). Assessment of bilingual children with inattention, over activity and impulsivity–Challenges and solutions. *International Electronic Journal of Elementary Education*, *3*(3), 193–212.

Patterson, J. L. (1999). What Bilingual Toddlers Hear and Say: Language Input and Word Combinations. *Communication Disorders Quarterly*, *21*(1), 32–38.

Pearson, B. Z. (2008). *Raising a bilingual child*. Living Language.

Penney, B. (2019). The English Indices of Deprivation 2019 (IoD2019)g. Statistical Release, Ministry of Housing, Communities Local Governments.

Plewis, I., Calderwood, L., Hawkes, D., Hughes, G., & Joshi, H. (2007). Millennium Cohort Study: Technical report on sampling. *London: Centre for Longitudinal Studies*.

Ray-Chaudhuri, S., & Xu, X. (2023). Are the kids alright? The early careers of education leavers since the COVID-19 pandemic. IFS Report R237, London: Institute for Fiscal Studies.

Rege, M., Telle, K., & Votruba, M. (2011). Parental job loss and children's school performance. *The Review of Economic Studies*, *78*(4), 1462–1489. URL `http://www.jstor.org/stable/41407068`

Reyes, I., & Hernández, A. E. (2006). Sentence interpretation strategies in emergent bilingual children and adults. *Bilingualism: Language and Cognition*, *9*(1), 51–69.

Rose, S., Badr, K., Fletcher, L., Paxman, T., Lord, P., Rutt, S., Styles, B., & Twist, L. (2021). Impact of School Closures and Subsequent Support Strategies on Attainment and Socio-Emotional Wellbeing in Key Stage 1. Research Report. *Education Endowment Foundation*.

Ruiz-Valenzuela, J. (2015). Job Loss at Home: Children's School Performance During the Great Recession in Spain. CEP Discussion Papers 1364, Centre for Economic Performance, LSE.

Rumbold, A., Giles, L., Whitrow, M., Steele, E., Davies, C., Davies, M., & Moore, V. (2012). The effects of house moves during early childhood on child mental health at age 9 years. *BMC public health*, *12*, 583.

Russell, M., & Lister, R. (2020). Moving, Always Moving. Tech. rep., University of Bath.

Schiller, P. (2010). Early Brain Development Research Review and Update.

Shonkoff, J. P., Phillips, D. A., Council, N. R., et al. (2000). Rethinking nature and nurture. In *From neurons to neighborhoods: The science of early childhood development*. National Academies Press (US).

Singer, D. G., & Revenson, T. A. (1997). *A Piaget primer: How a child thinks*. ERIC.

Snow, C. E., & Hoefnagel-Höhle, M. (1978). The Critical Period for Language Acquisition: Evidence from Second Language Learning. *Child Development*, *49*(4), 1114–1128.

Sorenson Duncan, T., & Paradis, J. (2020). How does maternal education influence the linguistic environment supporting bilingual language development in child second language learners of English? *International Journal of Bilingualism*, *24*(1), 46–61.

South, S., Haynie, D., & Bose, S. (2007). Student mobility and school dropout. *Social Science Research*, *36*(1), 68–94.

Spomer, M. L., & Cowen, E. L. (2001). A comparison of the school mental health referral profiles of young ESL and English-Speaking children. *Journal of Community Psychology*, *29*(1), 69–82.

Strand, S., Malmberg, L., & Hall, J. (2015). English as an additional language and educational achievement in England: An analysis of the National Pupil Database. Technical Report, Department of Education, University of Oxford.

Sun, H. (2019). Home Environment, Bilingual Preschooler's Receptive Mother Tongue Language Outcomes, and Social-Emotional and Behavioral Skills: One Stone for Two Birds? *Frontiers in Psychology*, *10*(1640), 1–13.

Sun, H., Yussof, N. T. B., Mohamed, M. B. B. H., Rahim, A. B., Bull, R., Cheung, M. W. L., & Cheong, S. A. (2021). Bilingual language experience and children's social-emotional and behavioral skills: a cross-sectional study of Singapore preschoolers. *International Journal of Bilingual Education and Bilingualism*, *24*(3), 324–339.

Taylor-Leech, K. (2013). Finding space for non-dominant languages in education: language policy and medium of instruction in Timor-Leste 2000–2012. *Current Issues in Language Planning*, *14*(1), 109–126.

Todd, P. E., & Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, *113*(485), 3–33.

Todd, P. E., & Wolpin, K. I. (2007). The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps. *Journal of Human Capital*, *1*(1), 91–136.

UNICEF (2018). An unfair start: Inequality in children's education in rich countries. *Innocenti Report Card 15*.

Werner, K., & Woessmann, L. (2021). The Legacy of Covid-19 in Education. Rationality and Competition Discussion Paper Series 291, CRC TRR 190 Rationality and Competition.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Yang, S., Yang, H., & Lust, B. (2011). Early childhood bilingualism leads to advances in executive attention: Dissociating culture and language. *Bilingualism: Language and Cognition*, *14*(3), 412–422.