UCD CENTRE FOR ECONOMIC RESEARCH

WORKING PAPER SERIES

2020

The Socioeconomic Gradient of Cognitive Test Scores: Evidence from Two Cohorts of Irish Children

David Madden, University College Dublin

WP20/20

June 2020

UCD SCHOOL OF ECONOMICS
UNIVERSITY COLLEGE DUBLIN
BELFIELD DUBLIN 4

The Socioeconomic Gradient of Cognitive Test Scores: Evidence from Two Cohorts of Irish Children

David Madden
(University College Dublin)

June 2020

Abstract: There is a well-established socioeconomic gradient in cognitive test scores for children. This gradient emerges at very early ages and there is also some evidence that it can widen as children age. We investigate this phenomenon with two longitudinal cohorts of Irish children who take such tests at ages ranging from 9 months to 17 years, using maternal education and equivalised income as our measure of socioeconomic resources. The gradient is observed from about 3 years and there is some tentative evidence that it widens as children get older. We have evidence on a wide range of tests and there is some evidence that the gradient is slightly stronger for tests involving crystalised as opposed to fluid intelligence. Exploiting the longitudinal nature of the data, we also investigate mobility across the distribution of test scores and there is some evidence that such mobility is less among poorer children raising the disturbing possibility that such children could become trapped in low achievement.

Keywords: Socioeconomics gradient; cognitive test score; achievement gap.

JEL Codes: I24, I30.

Corresponding Author: David Madden,

School of Economics,

University College Dublin,

Belfield,

Dublin 4,

Ireland.

Phone: 353-1-7168396

E-Mail: david.madden@ucd.ie

This research was not funded by any specific grant.

I am grateful to Kevin Denny and especially to Orla Doyle for comments on a very early draft. I remain responsible for any errors.

The Socioeconomic Gradient of Cognitive Achievement Test Scores: Evidence from Two Cohorts of Irish Children

1. Introduction

There is substantial evidence that the scores which children achieve in childhood cognitive achievement tests are systematically related to their socioeconomic status (SES).[1] These scores in turn are associated with important subsequent lifetime outcomes such as educational attainment, earnings and employment.[2] Moreover, this gradient by SES is visible from very early ages and there is some evidence that the gradient may become steeper as children age.[3] In some countries the gradient has steepened over time accompanied by rising income inequality (Duncan et al, 2019). The gradient of such scores thus can provide an important signal regarding future inequalities.

The ubiquity of this gradient, even in countries with well-developed fiscal and educational systems, and its persistence over time is a significant policy issue. Inequalities which develop very early on in life can become magnified implying that children can face a lifetime of disadvantage. In addition, such disadvantage can become transferred through generations, as children with parents with low educational attainment in turn have low attainment themselves and the pattern may then persist to subsequent generations (Black et al, 2005). It is interesting to note however that the historical role of education in promoting inter-class social mobility is disputed (see the discussion in Breen, 2019).

This paper explores the test score gradient in Ireland using various waves of the two cohorts of the *Growing Up in Ireland* (GUI) survey. This is a longitudinal data set which follows two cohorts, the Infant cohort born in the period December 2007-June 2008 and the Child Cohort

---

[1] For the rest of the paper we will refer to this phenomenon as the socioeconomic gradient of cognitive test scores. For the sake of brevity also we will simply use the term "test score" to refer to the variety of achievement tests which we will review and analyse.

[2] See for example Feinstein (2003), Dickerson and Popli (2016), Clotfelter et al. (2009) and the review by Currie and Almond (2011). For evidence for Ireland see Madden (2018) and Williams et al (2016).

[3] See Cunha and Heckman (2007) and Heckman and Mosso (2014).

born in the period November 1997-October 1998 (see Thornton et al, 2013 and Williams et al, 2009). GUI contains comprehensive information on a variety of cognitive tests and also has detailed data on SES. We thus have two longitudinal datasets with information on cognitive tests for children ranging in ages from 9 months to 17 years. As we explain below, while these are two distinct cohorts, we think results from the different cohorts are highly comparable and thus it is reasonable to provide analysis of the gradient over a long period of time.

We make four distinct contributions to the analysis of the socioeconomic gradient of test scores. First, as referred to above, we have data on the gradient from a total of seven different waves over the two datasets, ranging in age from 9 months to 17 years. This is an unusually long run of data for such a study and allows for a detailed analysis of how the gradient changes as children age. Secondly, we have data on a wide range of cognitive measures and so we can analyse the extent to which the gradient may be more pronounced for different sub-measures. Thirdly, the longitudinal nature of the data enables additional analysis of how the gradient changes. Since the gradient arises from the joint distribution of test scores and a measure of SES, changes in the gradient must arise from changes in either or both of the distribution of tests scores and SES. Longitudinal data allows us to decompose the overall change in the gradient into changes from these two sources. Finally, exploiting the longitudinal nature of the data we can analyse the extent to which children are mobile across the distribution of test scores. While the gradient reveals that poorer (richer) children have lower (higher) test scores, we may also be interested to know if it is the *same* poor/rich children who are at either end of the test score distribution or whether there is some degree of mobility.

Thus the purpose of this paper is primarily to explore in more detail the nature of this gradient in Ireland and how it evolves as children age. Our analysis does not examine potential pathways of the gradient nor do we analyse interventions designed to reduce the gradient.[4] In the next section of the paper we briefly review some of the international and Irish literature on the gradient. We then discuss our data and the different test scores we have over the different waves of the two cohorts of GUI. Since we are looking at children from ages 9 months to 17 years inevitably the specific measures will vary from wave to wave. The fact that the measures change has implications for the specific measures of socioeconomic gradient which we employ.

[4] We hope to address the former issue in future research. For a recent account of an intervention designed to improve cognitive and non-cognitive skills amongst disadvantaged Irish children see Doyle (2020).

Section 3 contains the results for the analysis of the gradient while section 4 presents results on the mobility of these achievement scores across waves of the GUI data. Section 5 provides concluding comments.

2. The Socioeconomic Gradient in Test Scores: A Brief Review of Background Literature

In this section we provide a brief review of the potential pathways via which the socioeconomic gradient can emerge. While we do not address any of these pathways directly in this paper they provide an essential background for our analysis.

We take as our starting point the family investment model of human capital production. Cognitive test scores can be seen as a measure of certain dimensions of human capital and hence a gradient in test scores reflects systematic achievement gaps in such human capital among children according to their SES. The production of such human capital for children is determined by the stock of human capital inherited from parents and then subsequent inputs of resources such as time and income (again for the most part provided by parents, up till early adulthood at least). Gaps which are observed may then reflect differential inheritance from parents, differential investments by parents and also presumably random shocks (which in turn might be related to parental SES e.g. if poorer children are more likely to experience adverse health shocks).

One of the most influential versions of this model is Cunha and Heckman (2007) whereby childhood endowments and parental investment are critical in the development of human capital. The build-up of human capital is sequential, with children mastering simple tasks before moving onto more complex ones. One of the key features of their model is what is known as *dynamic complementarity* whereby skills which are acquired in one period make investments in human capital in subsequent periods more productive. Intuitively it is simple to see how such a model can lead to a socioeconomic gradient which can become steeper over time. Children who are lucky enough to inherit a large stock of human capital will then find subsequent investments more productive. Since it seems plausible that children from well resourced families will on average inherit more human capital (either genetically or via material resources in the household) a gradient can emerge quite early and can steepen over time.[5] [6] While the basic nature-nurture distinction is now seen as overly simplistic it is still sometimes employed as a means of differentiating aspects of the parent-child human capital relationship. Bjorklund and Salvanes (2011) in their survey suggest that nature (pre-birth) factors and nurture (post birth) factors both account for about one third of the association between parent and child educational attainment, while also noting the likely interactions between nature and nurture.

Well-resourced and highly educated parents can also influence child human capital accumulation via parenting practice and other time investments and the basic human capital model thus been augmented to incorporate additional factors such as parenting skills, behaviours and beliefs (Doyle, 2020). There is evidence that lower SES parents engage in poorer parenting styles such as permissive or harsh parenting (Bradley and Corwyn, 2002). This can reflect both a lack of knowledge and also differential beliefs regarding parenting practice according to SES (Cunha et al, 2013).

The socioeconomic gradient of test scores and education is also consistent with the family stress model as outlined in Duncan et al (2019). In this model the principal pathway from low SES to poor test scores for children is via the high level of psychological stress experienced by parents in poor families (the model also acknowledges issues concerning lack of resources, credit constraints etc). Psychological stress can lead to the type of poor parenting practices referred to above which in turn can give rise to physiological stress responses from children which can harm development, including cognitive functioning (Duncan et al 2019). Poorer children may also suffer from environmental stress in the form of bad housing, sub-standard air quality and neighbourhood crime. If parents have limited cognitive capacity, then the demands imposed by poverty-related stress will reduce resources available for other tasks such as investment in the human capital of their children (Mani et al, 2013).

[5] There is also evidence that inheritability of IQ increases as children age which could also explain a steepening gradient (Davis et al, 2009).

[6] A recent paper by Belzil and Hansen (2020) suggests a decreasing role for family income in explaining educational transitions for US students at older ages, from 16 to 24.

So far in this brief run through the literature we have assumed that cognitive test scores are measuring a single dimension of intelligence. Of course there are different cognitive tests which measure different dimensions of intelligence. It may also be the case that the socioeconomic gradient may not be uniform across all dimensions of intelligence. Anger and Heineck (2010) examine intergenerational transmission of cognitive abilities between parents and children in Germany and distinguish between fluid intelligence, essentially related to innate abilities and problem solving and crystalised intelligence, related to accumulated learning (Cattell, 1963). They find significant transmission for both types of intelligence but tentative evidence that transmission is greater for crystalised intelligence. The wide range of cognitive tests in our dataset enables us to explore this issue in more detail.

Before moving on to discuss our data we briefly review evidence for Ireland. In the study most closely related to ours, Quigley and Nixon (2016) use data from waves 2 of both the Infant and Child cohort of GUI to regress test scores in verbal ability and verbal reasoning against a variety of variables including maternal education and family income. They find evidence of a socioeconomic gradient and also a role for variables such as access to books and parental reading, thus lending support for the family investment model of human capital of Cunha and Heckman (2007). However they do not analyse the gradient over a wider range of cognitive tests nor for a wider range of ages. In a similar study McNally et al (2019) also find roles for income, parenting practice and material resources (including books) in mediating the gradient between parental education and a vocabulary test score in wave 2 of GUI.

McMullin et al (2020) and McGinnity et al (2017) investigate the role of home learning activities (HLAs) and home learning environment (HLE) respectively in the relationship between social origins and cognitive development (as measured by vocabulary scores) for waves 2 and 3 of the Infant Cohort of GUI. In results which are echoed below, they confirm the existence of the social gradient but find relatively little independent role for HLAs and only a limited role for HLE, once controls for SES are included.

Madden (2018) used wave 1 of the GUI Child cohort to perform decomposition analysis of the gaps in test scores between 9 year olds differentiated by maternal education in reading and mathematics. Again, consistent with the family investment model he found that home rather than school characteristics played the dominant role in accounting for the gaps.

Much of the related work in Ireland has been on the socioeconomic gradient of educational attainment and access to third level education. Denny and Flannery (2017) review much of the work pertaining to access to third level education, pointing out that much of this is mediated by differential attainment in secondary school (Flannery and Cullinan, 2014, Denny, 2014). Other relevant contributions in this area include O'Connell et al (2006), Smyth and Hannon (2007) and the pioneering work of Clancy (1982). These studies all document significant and systematic differences in outcomes according to various measures of socioeconomic resources (e.g. income, social class, parental education).

To summarise this brief review of the background literature, the human capital/family investment model of Cunha and Heckman (2007) is consistent with evidence concerning the socioeconomic gradient in test scores. Parents with greater resources in terms of income and education, and reflected in their higher SES, can both bequeath and invest to a greater degree in the human capital of their children. These greater investments can come in the form of quantity (time and money) and also quality (superior parenting practices). Dynamic complementarities imply that initial gaps in cognitive attainments between children of different SES can become magnified over time. In addition, the family stress model provides alternative and complementary pathways whereby the socioeconomic gradient can emerge.

We now turn to explore this gradient in more detail for Ireland, starting off with a discussion of our data.

3. Data and Measures of Educational Outcome

Our data is four waves of the GUI Infant Cohort and three waves of the GUI Child Cohort. These data sets track the development of two cohorts of children born in Ireland, the Infant Cohort born in the period December 2007-June 2008 and the Child Cohort born in the period November 1997-October 1998 (see Thornton et al, 2013 and Williams et al, 2009).

The Infant Cohort consists of 11,000 children and the sampling frame used was the Child Benefit Register. This payment is made directly to the principal carer of the child (most typically the resident mother or step mother) and must be claimed within six months of the child being born, in the six months after the child becomes a member of the family or six months after the family become resident in Ireland.

The initial sample for the Infant Cohort consisted of just over 11,000 children. However the sample ultimately used for analysis is considerably smaller than this for a number of reasons. First, we choose to use a complete case balanced panel, hence we only include children for whom we have observations for all relevant variables at every wave. The exception is income, which is critical to our analysis. For the case of income we use conditional mean imputation to provide values for the missing observations. In addition we also drop children where data on the educational outcome, or those measures which are used to construct the educational outcome, are missing. As the education level of the primary caregiver is one of our measures of SES, we also drop observations where the primary caregiver changes between waves. After making these adjustments our working sample drops to 5668 (2767 male and 2901 female). With such a shrinkage of the sample there is clearly a danger of non-random attrition which can lead to bias as well as loss of precision. As outlined in McCrory et al (2013) wave to wave attrition in the GUI Infant Cohort is not random and hence in all the analysis we carry out we use the sampling weight for the latest available wave (wave 5) thus accounting for attrition from waves 1 to wave 5.

In the case of the Child Cohort the initial sample was 8568 children. The sampling frame of the data was the national primary school system, with 910 randomly selected schools participating in the study. Again we choose to work with a complete case balanced panel, consisting of only those children who were sampled in each of the three waves and imputed values where income is missing. We adopted the same procedure as with the Infant Cohort, dropping observations where the underlying educational data are missing and also where the primary caregiver changes between waves. This leaves us with an ultimate sample of 5326 (2585 male and 2741 female). As with the Infant Cohort we use wave 3 sampling weights to allow for non-random attrition.

We now turn to discuss the range of cognitive test scores available in each wave of the two cohorts of GUI. For each wave and cohort (with the exception of wave 4 of the Infant Cohort) children were given a range of tests. The range of these tests and whether they were administered by parents, GUI survey workers or by teachers in a classroom setting of course depended upon the age of the children. In some cases the tests were specifically related to the primary school curriculum (the Drumcondra reading and maths tests). With respect to the Infant Cohort, for the very youngest children, aged only 9 months, the tests were based on the Ages and Stages Questionnaire (Squires et al, 1997) and administered by the parents. For three year old children the tests were the British Ability Scales (BAS) Picture Similarities and

Naming Vocabulary tests (Elliot et al, 1996) and were administered by the GUI survey team. At five years of age the children again took the BAS Picture Similarities and Naming Vocabulary tests as well as a set of tests adapted from the UK Millenium Cohort covering Dispositions and Attitudes, Language for Communication and Thinking, Linking Sounds and Letters, Reading and Numeracy and again administered by survey workers from GUI. The final test we have for the Infant Cohort is the Drumcondra Reading Test which is a curriculum based test and was administered by teachers in the classroom.

Turning now to the older Child Cohort, the first set of tests administered to these children were the curriculum based Drumcondra Reading and Maths Tests again administered by teachers in the classroom. The availability of the Drumcondra Reading test for wave 5 of the Infant Cohort and wave 1 of the Child Cohort (in both cases aged 9) is very useful as it enables a comparison to be made between the two cohorts. If results for this common test are similar (as is the case as we will see below), then it provides some reassurance in terms of comparing results between the two cohorts. In wave 2 of the Child Cohort the tests administered were the Drumcondra Numerical and Verbal Ability tests. It should be noted that unlike the Drumcondra Reading and Maths tests, these are not curriculum based tests. The final set of tests we have is for wave 3 of the Child cohort. Three tests were carried out: a Cognitive Naming Test, a Cognitive Maths Test and Cognitive Vocabulary Test (details in Williams et al (2019). More details are available in appendix 1 and table 1 in that appendix also provides the rank correlations across the different subscales and components within each wave/cohort. In most cases these correlations are at least 0.3 and in some cases as high as 0.7, correlations which are comparable to Feinstein (2003).

In table 2 in appendix 1 we also show rank correlations for the composite measure across waves. We note that for the Infant cohort, with the exception of wave 1, rank correlations are around 0.3-0.45. They are higher for the Child cohort, up to 0.55 in some cases. It is not quite clear how to interpret these differing rank correlations. They may reflect the fact that the measures in the Child cohort are simply more homogenous or instead that there is less mobility across waves, or a combination of both factors. As we will see later it does seem to be the case that mobility is less in the Child cohort.

We thus have quite a wide range of cognitive test scores. We proceed by first of all by following Feinstein (2003) in using all the information available for each wave/cohort to construct a general measure via principal components analysis (PCA) and we carry out the bulk of our analysis on this measure. We then proceed to look at results for each individual test score.

PCA and the Composite Measure

PCA is the eigenvalue decomposition of the correlation matrix *R* of the different individual test score measures available in each wave/cohort. If we have, say, *n* measures, $x_1 \dots x_n$ then the first principal component, y_1 is given by

$$y_1 = a_{11}x_1 + a_{12}x_2 \dots + a_{1n}x_n$$

where a_{1i} are the weights which are chosen to maximise the variance of y_1 and must also satisfy the normalising constraint $\sum_{i=1}^{n} a_{1i}^2 = 1$.

Using the first principal component has the advantage of combining information from the different sources. As noted above, the rank correlations across the different measures seem to be sufficiently high to be confident that we are picking up a similar underlying process. In the analysis which follows we will refer to this as our *composite* measure, to distinguish it from the individual sub-components.

In appendix 1 we show the scree plots for the PCA. Using the rule of thumb that components where the eigen value exceeds unity should be selected we see that in nearly all cases it is only the first principal component which satisfies this condition. Table 3 appendix 1 also shows the fraction of variance explained by the first principal component. Where we have only two measures entering into the PCA then the first component explains about 75-80% of variance. When there are more measures e.g. Infant Cohort wave1 and wave 3, then the fraction of variance explained falls to 40-50%. In all instances the value of Kaiser-Meyer-Olkin test statistic for sampling adequacy for PCA meets the rule of thumb threshold of 0.5, though in some cases only barely.

The analysis which follows is rank-based i.e. looking at the gradient between the child's rank in the different cognitive measures and rank in terms of our SES measure. The idea here is that while the composite measure is not strictly the same from wave to wave (since the individual sub-components change), nevertheless a rank-based measure is hopefully less sensitive to the wave to wave variation in measure.

We now discuss our choice of measure of SES. In this paper we will use two measures. The first is education of the principal carer (since this is the mother in around 99% of cases we will refer to it as maternal education). This choice is motivated by our belief that it is likely to be more accurately measured than other contending variables such as social class. In addition, it seems likely that critical decisions regarding the child's educational habits and practises will either be taken by or at least heavily influenced by the mother and the highest educational level achieved by the mother is likely to affect these decisions.[7] [8]

We thus break down education into four categories: (1) up to and including completion of lower secondary schooling (2) completion of all secondary schooling (3) obtaining a post-secondary school diploma or cert and (4) completion of third level education.

While it seems plausible that a gradient would be observed with respect to maternal education, one disadvantage of such a measure is that with only four categories there will be many ties in rank of SES. Thus in addition we also present results using equivalized income. This is calculated via the answer to questions on total net household income from all members and sources after deductions for income tax and social insurance. If households cannot give an exact figure then they answer a sequence of questions where they are presented with cards where they select the range into which they believe that family income falls. As mentioned above our measures of gradient are all rank based. It seems reasonable to assume that even if households exact level of income is not measured with 100% accuracy that the *ranking* of households by income will be less prone to error.

We use *current* equivalized income as measured at each wave. One potential issue with this approach is that each single period measure of income may contain measurement error which in turn could attenuate the correlation with test scores. Taking the average measure of income over the waves for each cohort could provide a better approximation to permanent income (Rothstein and Wozny, 2013). The use of permanent income however would prevent us from carrying out some of the decomposition analysis below, in particular the extent to which changes in the socioeconomic gradient arise owing to changes in the distribution of test scores or changes in the distribution of income. Also it is arguable that such measurement error might

---

[7] Anger and Heineck (2010) find that in terms of intergenerational transmission of intelligence that maternal education appears to be more important that paternal education.

[8] Bukodi and Goldthorpe (2013) discuss how the relationship between educational attainment and social origins can differ according to the specific measure of social origin adopted. Their comparison is between parental class, parental social status and parental education. GUI does not have information upon parental class and in any event it is arguable that it would be measured with less accuracy than education.

be less problematic when employing rank correlations. In the analysis which follows we check the sensitivity of our results to whether we use current or permanent income.

Socioeconomic Gradient of the Composite Measure

We commence with some graphical analysis. Figures 1a-1h provides the cumulative distribution functions for the fractional rank of our composite measure for all waves/cohorts, stratified by maternal education. The first three graphs show the CDF for waves 1 to 3 of the Infant Cohort for the composite measure. The fourth and fifth graphs show the CDF for the Drumcondra Reading test for wave 5 of the Infant Cohort and wave 1 of the Child Cohort respectively, while the final three graphs provide the CDFs for the composite measure for the three waves of the Child Cohort.

The breakdown by maternal education is critical here, since if we just presented this graph for the sample as a whole then it would be a straight line with a slope of 1 (since we are effectively plotting rank against rank). If there was no difference in outcomes by maternal education then the CDFs would essentially be superimposed on each other. Differences in the CDF by maternal education essentially reflect the gap in achievement – the wider the gap, the greater is the gradient by maternal education.

If we look at figure 1a we see that while not exactly superimposed upon each other, the CDFs are very close.[9] Thus effectively there is no socioeconomic gradient for the measure for three month old infants. However, when we go to figure 1b which shows the CDFs for three years, the CDFS are clearly some distance apart. To interpret this, take for example the point corresponding to 0.5 on the horizontal axis. If we were just looking at the sample as a whole then the corresponding point on the vertical axis would also be 0.5. However comparing the CDFs for education levels 4 and 1, for example, we see that the vertical axis value for education level 1 is around 0.65, while that for education level 4 is about 0.45. This tells us that around 65% of children whose mothers have educational level 1 would be ranked in the bottom half of the overall distribution of test scores, but only about 45% of those in education level 4 would be so ranked. Thus those children whose mothers have the lowest education levels are disproportionately concentrated in the lower half of the distribution. Since the CDF for

[9] In the interests of visual clarity in figures 1a-1h we do not show the associated confidence intervals (available on request) but in the case of the measure for wave 1 Infant they clearly overlap indicating that we cannot reject the null that the CDFs are equal.

education level 4 lies below that for education level 1 for all values on the horizontal axis this implies that regardless of which overall fractional rank of test scores we choose there will always be a higher fraction from education level 1 below that level of achievement than education level 4. Thus education level 4 in a sense stochastically dominates education level 1.[10] While we do not show the confidence intervals for the CDFs (available on request), apart from some overlap between those for levels 3 and 4, they are sufficiently far apart that we can reject the null of equality.

Moving on through the graphs from figure 1c to figure 1h, in all cases the gaps between the CDFs by maternal education reveal the existence of the gradient. There seems to be some visual evidence that the gaps widen as children get older (we investigate this below when we look at rank correlations) but bear in mind that the underlying measures are not always directly comparable e.g. figure 1d shows the CDFs for reading only (as this was the only measure available for wave 5 of the Infant cohort). One interesting comparison is between figures 1d and 1e. Here we have the same measure for the same age but for a different cohort (the Child cohort measured in 2008 and the Infant cohort measured in 2018). The graphs look remarkably similar offering some support to the idea that we can make comparisons between the two cohorts.

However eyeballing is not always a reliable guide when making comparisons and so in table 1 we present the Spearman rank correlations between the composite measure and maternal education and we also present the rank correlations using equivalized income as the ranking SES variable. We choose to use the rank correlation coefficient as the measure of gradient rather than the concentration index (CI) often favoured by health economists when calculating the socioeconomic gradient for illness. In our application here, the expression for the CI would be $CI = \frac{2\,Cov(x_i, r_i)}{\bar{x}}$ where x_i represents the test score for child i, r_i is the fractional rank of the ranking socioeconomic variable and $\bar{x}$ is the mean of the test score. This measure however will be sensitive to precisely how the test score is measured and as we have seen this varies from wave to wave. The rank correlation coefficient is independent of the distribution of the underlying variables and so seems to be preferable in this case.

[10] Typically if we were searching for stochastic dominance then the horizontal axis would be the support of the distribution. However since the values of the support differ from wave to wave, owing to the range of measures we use, we choose to use instead the fractional rank for the complete sample.

The results from table 1 suggest that the gradient does increase with age. Regardless of whether education or income is used as the ranking SES variable, we see that between ages 3 and 5 the rank correlation is around 0.16-0.18. It increases to around 0.24 for nine year olds. Note that the correlation here is for reading only, and it is quite possible that the correlation for this specific domain is higher than for the more general measure anyway, and that we are not observing an increase in the gradient. However, when we look at the gradient for the composite measure for nine year olds (this time from the Child cohort) it is in the range 0.21-0.24 and the correlation for 13 and 17 year olds has increased to 0.22-0.27. In making these comparisons at all time we must bear in mind that the actual measures themselves, or the underlying measures from which the composite principal component is derived, differ from wave to wave. We also note that when we use permanent income as the SES ranking variable (the average of income over the waves for each cohort) the correlation tends to be higher.

Note also that the rank correlations for Drumcondra reading for wave 5 of the Infant cohort and wave 1 of the Child cohort are practically identical when education is used as the SES and very similar when income is used. This provides further reassurance that it is legitimate to compare results across the two cohorts.

The Evolution of the Gradient of the Composite Measure Over Time

Figure 2a provides another perspective on the evolution of the gradient with respect to maternal education. Again we partition our sample into four groups based upon maternal education. We then show on the vertical axis the average fractional rank for the composite measure for each education group. If there was no gradient, each group would have an average rank of 0.5. However the existence of a gradient implies that the average fractional rank by maternal education will differ and that this difference will be statistically significant. Figure 3a shows how these average ranks vary across wave and cohort. For the Infant cohort wave 1, when children are nine months old, there is no difference in rank by maternal education. However by 36 months with the Infant cohort wave 2 a statistically significant difference is clearly evident, in particular between education levels 1 and 2 and levels 3 and 4 and this gradient shows signs of widening as children get older, though it appears to remain reasonably stable throughout the Child cohort.

One of the advantages of this graph is that it gives us a clearer idea of how the gradient is developing. As children age from 9 months to 9 years the gradient primarily emerges owing to the relative decline of those children whose mothers have not completed secondary education and to a lesser extent for children whose mothers do complete second level education, but who do not obtain any more education. Children whose mothers have a Diploma/Cert (i.e. some post-secondary school education but not university) also show a slight relative decline and there is a clear improvement for children whose mothers have third level education.

Moving now to the GUI Child cohort, we see that the gaps which have emerged by age 9 appear to stabilize with virtually no statistically significant change as children age up to 17. It is interesting to note that the gaps remain stable even though the sample of children changes from the Infant to the Child cohort.

Similar to the decompositions which we discuss below of the change in the rank correlation coefficient, we note that changes in figure 2a can arise for either of two reasons. It may be that the composite measure scores for children of given maternal education change over time, or it may also be the case that maternal education itself can change, so that children with given scores now have different maternal education.

In order to get some idea of the relative importance of these factors figure 2b reproduces figure 2a except that now we "freeze" children at the level of maternal education in wave 1 of the respective cohorts. There seems to be some evidence that the gaps widen slightly as we move through the Child cohort but overall the story is consistent with the decomposition carried out in table 2, whereby the bulk of the emerging socioeconomic gradient arises from a relative deterioration in the scores for children of less well educated mothers.

The Decomposition of the Change in the Gradient Over Time

Given the availability of longitudinal data it is possible to dig a little deeper into how these correlations have changed over time. The rank correlation is a statistic arising from the joint distribution of whatever achievement measure is used and SES. Thus any change in this over time must arise from changes in the distribution of the score and/or SES. In their analysis of how the socioeconomic gradient of health changes over time Allanson et al (2010) show that a change in the concentration index can arise owing to changes in the health outcome conditional on a given distribution of income (or whatever ranking variable is used), or changes in the income variable conditional on a given distribution of health, or a combination of both these factors.

A similar decomposition can be applied to changes in the rank correlation coefficient between our composite score (x) and SES. Thus suppose we are looking at changes between waves 1 and 2 we have: $r^{2,2}_{x,SES} - r^{1,1}_{x,SES} = \left(r^{2,2}_{x,SES} - r^{2,1}_{x,SES}\right) + (r^{2,1}_{x,SES} - r^{1,1}_{x,SES})$. In this expression the term $r^{2,1}_{x,SES}$ is the rank correlation coefficient between the wave 2 score and wave 1 SES.

Thus the first term in brackets on the right-hand side of the above expression shows the change in correlation arising from changes in the SES ranking conditional on a given ranking of the composite score (that of wave 2) – in table 2 we label this "Change in SES". The second term gives us the change arising from a change in the ranking of the score conditional upon a given ranking of SES (wave 1 SES) and we label this "Change in Score".[11]

Table 2 provides this decomposition (and figures 2a-2d provide the same information in graphical form). In order to make sense of the numbers let's take as an example the change in the rank correlation coefficient for the Child cohort between waves 2 and 1, where income is the SES measure. Overall the rank correlation increases by 0.0291. The term "Change in SES" however shows a fall of 0.0193. Thus if there had been no change in the rank of the score by child and we were merely looking at the change in gradient arising from changes in income ranks, then the correlation would fallen by 0.0193. Or in other words, incomes for children with (relatively) poorer composite scores have improved slightly (or incomes for children with relatively better composite scores have deteriorated slightly) and these forces act to reduce the gradient. The second term, "Change in Score" thus accounts for an increase in the rank correlation of 0.0607. Thus it is the (relative) deterioration of the score for poorer children which is the overwhelming driver of the increase in the socioeconomic gradient.

Table 2 and figures 3a-3d reveal that most wave to wave changes show an increase in the gradient and in all these cases the principal driver is a deterioration in the score for poorer children. The only exception again is between waves 2 and 3 for the Infant cohort and here the absolute changes in the gradient are quite small.

We also note that while term 1, the part of the decomposition arising from a change in the distribution of the SES measure, typically plays a very minor role, it tends to be relatively more important when the SES measure is income rather than maternal education. This presumably reflects the fact that maternal education tends to be much more stable over time than income.

[11] Of course as in any "index number" type issue we could also add and subtract $r^{1,2}_{m,SES}$ which is the rank correlation between the measure in wave 1 and SES in wave 2. The precise decomposition differs but the proportional contribution of each component is very similar.

Overall, these results seem quite intuitive. Given an increase in the gradient, it seems more plausible that this arises from a relative deterioration in the composite test score for poorer children rather than a relatively deteriorating economic situation for children with poor composite test scores. Poor test scores for children are unlikely to worsen the family economic situation as children effectively contribute little to that situation. On the other hand a declining economic situation could adversely impact testt scores for children via any or all of the pathways discussed in section 2 of this paper.

Results for the Individual Test Scores

The results presented above all refer to the socioeconomic gradient for the composite score derived from PCA. However, one of the contributions of this paper is to provide evidence on the gradient for the individual achievement scores. As discussed in section 2 it is possible that the gradient might differ for different individual cognitive measures, in particular the distinction between fluid and crystalised measures of intelligence pointed out by Anger and Heineck (2010).

Rather than reproduce the graphs of the CDFs by maternal education instead in table 3a-3b we reproduce table 1 for each individual cognitive measure. The results provide tentative support for the findings of Anger and Heineck. Going down each column we can try to locate those individual measures which seem to correspond most closely to fluid or crystalised intelligence. For wave 1 of the Infant cohort we see that the highest absolute rank correlation is for ASQ Communication, but here the gradient is not in the "expected" direction. Other rank correlations for wave 1 are all quite small in absolute size and not all are statistically significant. Given the combination of positive and negative correlations we can see how the rank correlation for the overall composite score was low in absolute size and not statistically significant. Wave 1 scores are measured at home, so it is possible that measurement error is an issue here.

For the other waves of the Infant cohort, the results are mixed. The gradient is higher for vocabulary as opposed to naming for wave 2, yet for wave 3 the correlations are very similar. For the other wave 3 individual measures we see the highest correlation for language, yet no real difference between the correlations for reading and mathematics.

By the time we come to the Child cohort there does seem to be a distinction between reading and language based tests compared to numerical/mathematics based tests and the results are consistent with the Anger and Heineck (2010) findings. For all three waves of the child cohort, when we have two or more tests carried out at the same time, the rank correlation is higher for the test which seems to more closely correspond to what we would regard as crystalised intelligence.

Carrying out the decomposition of the change in individual tests score is more difficult as the measures differ from year to year and we really would be comparing apples with oranges. The only consecutive waves where exactly the same test was used is between waves 2 and 3 of the Infant Cohort where the tests are the BAS Picture Similarities and Naming Vocabulary Tests. We show the results for these decompositions in table 4. Again, for the bulk of the wave to wave changes the greater part arises from a change in the distribution of the test score rather than from the distribution of the SES measure. The only exception here is for the change in the rank correlation for picture similarities when income is the SES measure. However the absolute change here is the smallest of all the wave to wave changes so we would not be inclined to read too much into this.

4. Mobility Across the Measures

We now turn to examine mobility across the waves for composite test scores. Mobility per se is of interest and we will also examine the extent to which mobility can impact upon the socioeconomic gradient. Clearly this can in principle work in either direction. Greater upward mobility for children from low SES backgrounds and/or greater downward mobility for children from high SES backgrounds can reduce the gradient, and of course should these forces work in opposite directions then the gradient can increase. We start off with some graphical analysis, followed by some statistical analysis.

The primary graphical approaches we employ are the plots of local linear kernel regressions of composite score in period t against score in period t-1 (using the Epanechnikov kernel and rule of thumb bandwidth). Such plots were used by Black et al (2015) in their analysis of *inter*-generational mobility in wealth in Sweden although here we apply them to *intra*-generational test scores. To the best of our knowledge, they have not been used before to analyse mobility across educational outcomes but they are a useful graphical tool to provide an insight into such mobility.

Figures 4a-4f show these plots for all pairwise wave-to-wave comparisons in the Infant cohort while figures 5a-5c provides the plots for the child cohort. In all of these plots we see an upward sloping relationship revealing persistence in the data in that rank in period t is a good predictor of rank in period t+1. It is noticeable however that pairwise ranks involving wave 1 of the Infant cohort show a weaker relationship. This is consistent with the generally non-existent socioeconomic gradient for this wave. By wave 2 of the Infant Cohort a gradient has been established but rank in wave 1 in only a weak predictor of rank in subsequent periods. In contrast, when we look at the plots for pairwise comparisons between waves 2, 3 and 5 of the Infant Cohort and for all pairwise comparisons in the Child Cohort we see a stronger and well-determined (in the sense of narrow confidence intervals) relationship between each wave. For the most part this relationship seems linear, but there is some evidence in the Infant Cohort that the slope of the relationship and hence persistence is stronger (and mobility weaker) at lower ranks.

Tables 5a and 5b provides the results from linear rank-rank regressions of the form $r_{t,i} = \alpha + \beta r_{t-1,i} + \varepsilon_i$ where $r_{t,i}$ represents the rank of person i in period t for the composite test score, and the β coefficient gives the rank-rank slope. There are a number of features of the results worth noting. First of all, looking at results from the child cohort, the rank-rank slopes when wave 1 is the "base" rank are clearly lower with values around 0.1, compared to values for the other regressions at around 0.3. This reflects the results from the plots of the local polynomial regressions that the degree of persistence from wave 1 was quite limited or to look at it another way, there was a degree of churning in fractional rank following wave 1.

Even allowing for the fact that the measures underlying the fractional ranks differ from wave to wave, it appears as though persistence is increasing and hence mobility decreasing as children become older. Hence the slopes are around 0.3 between waves 2 and 3 of the Infant Cohort increasing to 0.43 between waves 3 and 5. Then moving onto the Child Cohort we see slopes in the region of 0.6. These results suggest that interventions to try to reduce a socioeconomic gradient in educational outcomes might be best introduced when children are young, since by the time they reach adolescence mobility as measured by rank-rank slopes has reduced considerably.

What about the possibility that mobility is lower amongst lower ranked children, as suggested in the visual inspection of the local polynomial plots? To investigate this we estimate a modified version of the rank-rank relationship $r_{t,i} = \alpha + \beta r_{t-1,i} + \gamma r_{t-1,i} * I_{p_{t-1}<0.2} + \varepsilon_i$ where $I_{p_{t-1}<0.2}$ is a dummy variable which takes on a value of 1 if an observation has a fractional rank in the composite score of less than 0.2. The significance of the γ coefficient then indicates if the slope is higher for the lowest 20% in wave t-1. These results are presented in tables 6a and 6b for the Infant and Child cohorts respectively. They show that there is no evidence of less mobility at lower ranks for rank-rank regressions between waves 1, 2 and 3 for the Infant or the Child Cohort. It is only for the Infant Cohort when the rank-rank regression is between waves 5 and waves 2 and 3 that we observe such a phenomenon. In these cases though there seems to be quite a sizeable difference in the rank-rank slope for the lowest 20%. For the top 80% the rank-rank slope ranges from 0.334 to 0.437. The rank-rank slope for the bottom 20% is about 0.17 higher i.e. almost half as big again.

A further insight on mobility can be gained from looking at the summary mobility indices derived from transition matrices. In tables 1a-1f and 2a-2c in Appendix 2 we look at a sequence of wave to wave transition matrices of fractional rank of composite test score by quintile for the Infant and Child cohorts respectively. For example in table 1a, the top left entry in the matrix is 0.27. This reveals that of the population who were originally in the bottom quintile of the wave 1 outcome, 27% of that group stayed in this quintile, 21% moved up the next highest quintile, while 15% moved all the way up to the highest quintile. A lack of mobility is reflected in high values along the main diagonal, indicating that most people stayed in the same quintile. Thus a summary measure of mobility which has been suggested by Shorrocks (1978) is $\frac{m-Tr(M)}{m-1}$ where *m* refers to the dimensionality of the transition matrix (5, in this case) and *Tr(M)* is the trace of the transition matrix, *M*. This provides an index whose lower bound is clearly zero (since the proportion along the main diagonal would be unity for each quintile) and whose upper bound is $\frac{m}{m-1}$, since in this case the entry for each element along the main diagonal is zero.

Tables 7a and 7b provide the Shorrocks index for the Infant and Child cohorts respectively. We see that for children mobility is greatest between waves 1 and subsequent waves. Mobility between waves other than wave 1 is more limited. Table 7b shows that mobility overall for the Child cohort is lower than for the Infant cohort, again entirely consistent with results from the rank regressions and the local polynomial plots.

In tables 8a-8b we present the same results except this time by maternal education. Overall, the indices are typically quite close together so that in many comparisons the difference by educational level is not statistically significant. For the Infant Cohort however, it is always the case that the index for educational level 1 (only completed primary or lower secondary education) is less than for other educational levels and in many cases this difference is statistically significant. That pattern is not quite so pronounced for the Child cohort except for the wave 2 – wave 3 transitions.

What implications do the mobility results have for the socioeconomic gradient? The results discussed above provide some evidence that mobility in test scores is less among children with lower SES ranks. As we discussed in the introduction, the existence of the socioeconomic gradient is a significant policy challenge. This challenge arguably becomes even more acute if, as the evidence suggests here to some extent at least, it is the *same* children over time who are in a condition of low test score achievement and low SES. Results from Doyle (2020) indicates that interventions via improvements in parenting skills can raise IQ and the presence of children who are trapped in a low achievement-low SES situation makes the arguments for such interventions more compelling.

5. Conclusions

This paper has provided further evidence on the socioeconomic gradient across cognitive test scores for two cohorts of children ranging in age from 9 months to 17 years. Our principal contribution compared to earlier work in the area is that we present results for a wider range of ages, from well before formal schooling up until just before the completion of second level schooling. We also present results for a wide range of test scores, scores which measure different dimensions of intelligence as well as evidence on a composite measure derived from PCA. The longitudinal nature of our data permits two further innovations: we can decompose changes in the gradient into changes arising from changes in the distribution of test scores (conditional upon a given distribution of SES) and changes in the distribution of SES (conditional upon a given distribution of test scores). We can also explicitly examine mobility along the distribution of test scores across waves of GUI data.

Our results indicate that a clear gradient for the composite test score sets in by 3 years. The gradient does not appear to change by 5 years, but by the time we reach 9 years it has widened. This widening appears to continue with respect to maternal education on into adolescence when we analyse the second cohort, but not so much when equivalised income is used as the measure of SES. Thus evidence of a steepening gradient is only tentative as comparisons between the two cohorts should be made with care.

In terms of the range of different test scores which we analyse, the results are similar to those for the composite score with tentative evidence that the gradient is stronger for language and vocabulary based measures which are related to crystalised intelligence as opposed to numerical type measures which are more closely related to fluid based intelligence.

In terms of the breakdown of the change in the gradient our analysis here shows that the vast bulk of the change in gradient arises from changes in the distribution of test scores conditional upon a given distribution of SES as opposed to the opposite. This seems plausible given (a) that a measure of SES such as maternal education is unlikely to show great change and (b) it seems more likely that changes in a family's resources would impact upon child test scores rather than vice versa.

Finally, we provide some analysis of mobility in terms of children moving up and down in rank by test scores. Mobility appears to be greater in the younger cohort of children. There is also tentative evidence that within this cohort, mobility is less among the lower ranked children. This raises the disturbing possibility that children who are already educationally disadvantaged may become trapped in this situation.

Table 1: Spearman Rank Correlations

	Education	Income	Income (permanent)
Wave 1, Infant	-0.0141	-0.0137	-0.0221
Wave 2, Infant	0.1595	0.1695	0.1807
Wave 3, Infant	0.1625	0.1525	0.1731
Wave 5, Infant (reading only)	0.2448	0.2390	0.2492
Wave 1, Child (reading only)	0.2433	0.2195	0.2533
Wave 1, Child	0.2428	0.2297	0.2606
Wave 2, Child	0.2958	0.2588	0.2945
Wave 3, Child	0.2705	0.2180	0.2568

Table 2: Decomposition of Change in Rank Correlations

	Infant Cohort					
	Maternal Education			**Income**		
	Total Change	**Change in SES**	**Change in Scores**	**Total Change**	**Change in SES**	**Change in Scores**
W3-W1	0.1766	0.0100	0.1666	0.1662	-0.0079	0.1741
W3-W2	0.0030	0.0017	0.0013	-0.0170	-0.0119	-0.0051
W2-W1	0.1736	0.0073	0.1663	0.1832	0.0023	0.1809
	Child Cohort					
	Maternal Education			**Income**		
	Total Change	**Change in SES**	**Change in Scores**	**Total Change**	**Change in SES**	**Change in Scores**
W3-W1	0.0277	0.0007	0.0270	-0.0117	-0.0071	-0.0046
W3-W2	-0.0253	0.0051	-0.0304	-0.0408	-0.0127	-0.0281
W2-W1	0.0530	-0.0025	0.0555	0.0291	-0.0193	0.0617

Table 3a: Rank Correlation Coefficients – Maternal Education

	Infant Cohort				Child Cohort		
	W1	**W2**	**W3**	**W5**	**W1**	**W2**	**W3**
ASQ Communication	-0.1397						
ASQ Gross Motor	0.0307						
ASQ Fine Motor	0.0507						
ASQ Problem Solving	0.0034						
ASQ Personal Social	0.0065						
BAS Picture Similarities		0.1155	0.1022				
BAS Naming Vocabulary		0.1545	0.1120				
Language			0.1401				
Linking			0.1246				
Reading			0.1040				
Numbers			0.1163				
Drumcondra Reading				0.2448	0.2433		
Drumcondra Maths					0.1941		
Drumcondra Numerical						0.2539	
Drumcondra Verbal						0.2710	
Cognitive Naming							0.1622
Cognitive Maths							0.1990
Cognitive Vocab							0.2337

Table 3b: Rank Correlation Coefficients – Income

	Infant Cohort				Child Cohort		
	W1	W2	W3	W5	W1	W2	W3
ASQ Communication	-0.1145						
ASQ Gross Motor	-0.0279						
ASQ Fine Motor	0.0228						
ASQ Problem Solving	0.0294						
ASQ Personal Social	0.0335						
BAS Picture Similarities		0.0963	0.1092				
BAS Naming Vocabulary		0.1796	0.1380				
Language			0.1576				
Linking			0.0915				
Reading			0.0842				
Numbers			0.0919				
Drumcondra Reading				0.2390	0.2195		
Drumcondra Maths					0.1903		
Drumcondra Numerical						0.2140	
Drumcondra Verbal						0.2456	
Cognitive Naming							0.1222
Cognitive Maths							0.1636
Cognitive Vocab							0.1898

Table 4: Decomposition of Change in Rank Correlations

	Picture Similarities					
	Maternal Education			Income		
	Total Change	Change in SES	Change in Scores	Total Change	Change in SES	Change in Scores
W3-W2	-0.0133	-0.0042	-0.0091	0.0129	0.0116	0.0013
	Naming Vocabulary					
	Maternal Education			**Income**		
	Total Change	**Change in SES**	**Change in Scores**	**Total Change**	**Change in SES**	**Change in Scores**
W3-W2	-0.0425	-0.0021	-0.0404	-0.0416	-0.0087	-0.0329

Table 5a: Rank-rank regression slopes – Infant Cohort

	Wave 2-Wave 1	Wave 3 – Wave 1	Wave 3 - Wave 2	Wave 5 – Wave 1	Wave 5 – Wave 2	Wave 5 – Wave 3
β	0.130*** (0.013)	0.101*** (0.013)	0.317*** (0.013)	0.038*** (0.013)	0.334*** (0.013)	0.437*** (0.012)

*** p<0.01, ** p<0.05, * p<0.1

Table 5b: Rank-rank regression slopes – Child Cohort

	Wave 2 – Wave 1	Wave 3 – Wave 1	Wave 3 – Wave 2
β	0.681*** (0.010)	0.569*** (0.011)	0.692*** (0.010)

*** p<0.01, ** p<0.05, * p<0.1

Table 6a: Rank-rank regression slopes – lowest 20% coefficient – Infant Cohort

	Wave 2-Wave 1	Wave 3 – Wave 1	Wave 3 - Wave 2	Wave 5 – Wave 1	Wave 5 – Wave 2	Wave 5 – Wave 3
γ	0.155 (0.094)	0.095 (0.099)	0.119 (0.09)	0.088 (0.095)	0.173* (0.089)	0.171** (0.086)

*** p<0.01, ** p<0.05, * p<0.1

Table 6b: Rank-rank regression slopes – lowest 20% coefficient – Child Cohort

	Wave 2 – Wave 1	Wave 3 – Wave 1	Wave 3 – Wave 2
γ	0.049 (0.072)	0.066 (0.081)	0.084 (0.071)

*** p<0.01, ** p<0.05, * p<0.1

Table 7a: Shorrocks Mobility Index - Infant Cohort

Wave 2-Wave 1	Wave 3 – Wave 1	Wave 3 - Wave 2	Wave 5 – Wave 1	Wave 5 – Wave 2	Wave 5 – Wave 3
0.950 (.007)	0.975 (.007)	0.893 (.007)	0.995 (.007)	0.896 (.007)	0.845 (.008)

Table 7b: Shorrocks Mobility Index – Child Cohort

Wave 2 – Wave 1	Wave 3 – Wave 1	Wave 3 – Wave 2
0.742 (.008)	0.812 (.008)	0.742 (.008)

Table 8a: Shorrocks Mobility Index by Maternal Education - Infant Cohort

	Wave 2-Wave 1	Wave 3 – Wave 1	Wave 3 - Wave 2	Wave 5 – Wave 1	Wave 5 – Wave 2	Wave 5 – Wave 3
Level 1	0.921 (0.024)	0.971 (0.028)	0.882 (0.028)	0.975 (0.022)	0.874 (0.029)	0.812 (0.032)
Level 2	0.971 (0.013)	0.985 (0.012)	0.904 (0.014)	1.013 (0.012)	0.925 (0.014)	0.856 (0.014)
Level 3	0.948 (.015)	0.972 (0.015)	0.918 (0.014)	0.995 (0.015)	0.913 (0.015)	0.865 (0.016)
Level 4	0.951 (0.011)	0.970 (0.011)	0.911 (0.011)	0.986 (0.011)	0.910 (0.011)	0.887 (0.012)

Table 8b: Shorrocks Mobility Index by Maternal Education - Child Cohort

	Wave 2-Wave 1	Wave 3 – Wave 1	Wave 3 -Wave 2
Level 1	0.770 (0.023)	0.845 (0.022)	0.699 (0.029)
Level 2	0.751 (0.015)	0.815 (0.014)	0.767 (0.015)
Level 3	0.767 (0.016)	0.837 (0.016)	0.787 (0.016)
Level 4	0.760 (0.017)	0.828 (0.016)	0.739 (0.016)

References:

Allanson, P., U. Gerdtham, & D. Petrie, (2010). Longitudinal analysis of income-related health inequality. *Journal of Health Economics*, Vol. 29, pp. 78-86.

Anger, S. and Heineck, G., (2010). Do smart parents raise smart children? The intergenerational transmission of cognitive abilities. *Journal of Population Economics*, *23*(3), pp.1105-1132.

Belzil, C. and Hansen, J., (2020). The Evolution of the US Family Income-Schooling Relationship and Educational Selectivity. *Journal of Applied Econometrics*.

Björklund, A. and Salvanes, K.G., (2011). Education and family background: Mechanisms and policies. In *Handbook of the Economics of Education* (Vol. 3, pp. 201-247). Elsevier.

Black, S.E., Devereux, P.J. and Salvanes, K.G., (2005). Why the apple doesn't fall far: Understanding intergenerational transmission of human capital. *American economic review*, *95*(1), pp.437-449.

Black, S., P. Devereux, P. Lundborg, and K. Majlesi (2015) "Poor Little Rich Kids? The Determinants of the Intergenerational Transmission of Wealth1." (2015).

Bradley, R.H. and Corwyn, R.F., (2002). Socioeconomic status and child development. *Annual review of psychology*, *53*(1), pp.371-399.

Breen, R., (2019). Education and intergenerational social mobility in the US and four European countries. *Oxford Review of Economic Policy*, *35*(3), pp.445-466.

Bukodi, E. and Goldthorpe, J.H., (2013). Decomposing ‘social origins’: The effects of parents’ class, status, and education on the educational attainment of their children. *European Sociological Review*, *29*(5), pp.1024-1039.

Cattell, R.B., (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of educational psychology*, *54*(1), p.1.

Clancy, P. (1982). *Participation in higher education: a national survey*. Higher Education Authority.

Clotfelter, C., H. Ladd and J. Vigdor (2009): The Academic Achievement Gap in Grades 3 to 8. *Review of Economics and Statistics*. Vol. 91, pp. 398-419.

Cunha, F. and Heckman, J., (2007). The technology of skill formation. *American Economic Review*, *97*(2), pp.31-47.

Cunha, F., Elo, I. and Culhane, J., (2013). *Eliciting maternal expectations about the technology of cognitive skill formation* (No. w19144). National Bureau of Economic Research.

Currie, J. and Almond, D., 2011. Human capital development before age five. In *Handbook of labor economics* (Vol. 4, pp. 1315-1486). Elsevier.

Davis, O.S., Haworth, C.M. and Plomin, R., (2009). Dramatic increase in heritability of cognitive development from early to middle childhood: An 8-year longitudinal study of 8,700 pairs of twins. *Psychological Science*, *20*(10), pp.1301-1308.

Deary, I. J., S. Strand, P. Smith, and C. Fernandes (2007). Intelligence and educational achievement. *Intelligence*, Vol. 35(1), pp. 13-21. doi:http://dx.doi.org/10.1016/j.intell.2006.02.

Denny, K. (2014). The effect of abolishing university tuition costs: Evidence from Ireland. *Labour Economics*, **Vol.** 26, pp. 26-33.

Denny K., and D. Flannery (2017). The economics of higher education participation. In Cullinan J., Flannery D. (Eds*) Economic Insights on Higher Education Policy in Ireland.* Cham: Palgrave Macmillan.

Dickerson, A. and 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), pp.535-558.

Doyle, O., (2020). The first 2,000 days and child skills. *Journal of Political Economy*, *128*(6), pp.2067-2122.

Duncan, G., Magnuson, K., Murnane, R. and Votruba-Drzal, E., (2019). Income Inequality and the Well-Being of American Families. *Family Relations*, *68*(3), pp.313-325.

Elliott, C. D., P. Smith, & K. McCulloch, (1996). *British ability scales second edition (BAS II).* Windsor, Berkshire: NFER-Nelson.

Feinstein, L. (2003): Inequality in the Early Cognitive Development of British Children in the 1970 Cohort. *Economica*. Vol. 70, pp. 73-97.

Flannery, D., and J. Cullinan (2014). Where they go, what they do and why it matters: the importance of geographic accessibility and social class for decisions relating to higher education institution type, degree level and field of study. *Applied Economics*, Vol. 46 pp. 2952-2965.

Hannan, D., E. Smyth, J. McCullagh, R. O'Leary and D. McMahon (1996). *Coeducation and Gender Equality: Exam Performance, Stress and Personal Development*. Dublin: Oak Tree Press.

Heckman, J.J. and Mosso, S., (2014). The economics of human development and social mobility. *Annu. Rev. Econ.*, *6*(1), pp.689-733.

Jacob, B. and Rothstein, J., (2016). The measurement of student ability in modern assessment systems. *Journal of Economic Perspectives*, Vol. 30, pp.85-108.

Lord, F. M. (1980). *Applications of Item Response Theory to Practical Testing Problems.* Hillsdale, NJ: Lawrence Erlbaum.

Madden, D. (2018). Bridging the gaps: inequalities in children's educational outcomes in Ireland. *The Journal of Economic Inequality*, Vol. 16, pp. 103-122.

Mani, A., Mullainathan, S., Shafir, E. and Zhao, J., (2013). Poverty impedes cognitive function. *Science*, *341*(6149), pp.976-980.

McCrory, C., J. Williams, A. Murray, A. Quail, and M. Thornton (2013). "Growing Up in Ireland NLSCI. Design, instrumentation and procedures for the infant cohort at wave two (3 years)." In *Technical report 3*. Department of Children and Youth Affairs, Dublin.

McGinnity, F., McMullin, P., Murray, A. and Russell, H., (2017). Social inequality in cognitive outcomes in Ireland: What is the role of the home learning environment and childcare?. In *Childcare, Early Education and Social Inequality*. Edward Elgar Publishing.

McMullin, P., McGinnity, F., Murray, A. and Russell, H., (2020). What You Do Versus Who You Are: Home-Learning Activities, Social Origin and Cognitive Skills among Young Children in Ireland. *European Sociological Review*.

McNally, S., McCrory, C., Quigley, J. and Murray, A., (2019). Decomposing the social gradient in children's vocabulary skills at 3 years of age: a mediation analysis using data from a large representative cohort study. *Infant Behavior and Development*, *57*, p.101326.

Murray, A., C. McCrory, M. Thornton, J. Williams, A, Quail, L. Swords et al (2011). *Growing Up in Ireland: National Longitudinal Study of Children: Design, Implementation and Procedure for the Child Cohort.* Dublin: Office of the Minister for Children and Youth Affairs.

Nixon, E., L. Swords and A. Murray (2013): *Growing Up in Ireland: Parenting and Infant Development. Infant Cohort. Report 3.* Dublin: Office of the Minister for Children and Youth Affairs.

O'Connell, P.J., Clancy, D. and McCoy, S. (2006). *Who Went to College in 2004? A National Survey of New Entrants to Higher Education.* Dublin: Higher Education Authority.

Quigley, J, and E. Nixon (2016). "Parental Investment and Child Development", in Williams, J., Nixon, E. and Smyth, E., eds. *Cherishing All the Children Equally?: Children in Ireland 100 Years on from the Easter Rising*. Oak Tree Press.

Rothstein, J., and N. Wozny (2013). "Permanent income and the black-white test score gap." *Journal of Human Resources,* Vol. 48, pp. 510-544.

Shorrocks, A.F., (1978). The measurement of mobility. *Econometrica: Journal of the Econometric Society*, pp.1013-1024.

Singh, A., C.J. Yeh, and S. Blanchard (2017). Ages and stages questionnaire: a global screening scale. *Boletín Médico Del Hospital Infantil de México (English Edition)*, Vol. 74, pp. 5-12.

Smyth, E. and C. Hannon (2007). School processes and the transition to higher education. *Oxford Review of Education*, Vol. 33, pp. 175-194.

Squires, J., Bricker, D. and Potter, L., (1997). Revision of a parent-completed developmental screening tool: Ages and Stages Questionnaires. *Journal of pediatric psychology*, *22*(3), pp.313-328.

Thornton, M., J. Williams, C.McCrory, A. Murray, A., and A. Quail, (2013). *Growing Up in Ireland: Design, Instrumentation and Procedures for the Infant Cohort at Wave 1 (9 Months)*. Dublin: Department of Children and Youth Affairs

Thornton, M., J. Williams, C.McCrory, A. Murray, A., and A. Quail, (2016). *Growing Up in Ireland: Design, Instrumentation and Procedures for the Child Cohort at Wave 2 (13 Years)*. Dublin: Department of Children and Youth Affairs.

Williams, J., S. Greene, E. Doyle et al (2009). *The Lives of 9 Year Olds: Growing Up in Ireland, National Longitudinal Study of Children (Report 1 of the Child Cohort)*. Dublin; The Stationery Office.

Williams, J., Nixon, E. and Smyth, E., (2016). *Cherishing All the Children Equally?: Children in Ireland 100 Years on from the Easter Rising*. Oak Tree Press.

Williams, J., M. Thornton, M. Morgan, A. Quail, E. Smyth, D. Murphy and D. O Mahony (2018). *Growing Up in Ireland. The Lives of 13 year Olds. Report 6.* Dublin; The Stationery Office.

Williams, J., A. Murray, D. O Mahony, A. Quail, C. O'Reilly, M. Thornton and M. Neary (2019). Growing Up in Ireland. National Longitudinal Study of Children. Technical Series Number 2019-7.

Figure 1a: CDFs by maternal education, Wave 1 (9 months) , GUI Infant

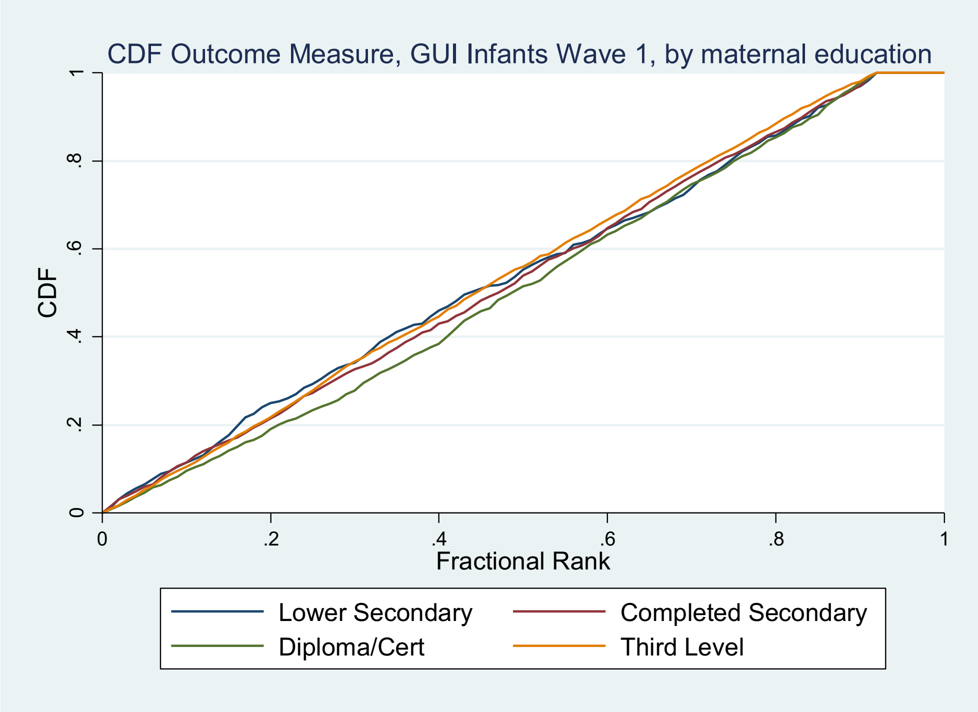


Figure 1b: CDFs by maternal education, Wave 2 (3 years), GUI Infant

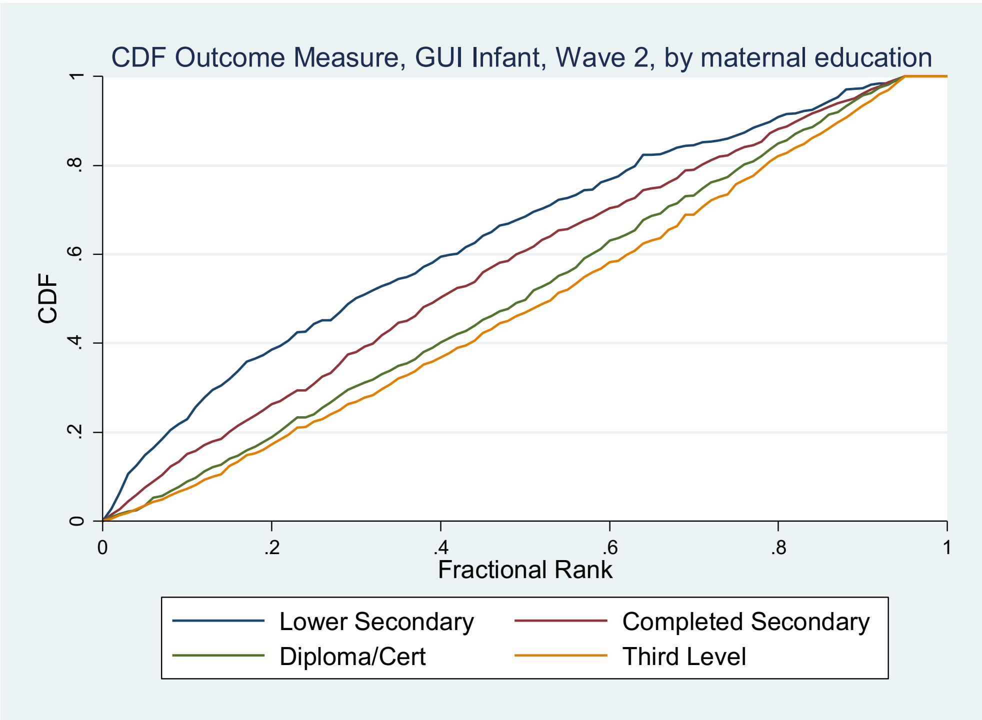


Figure 1c: CDFs by maternal education, Wave 3 (5 years), GUI Infant

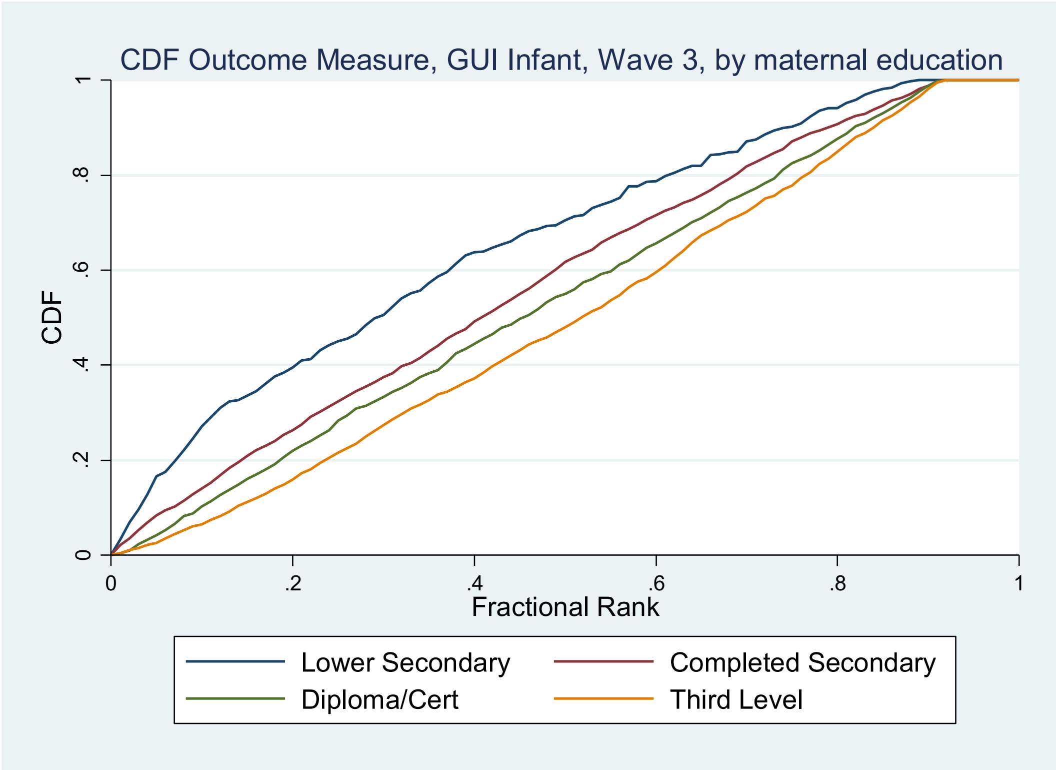


Figure 1d: CDFs by maternal education, Wave 5 (9 years), GUI Infant – note, reading only

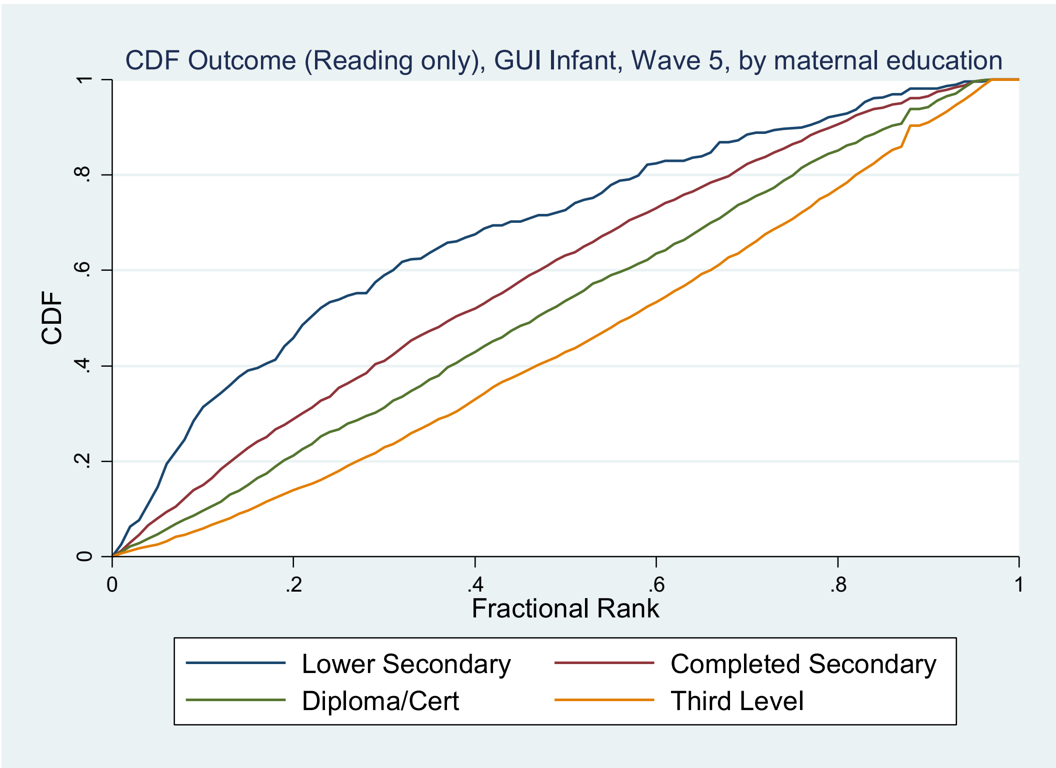


Figure 1e: CDFs by maternal education, Wave 1 (9 years), GUI Child, note, reading only

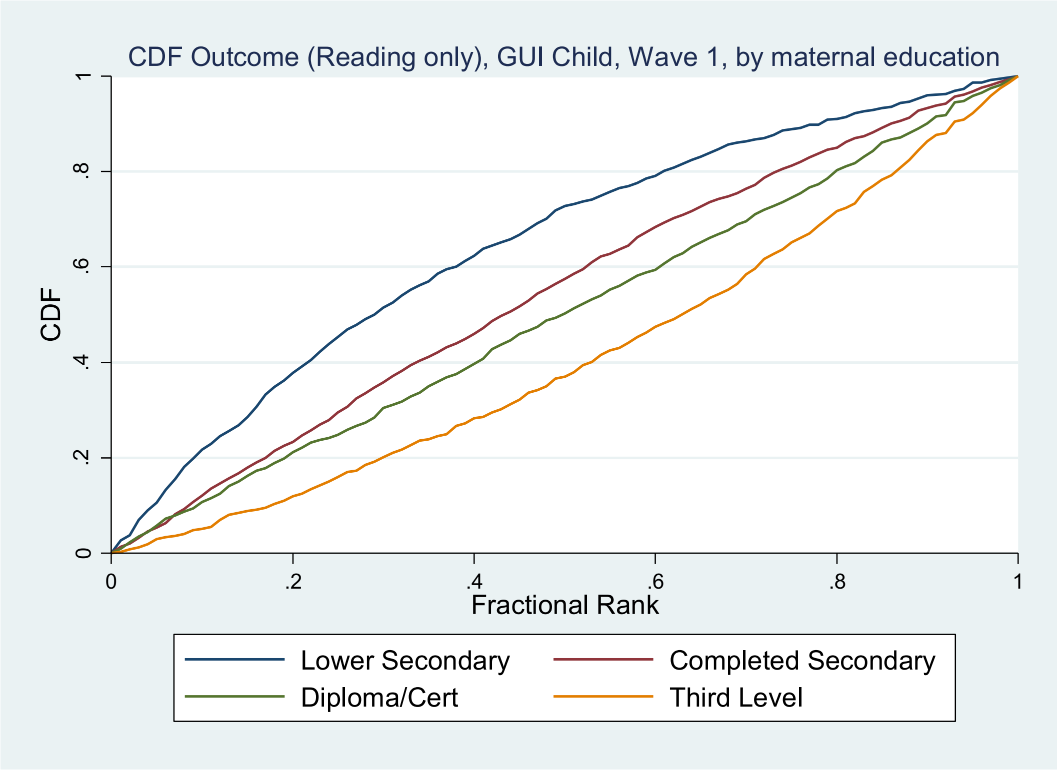


Figure 1f: CDFs by maternal education, Wave 1 (9 years), GUI Child

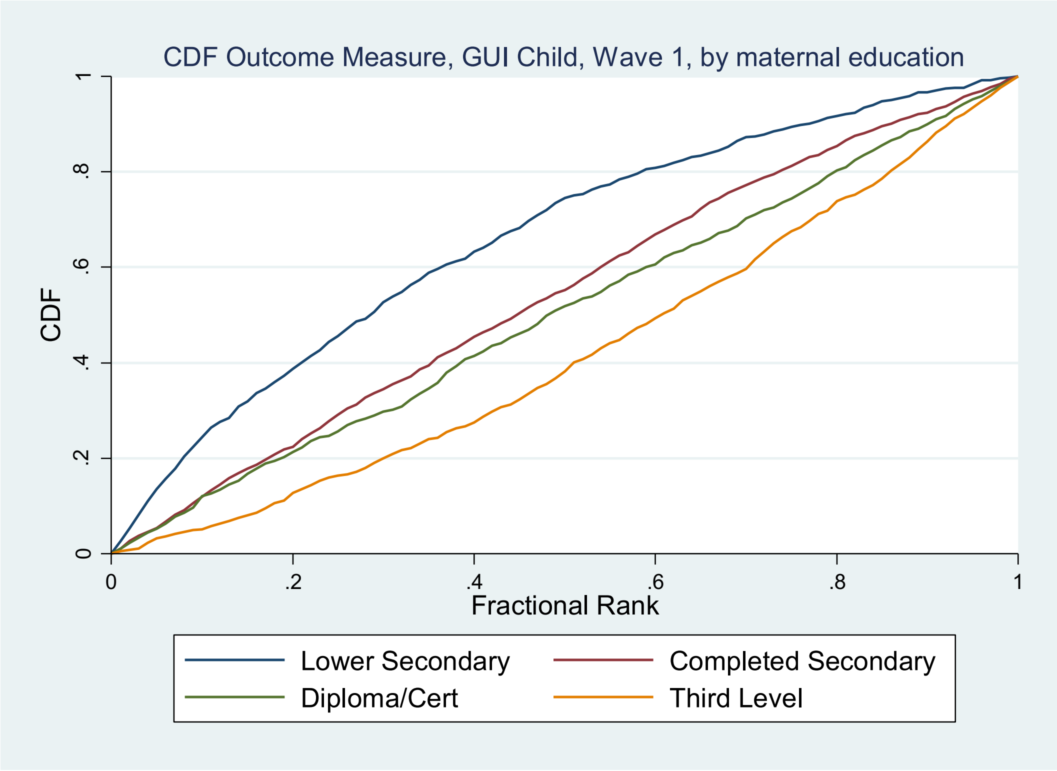


Figure 1g: CDFs by maternal education, Wave 2 (13 years), GUI Child

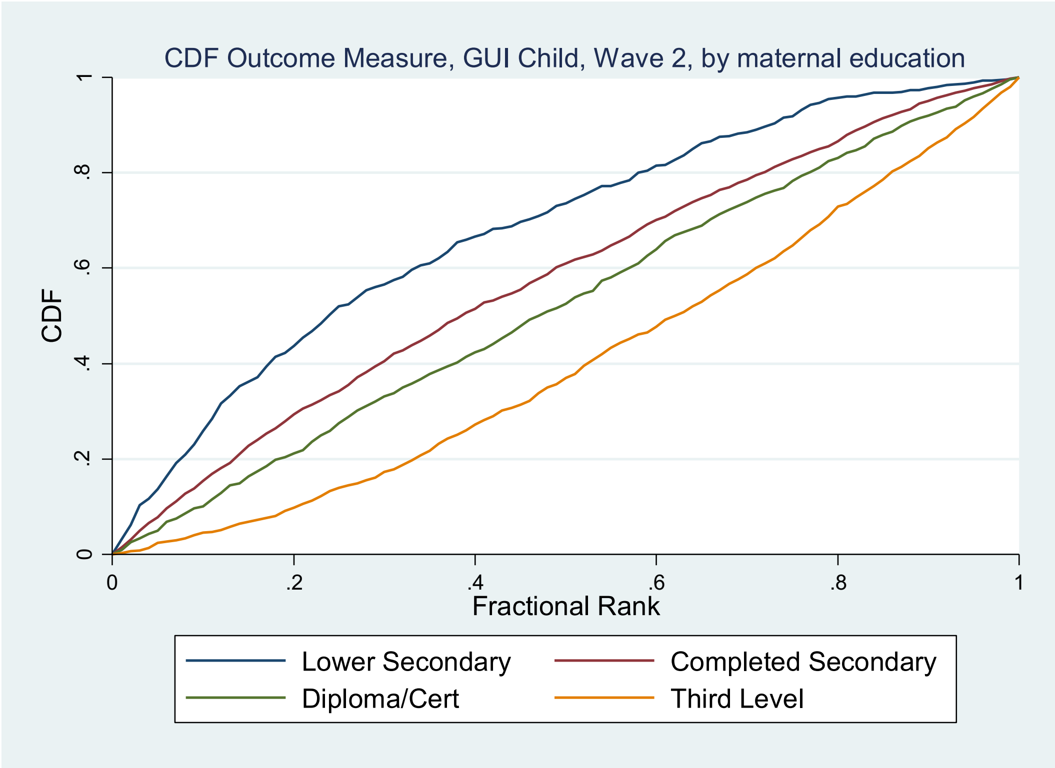


Figure 1h: CDFs by maternal education, Wave 3 (17 years), GUI Child

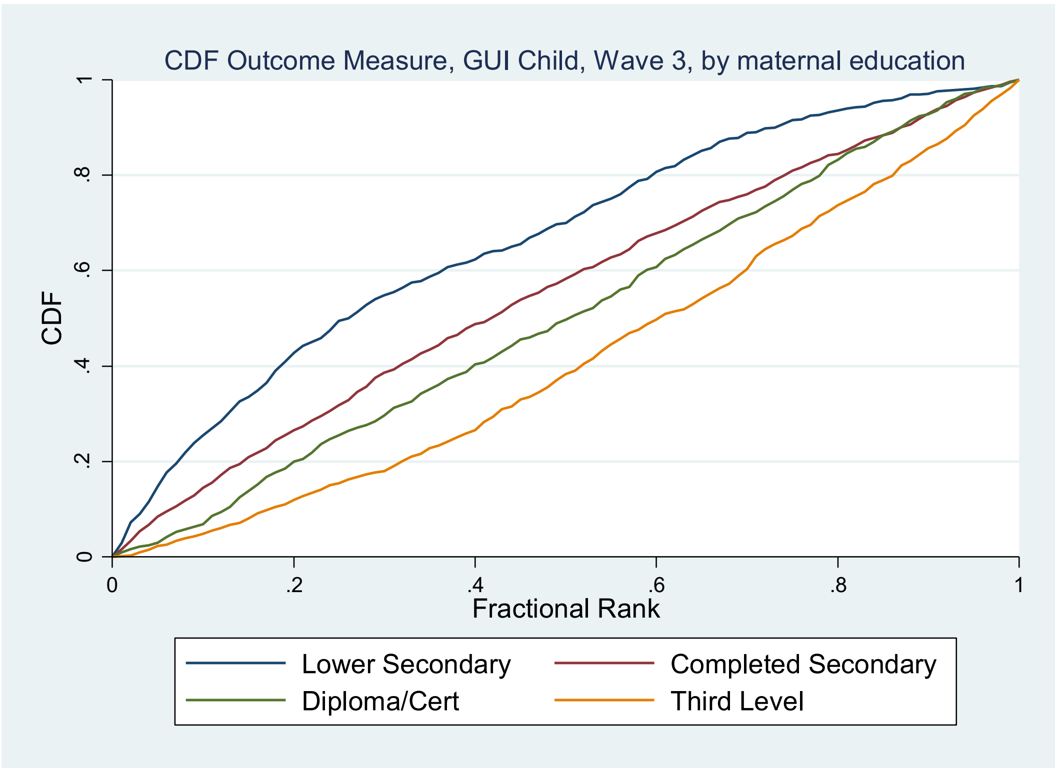


Figure 2a: Average Fractional Rank by Maternal Education

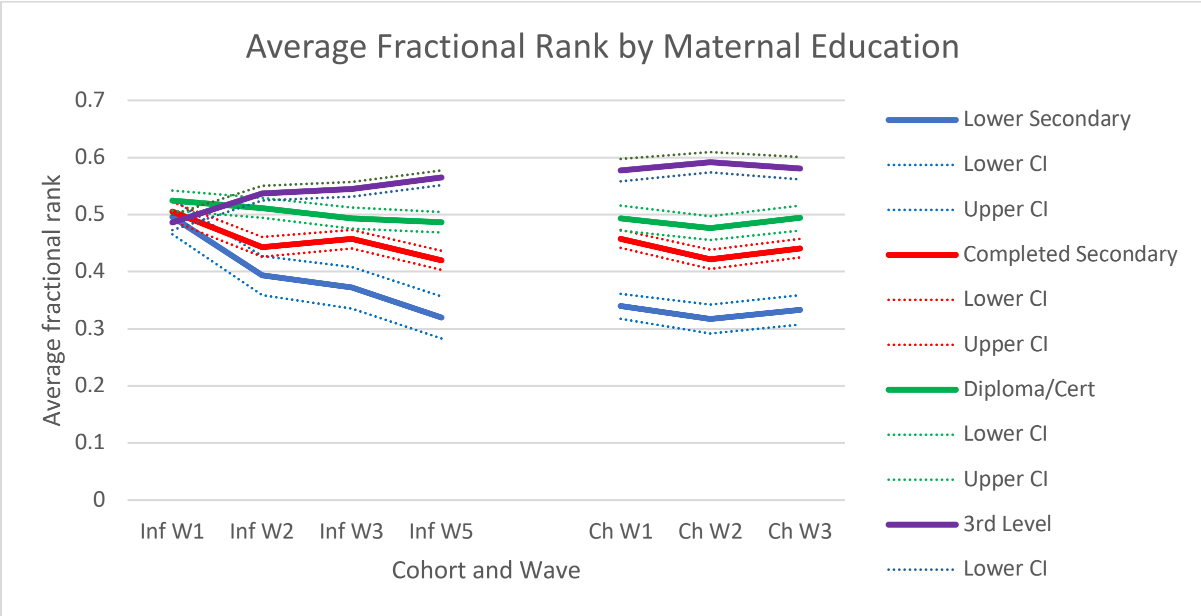


Figure 2b: Average Fractional Rank by Wave 1 Maternal Education

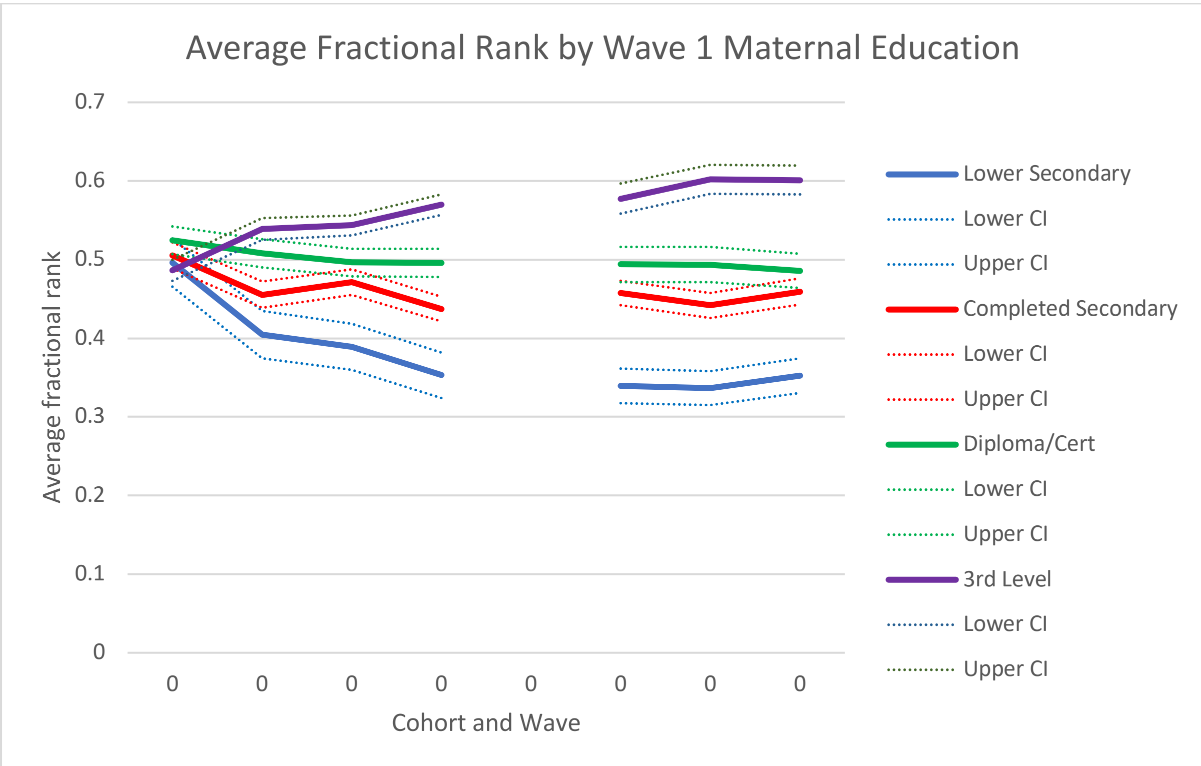


Figure 3a: Decomposition of Change in Rank Correlation, Infant Cohort, Maternal Education

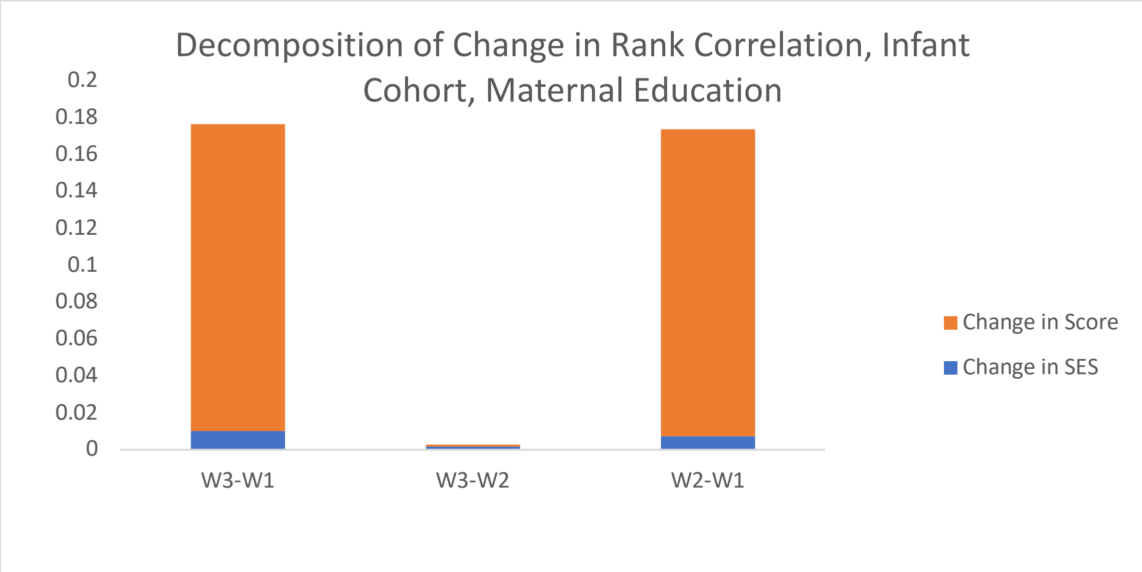


Figure 3b: Decomposition of Change in Rank Correlation, Infant Cohort, Income

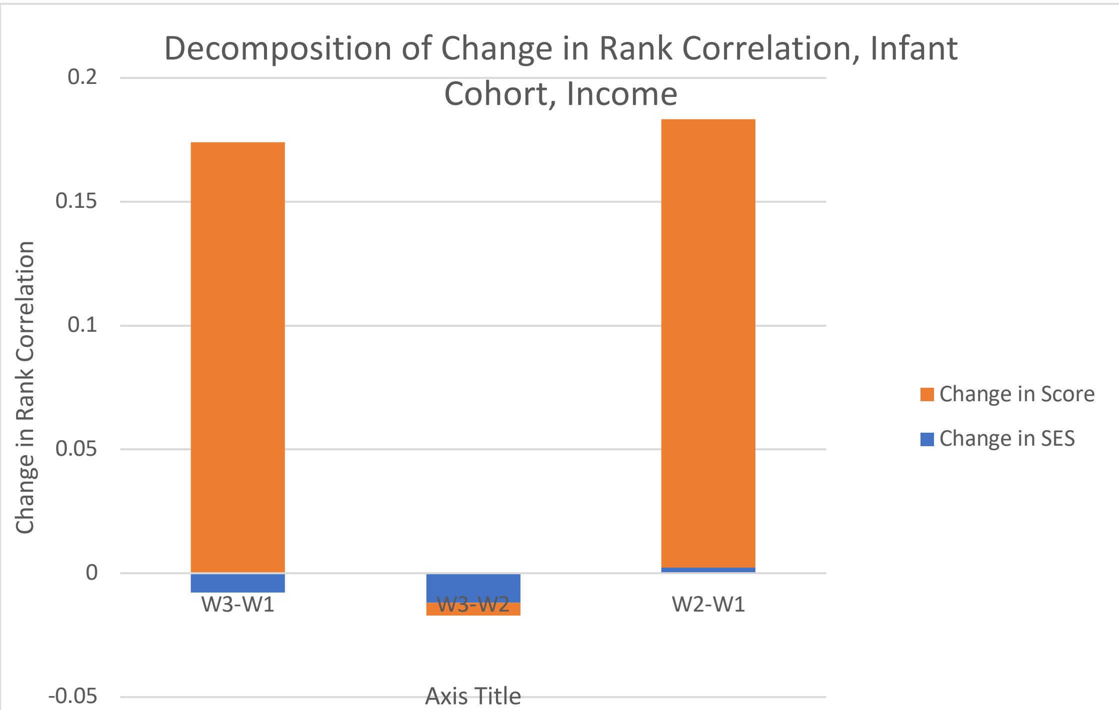


Figure 3c: Decomposition of Change in Rank Correlation, Child Cohort, Maternal Education

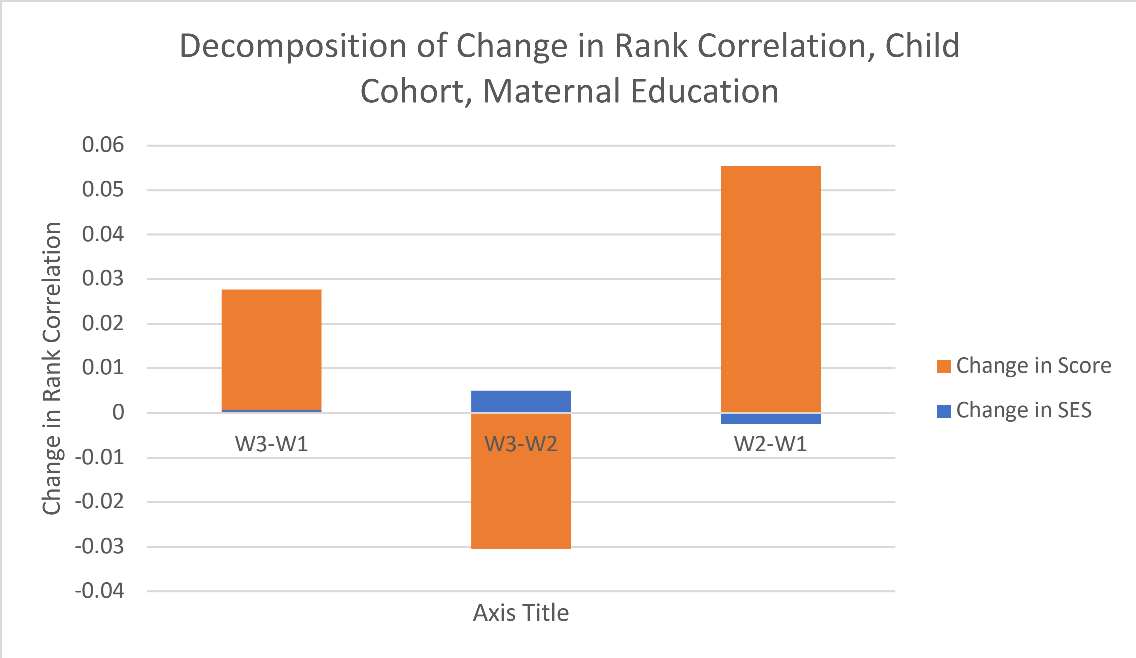


Figure 3d: Decomposition of Change in Rank Correlation, Child Cohort, Income

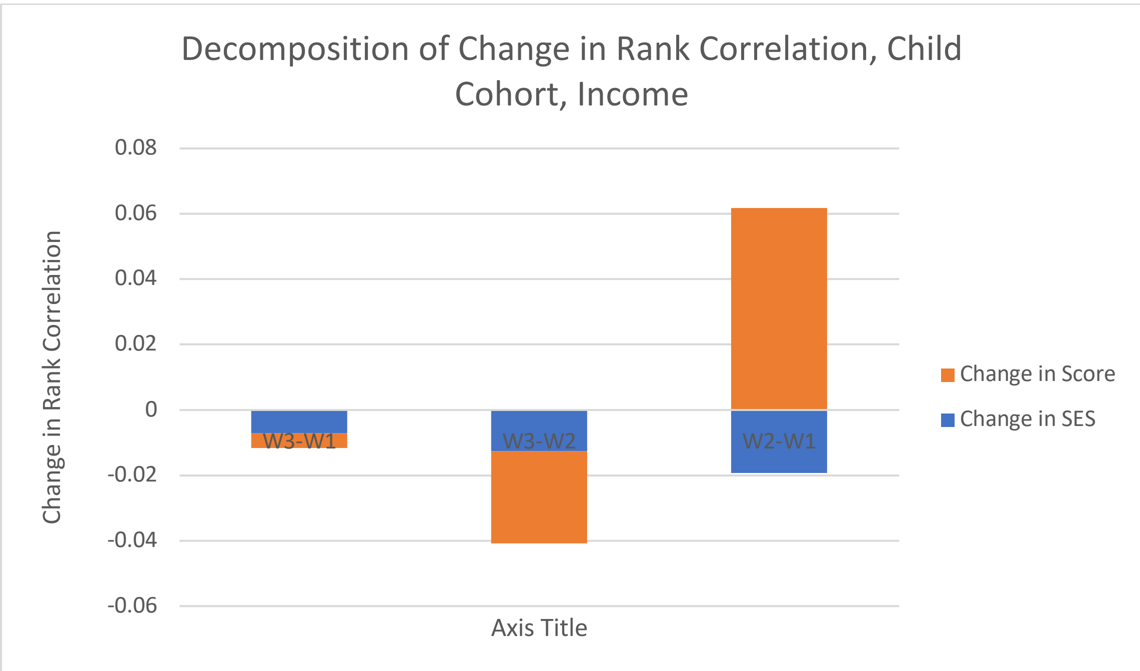


Figure 4a: Local Polynomial Rank Regression, Infant Cohort, Wave 2 – Wave 1

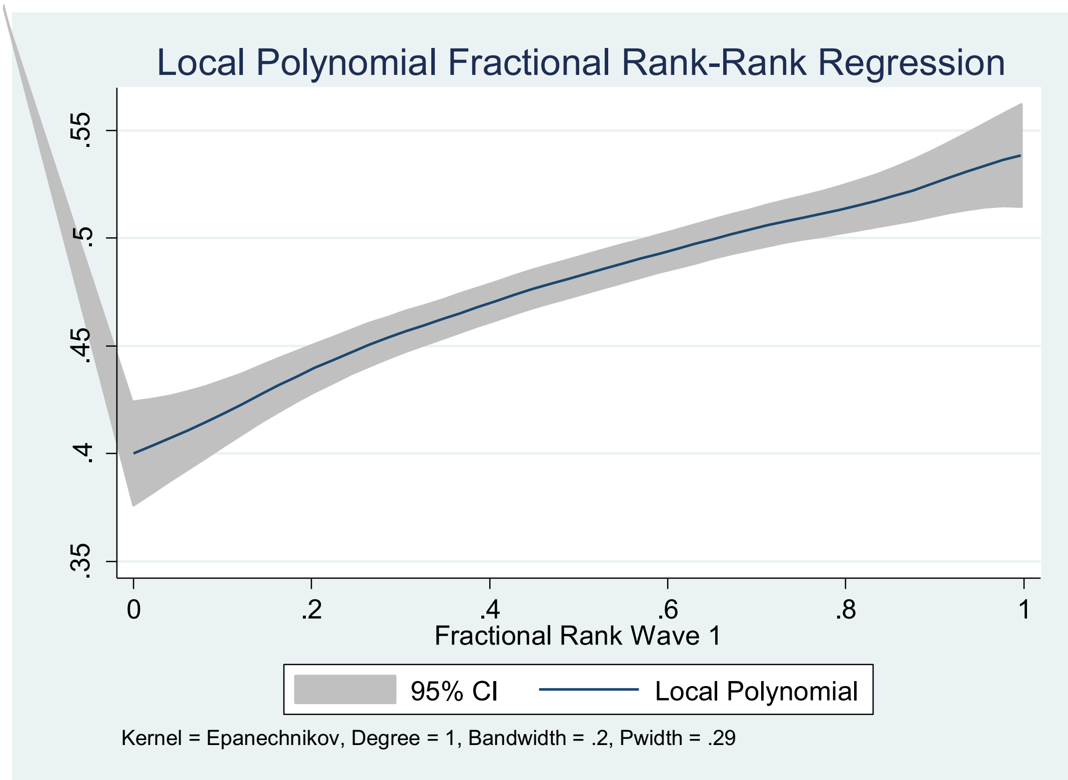


Figure 4b: Local Polynomial Rank Regression, Infant Cohort, Wave 3 – Wave 1

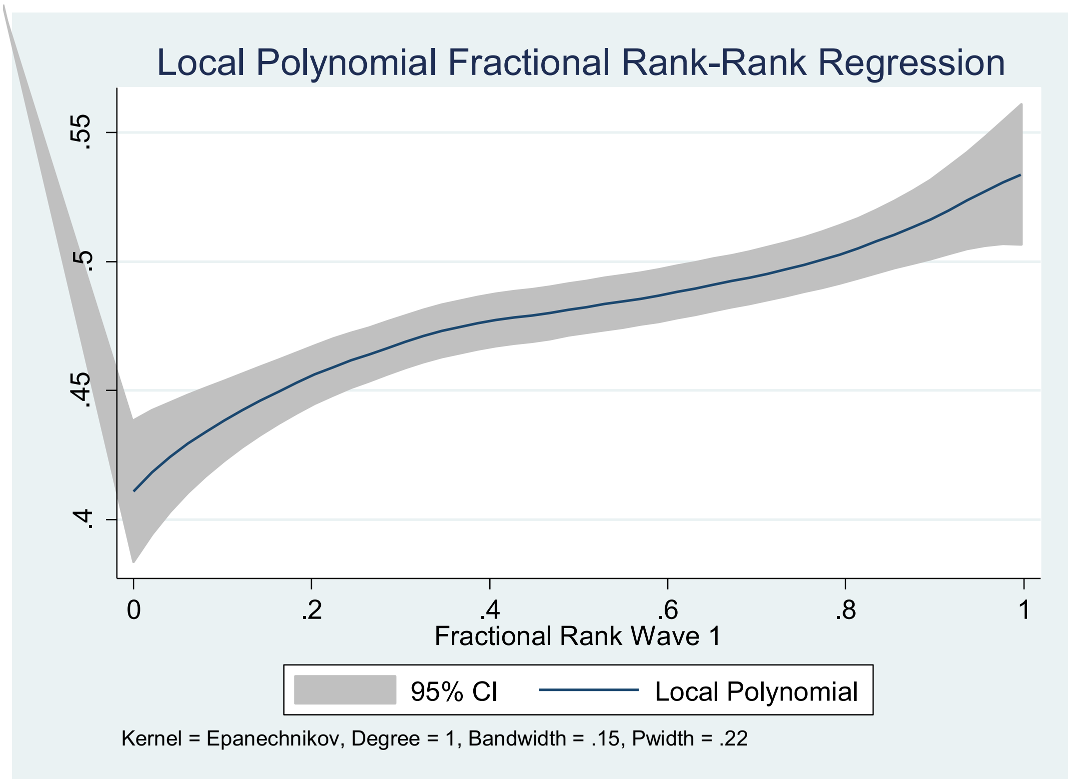


Figure 4c: Local Polynomial Rank Regression, Infant Cohort, Wave 5 – Wave 1

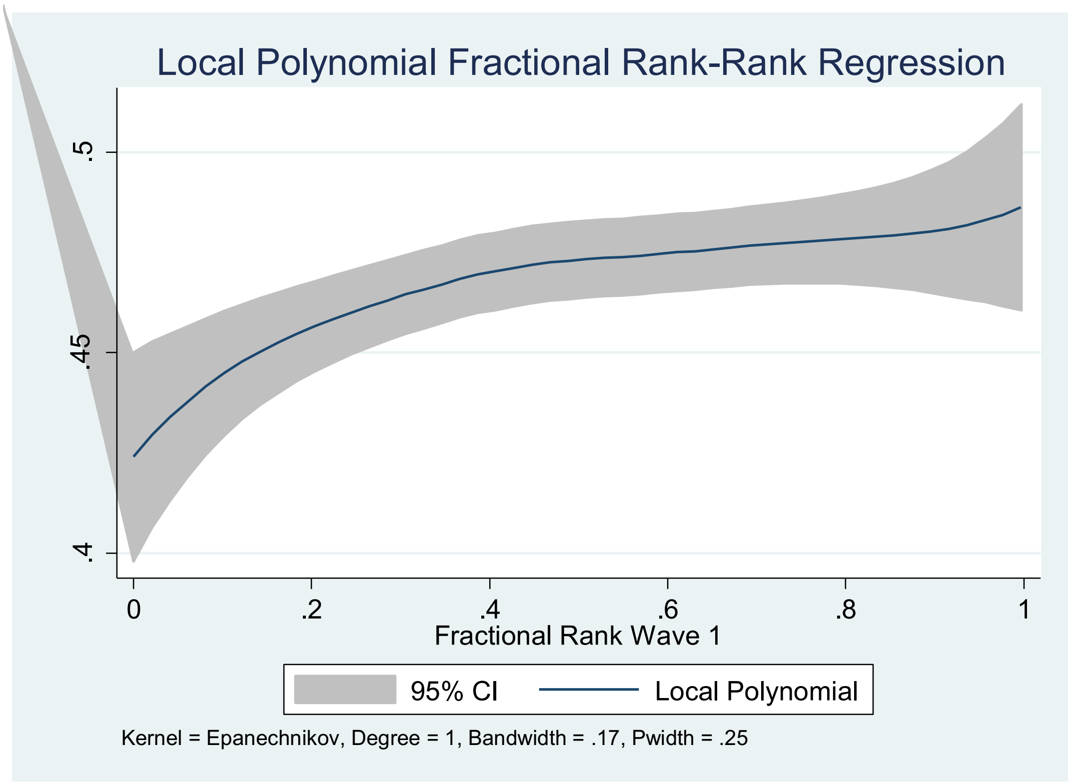


Figure 4d: Local Polynomial Rank Regression, Infant Cohort, Wave 3 – Wave 2

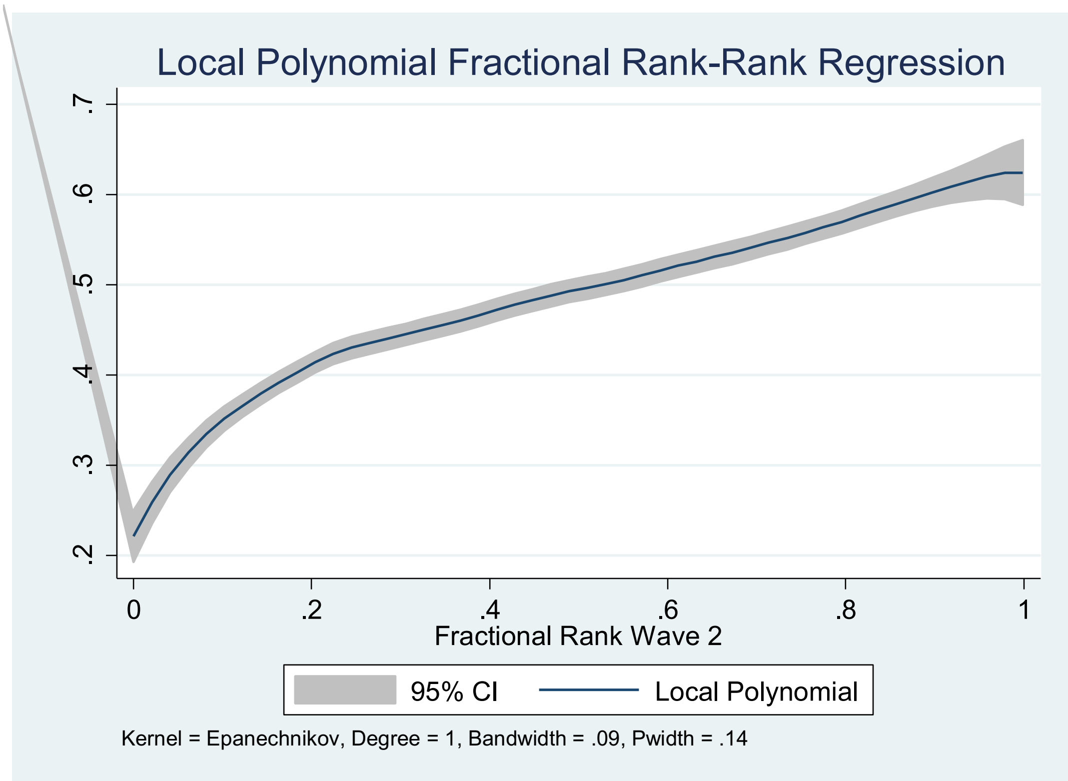


Figure 4e: Local Polynomial Rank Regression, Infant Cohort, Wave 5 – Wave 2

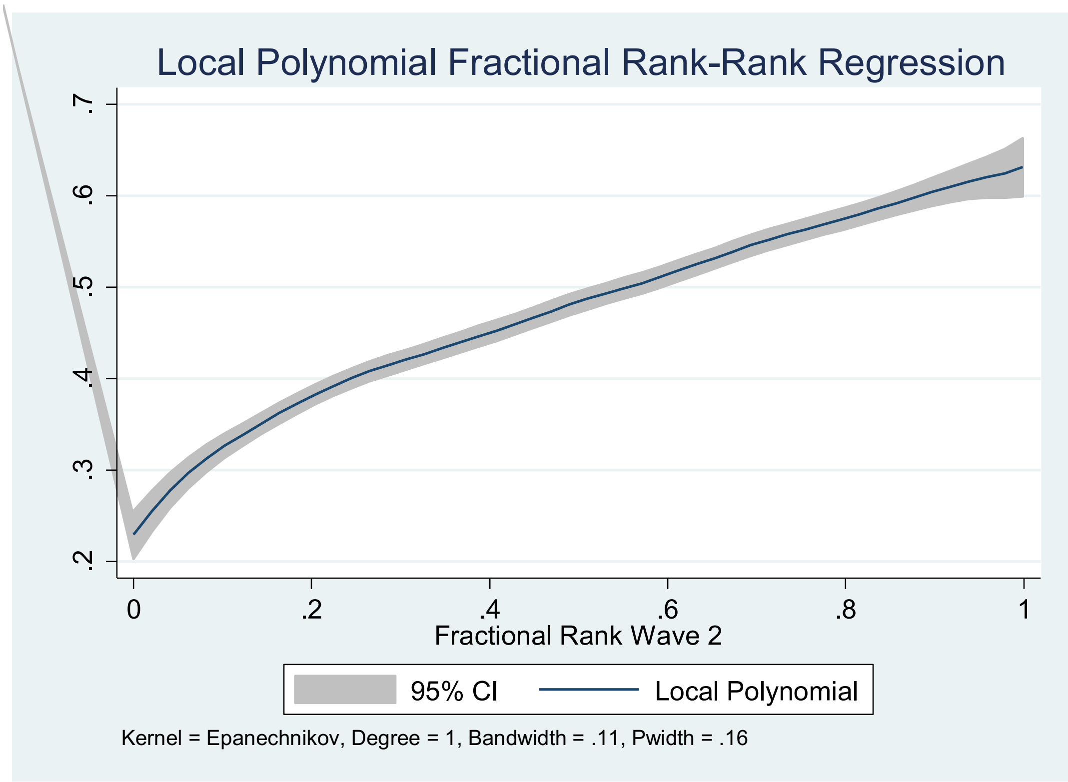


Figure 4f: Local Polynomial Rank Regression, Infant Cohort, Wave 5 – Wave 3

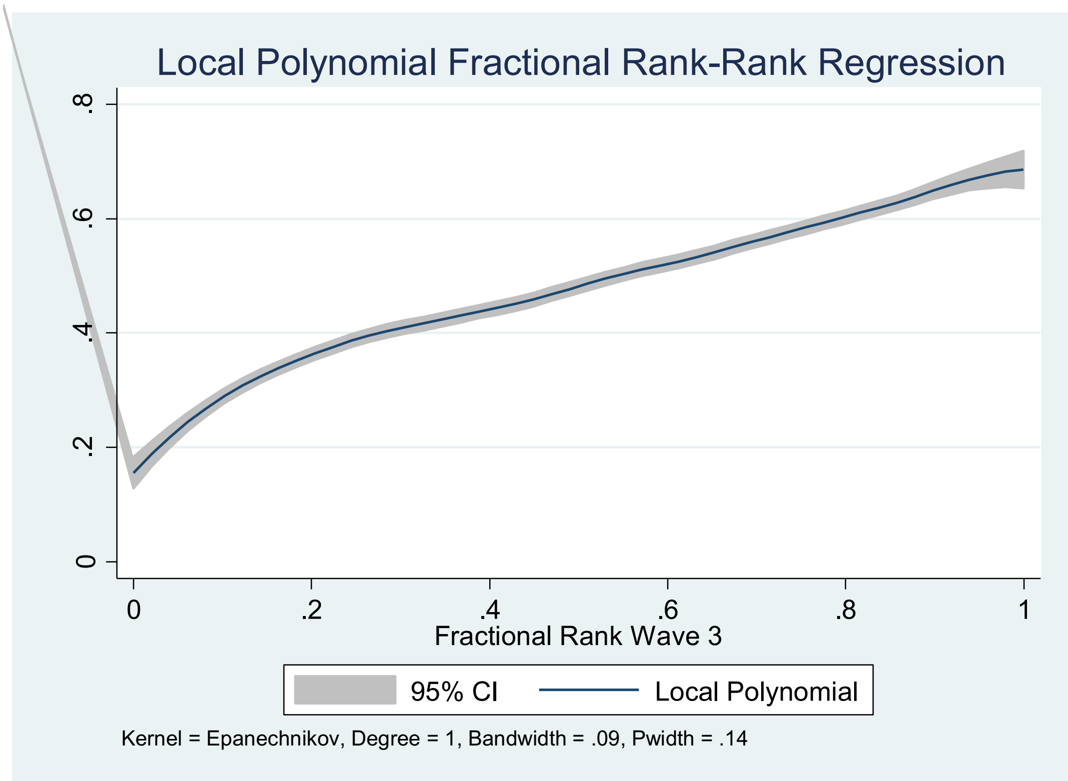


Figure 5a: Local Polynomial Rank Regression, Child Cohort, Wave 2 – Wave 1

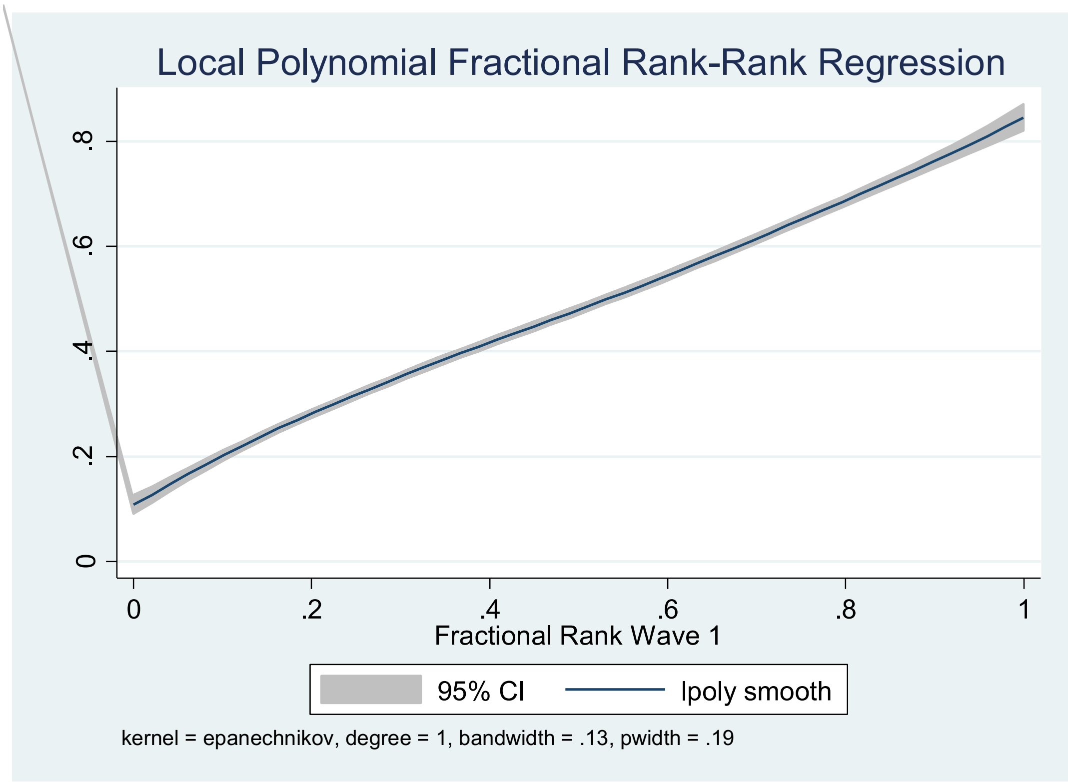


Figure 5b: Local Polynomial Rank Regression, Child Cohort, Wave 3 – Wave 1

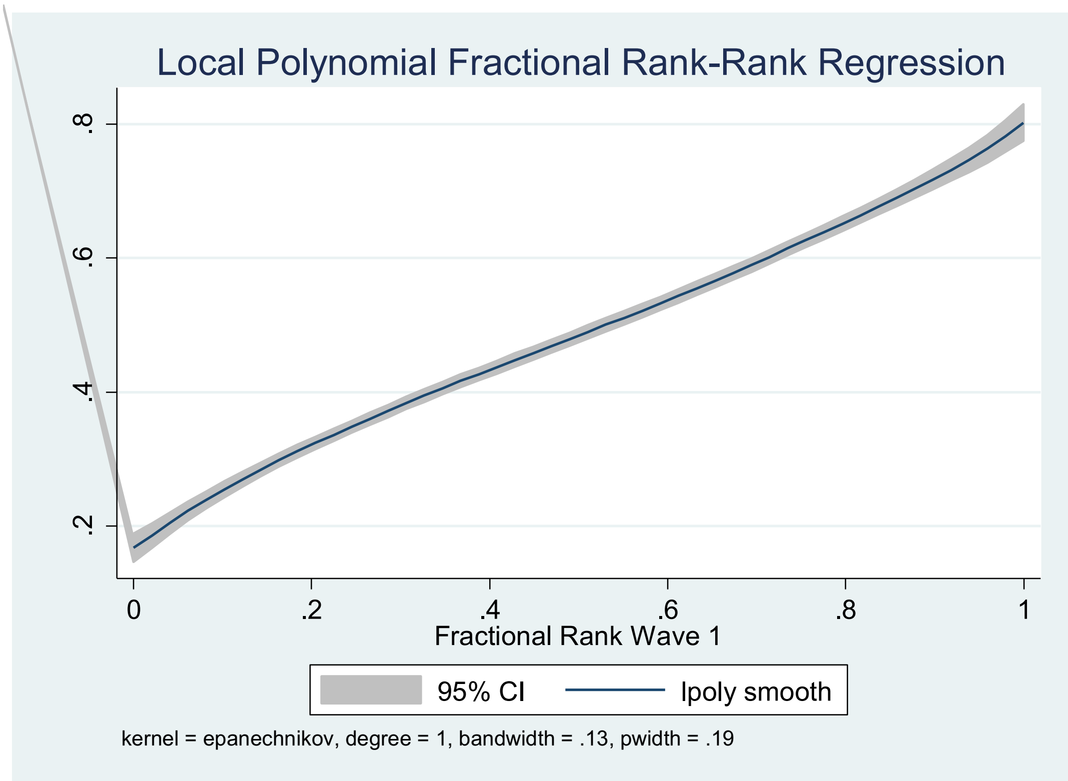
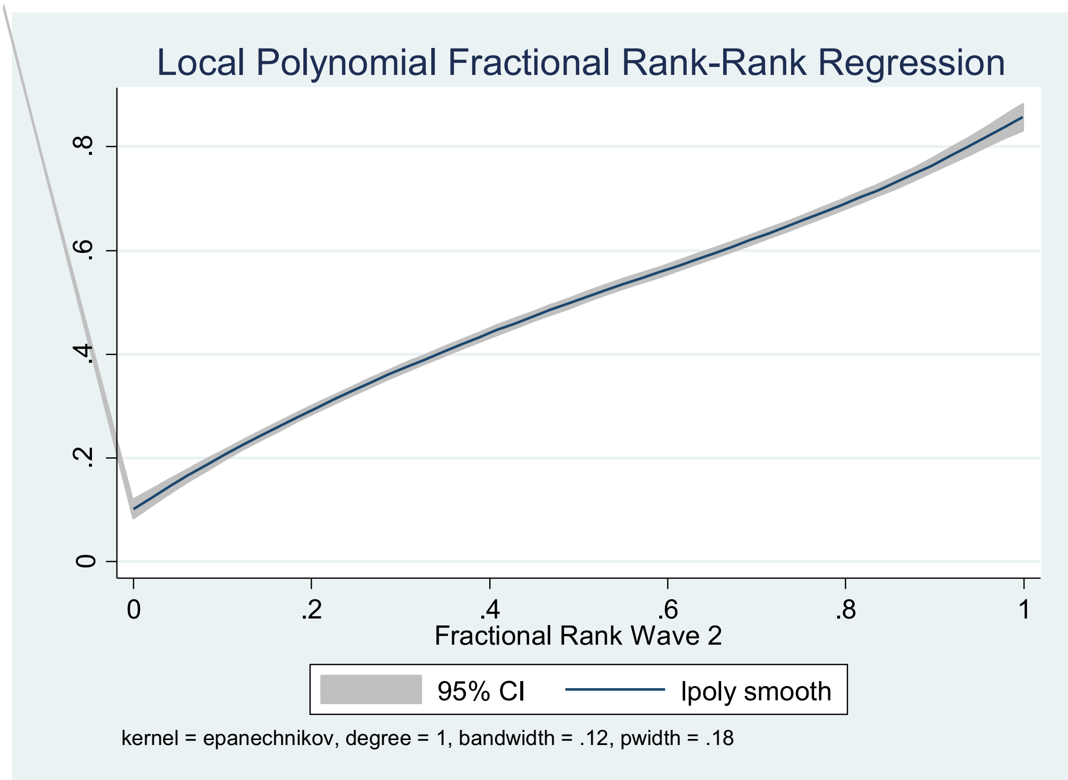


Figure 5c: Local Polynomial Rank Regression, Child Cohort, Wave 3 – Wave 2



Appendix 1 - Measures of Educational Outcomes

Infant Cohort	**Outcome**
Wave 1 – age 9 months	ASQ-2 measures - subscales in the areas of communication, gross motor, fine motor, problem solving and personal/social domains, with each subscale having a range from 0 to 60 (Squires et al, 1997). Organised as a separate set of questionnaires for 19 different age intervals ranging from 4 to 60 months. As the infants in GUI Infant are aged 9 months the results from the 10 month interval are used.
Wave 2 – age 3 years	Picture Similarities Scales and Naming Vocabulary Scales from the British Abilities Scales (Elliot et al, 1996) measuring reasoning/problem-solving and vocabulary respectively - standardised scores.
Wave 3 – age 5 years	Picture Similarities Scales and Naming Vocabulary Scales from the British Abilities Scales (Elliot et al, 1996) measuring reasoning/problem-solving and vocabulary respectively - standardised scores. Also teacher based achievement scales adapted from the UK Millenium Cohort Study and based upon the Foundation Stage Profile in England. There are five subscales covering Dispositions and Attitudes, Language for Communication and Thinking, Linking Sounds and Letters, Reading and Numeracy.
Wave 5	Test in reading administered by the GUI fieldworkers at school. Known in Ireland as the Drumcondra tests and a feature of the Irish educational system for a number of years and linked to the national curriculum. Logit scores from test are used, obtained via Item Response Theory.
Child Cohort	
Wave 1	Test in reading and maths administered by the GUI fieldworkers at school. Known in Ireland as the Drumcondra tests and a feature of the Irish educational system for a number of years and linked to the national curriculum. Logit scores from test are used, obtained via Item Response Theory.
Wave 2	Shortened versions of the Drumcondra Reasoning Test focussing on items related to numerical ability and verbal reasoning. These are measures of cognitive ability or aptitude rather than performance in school or academic achievement and the content of the test is not related to the school curriculum. Logit scores from test are used, obtained via Item Response Theory.
Wave 3	Cognitive Naming Test, Cognitive Maths Test and Cognitive Vocabulary Test, details in Williams et al (2019).

Table 1: Rank Correlations Across Different Subscales/Components

Infant Cohort

Wave 1 – aged 9 months

	ASQ Communication	ASQ Gross Motor	ASQ Fine Motor	ASQ Problem Solving	ASQ Personal Social
ASQ Communication	1.000				
ASQ Gross Motor	0.2386	1.000			
ASQ Fine Motor	0.2014	0.1818	1.000		
ASQ Problem Solving	0.2712	0.2034	0.3561	1.000	
ASQ Personal Social	0.3346	0.2610	0.2343	0.2987	1.000

Wave 2 – aged 3 years

	BAS Picture Similarities	BAS Naming Vocabulary
BAS Picture Similarities	1.000	
BAS Naming Vocabulary	0.3587	1.000

Wave 3 – aged 5 years

	Language	Linking	Reading	Numbers	Prob. Solv.	Nam. Voc.
Language	1.000					
Linking	0.4216	1.000				
Reading	0.4289	0.7071	1.000			
Numbers	0.3450	0.5595	0.6218	1.000		
Prob. Solv.	0.1171	0.1454	0.1342	0.1466	1.000	
Nam. Voc.	0.2489	0.2168	0.2117	0.1776	0.2738	1.000

Child Cohort

Wave 1 – aged 9 years

	Drumcondra Maths	Drumcondra Reading
Drumcondra Maths	1.000	
Drumcondra Reading	0.5851	1.000

Wave 2 – aged 13 years

	Drumcondra Numerical	Drumcondra Verbal
Drumcondra Numerical	1.000	
Drumcondra Verbal	0.5496	1.000

Wave 3 – aged 17 years

	Cognitive Naming	Cognitive Maths	Cognitive Vocab
Cognitive Naming	1.000		
Cognitive Maths	0.2357	1.000	
Cognitive Vocab	0.2969	0.3709	1.000

Table 2: Rank Correlation of Composite Measure Across Waves

Infant Cohort

	Wave 1	Wave 2	Wave 3	Wave 5
Wave 1	1.000			
Wave 2	0.130	1.000		
Wave 3	0.101	0.317	1.000	
Wave 5	0.039	0.334	0.437	1.000

Child Cohort

	Wave 1	Wave 2	Wave 3
Wave 1	1.000		
Wave 2	0.682	1.000	
Wave 3	0.569	0.692	1.000

We now discuss these tests in more detail, commencing with the Infant Cohort.

Infant Cohort

Wave 1

Children are aged 9 months in this wave. The test score is the Ages and Stages Questionnaire 2nd edition (ASQ-2, see Nixon et al, 2013, Elliot et al 1996) which has been applied in many different countries (Singh et al, 2017). ASQ-2 is primarily designed as a screening rather than as a diagnostic tool and it consists of five subscales in the areas of communication, gross motor, fine motor, problem solving and personal/social domains, with each subscale having a range from 0 to 60. It is organised as a separate set of questionnaires for 19 different age intervals ranging from 4 to 60 months. As the infants in GUI Infant are aged 9 months the results from the 10 month interval are deemed to be most appropriate.

Wave 2

Wave 2 of GUI Infant surveys the children at aged 3. It uses two standardised tests administered by the interviewer in the home. These are the Picture Similarities Scales and Naming Vocabulary Scales from the British Abilities Scales (Elliot et al, 1996) measuring reasoning/problem-solving and vocabulary respectively. In our analysis we use the standardised rather than raw scores.

Wave 3

By wave 3 the children in the infant cohort of GUI were aged 5 and had just started attending school. Consequently there are two educational outcomes available in this wave. The first of these are the Picture Similarities Scales and Naming Vocabulary Scales from the British Abilities Scales also collected in wave 2. In addition there are also teacher based achievement scales adapted from the UK Millenium Cohort Study and based upon the Foundation Stage Profile in England. There are five subscales covering Dispositions and Attitudes, Language for Communication and Thinking, Linking Sounds and Letters, Reading and Numeracy.

Wave 5

Wave 4 of GUI Infant was carried out when the children from that cohort were aged 7. However no data was collected in that wave which could be regarded as an educational outcome hence the next data we have for the Infant cohort is from wave 5 when the children were aged 9. In that wave the children undertook the Drumcondra reading test. As that information was also collected for wave 1 of the child cohort we describe it in detail in the next section.

Child Cohort

Wave 1

In wave 1 of the child cohort the vast majority of the children were aged 9 and part of the survey consisted of tests in mathematics and reading which were administered by the GUI fieldworkers at school. These tests are known in Ireland as the Drumcondra tests and have been a feature of the Irish educational system for a number of years and are linked to the national curriculum. These are administered on an annual basis to all children in the primary school system. However, the particular tests for the GUI survey had not been seen by schools, teachers or pupils in advance of their use in GUI, thus it seems unlikely that students would have been intensively prepared for these tests, although they would have had some familiarity with tests of this kind from previous years.[12] It should be noted that the Drumcondra tests have no implications for further progression in the school system. The particular cohort of nine year

[12] For more details on these tests see Murray et al (2011).

olds in the GUI survey were spread over three different school grades (2nd, 3rd and 4th class) and three different levels of the test were administered, with the majority of the children in 3rd class (roughly equivalent to grade 3 in the US).

The test scores used for this wave are the results from these tests in maths and reading. As the tests were administered at three different levels it was necessary to standardise the results, hence the data we use are the logit scores which were obtained from the original raw data using the principles of Item Response Theory (see Lord, 1980). Results from tests at this age (and earlier) have been shown to have predictive power for subsequent later-life outcomes in areas such as education and health (Feinstein, 2003). It is important to note that the tests administered in wave 1 are *achievement* tests, based on the existing Irish primary school curriculum and essentially measures the amount the child would have learned at school up to then.

Wave 2

In wave 2 the children were now mostly aged 13 and the vast majority had entered the secondary school system. The secondary school system (which lasts from the ages of about 12-13 to 18) is more diverse in terms of curriculum and students have choice regarding what subjects they take (though practically every student will take Mathematics and English). The tests administered in wave 2 of GUI were shortened versions of the Drumcondra Reasoning Test focussing on items related to numerical ability and verbal reasoning. Thus critically they are measures of *cognitive ability or aptitude* rather than performance in school or academic achievement and the content of the test was *not* related to the school curriculum. As with wave 1, the scores which formed the basis of the composite measure are the logit scores from the test again obtained via Item Response Theory.

It must be stressed that ability/aptitude and achievement tests differ (see Jacob and Rothstein, 2016, and Williams et al 2018). Aptitude tests refer to scholastic ability not related to the school curriculum. Since they do reflect the acquisition of certain skills it is highly likely that they will be influenced by the environment (school and home) where these skills are acquired but they are not specifically linked to the school curriculum. Achievement tests however measure performance and will be strongly influenced by school and home factors. The two measures are generally agreed to be quite strongly correlated (see Deary, 2007) and Hannan (1996) finds that verbal and numerical performance in the Differential Aptitude Test was highly predictive of subsequent achievement in the Junior Certificate. The Drumcondra Reasoning Tests were also chosen on the basis that they would provide some comparability across the waves of GUI (Thornton et al, 2016). However it is important to bear in mind the warning of Williams et al (2018): *"Although the 13 year old's results on the Drumcondra Reasoning Tests may be correlated with their academic achievement or school performance, it is important to emphasise the conceptual difference between the cognitive measure of ability captured by the DRT and a measure of school achievement or performance."*

Wave 3

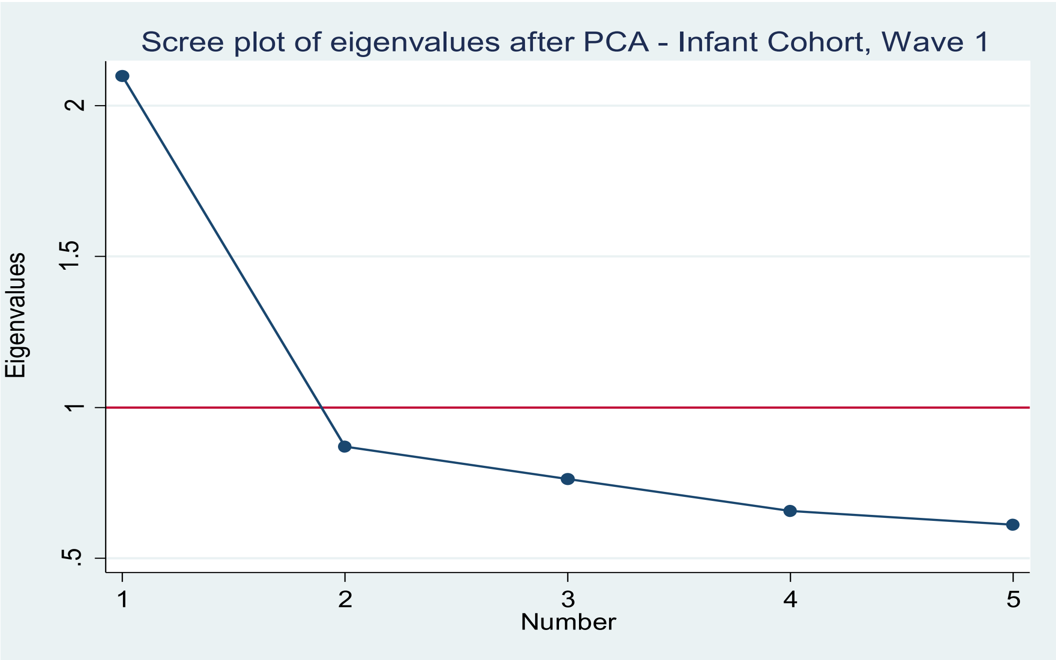
Wave 3 of GUI Child Cohort has outcomes from three cognitive tests. These are the Cognitive Naming Tasks, Cognitive Maths Score and Cognitive Vocabulary Test.

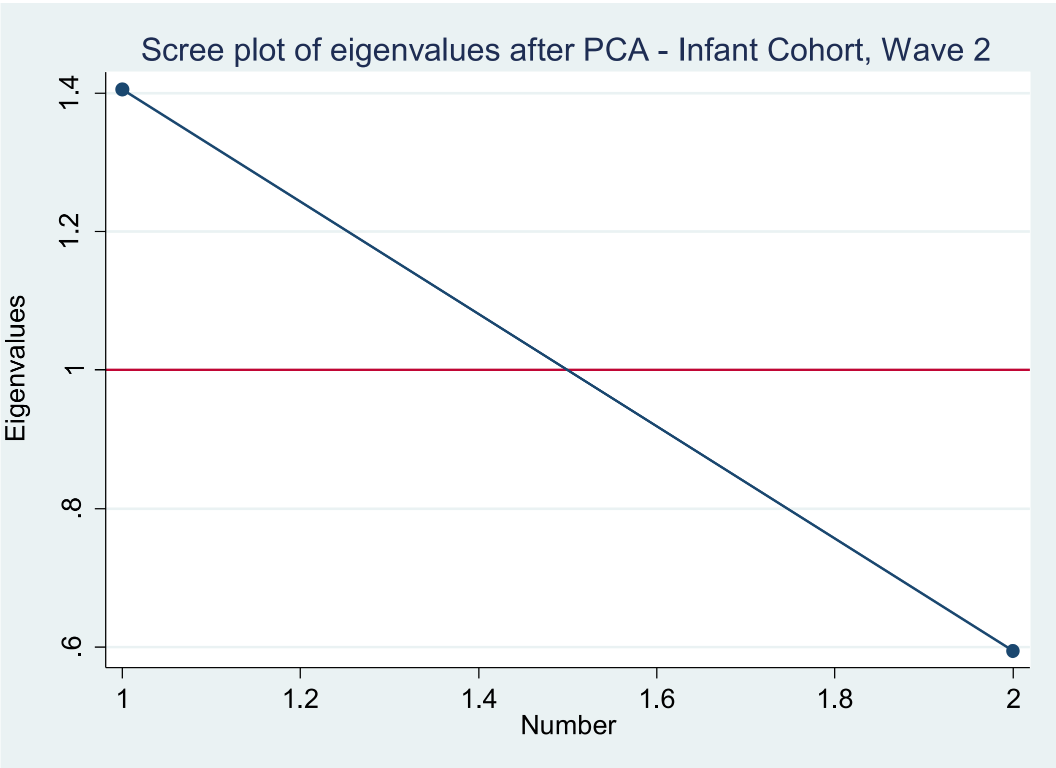
The Naming Task, also known as the Semantic Fluency Test involved the participant naming as many animals as they could think of in one minute and draws on general knowledge in long term memory. The Maths test involved three short questions aimed at testing the participant's ability to perform simple mathematical calculations and they also test financial literacy. The Vocabulary test consists of 20 words sharply increasing in difficulty. Each word is accompanied by five other words and the participant has to choose the word closest in meaning to the target word. Further details of the tests are available in Williams et al (2019).

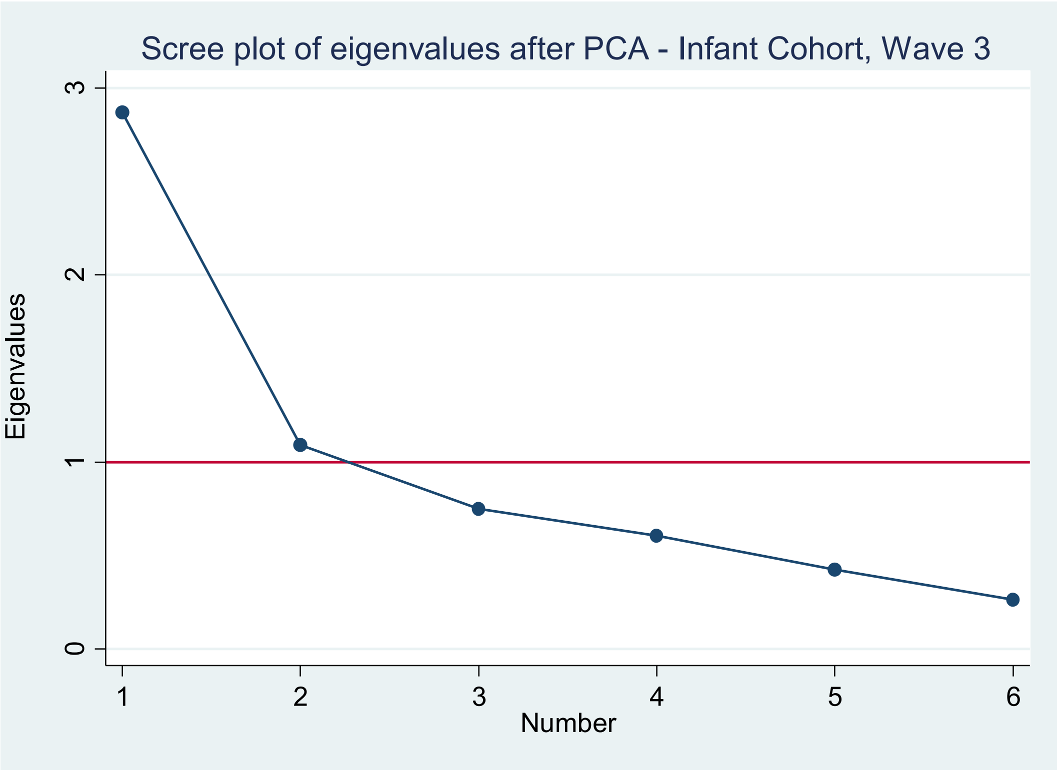
Table 3: Fraction of Variance Explained by 1st Principal Component

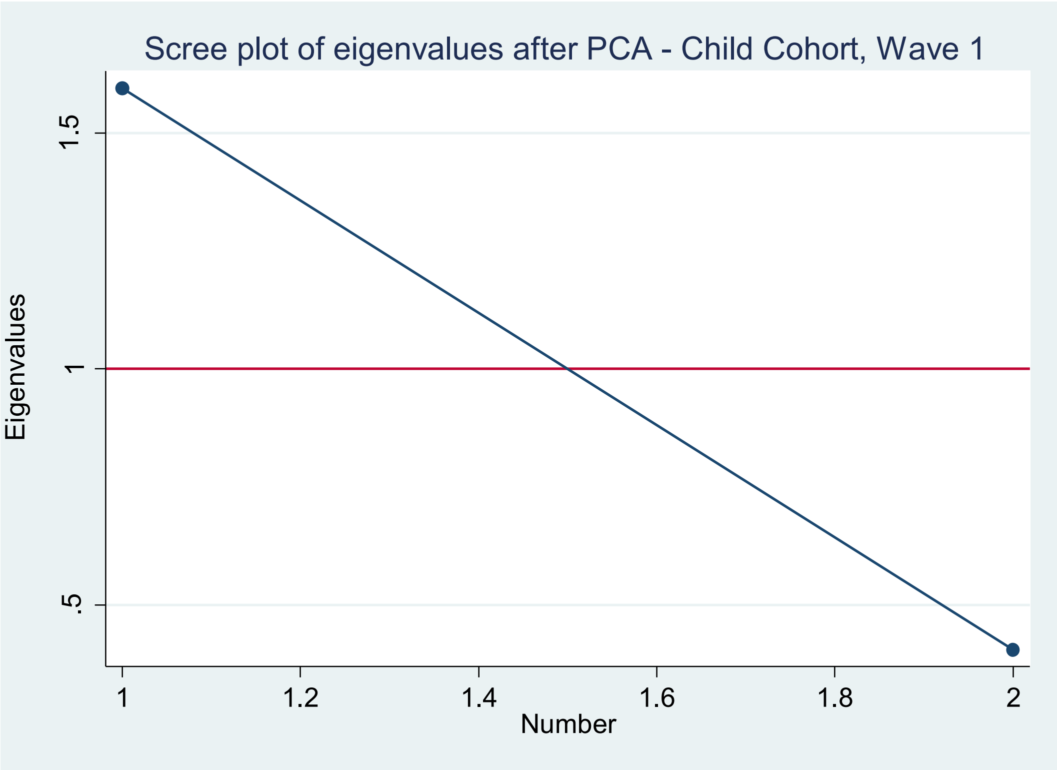
	Infant Cohort			Child Cohort		
	Wave 1	**Wave 2**	**Wave 3**	**Wave 1**	**Wave 2**	**Wave 3**
Fraction of variance	0.420	0.703	0.478	0.798	0.773	0.551
KMO	0.742	0.500	0.794	0.500	0.500	0.6234

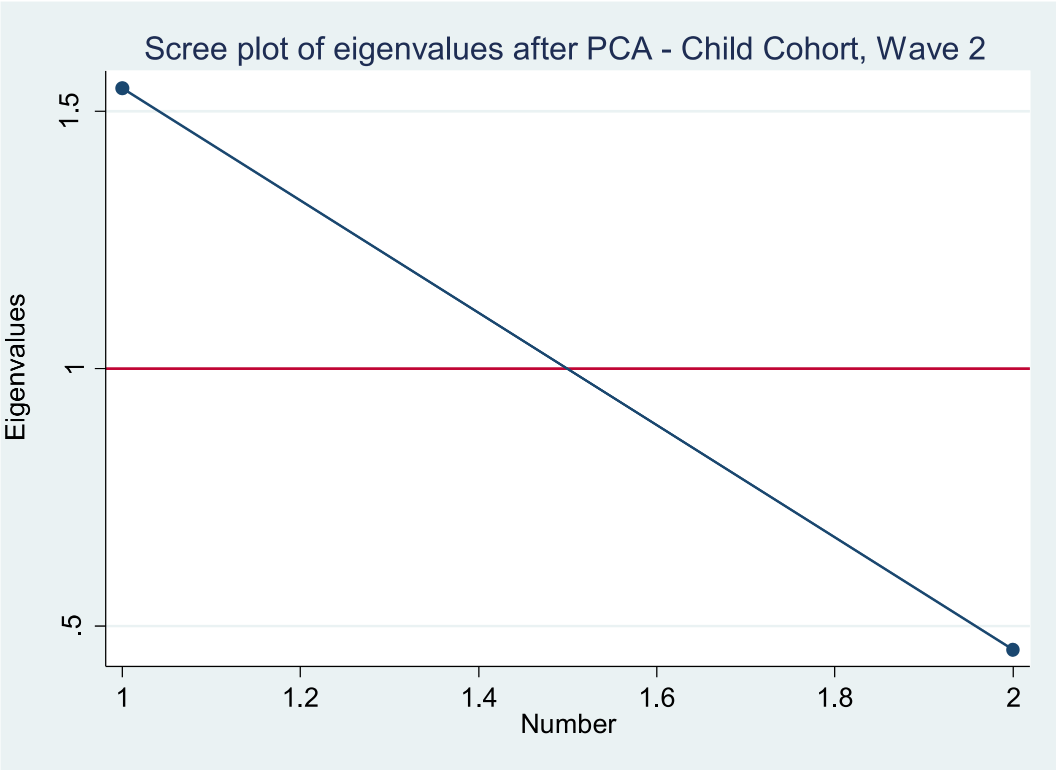
Scree Plots for PCA

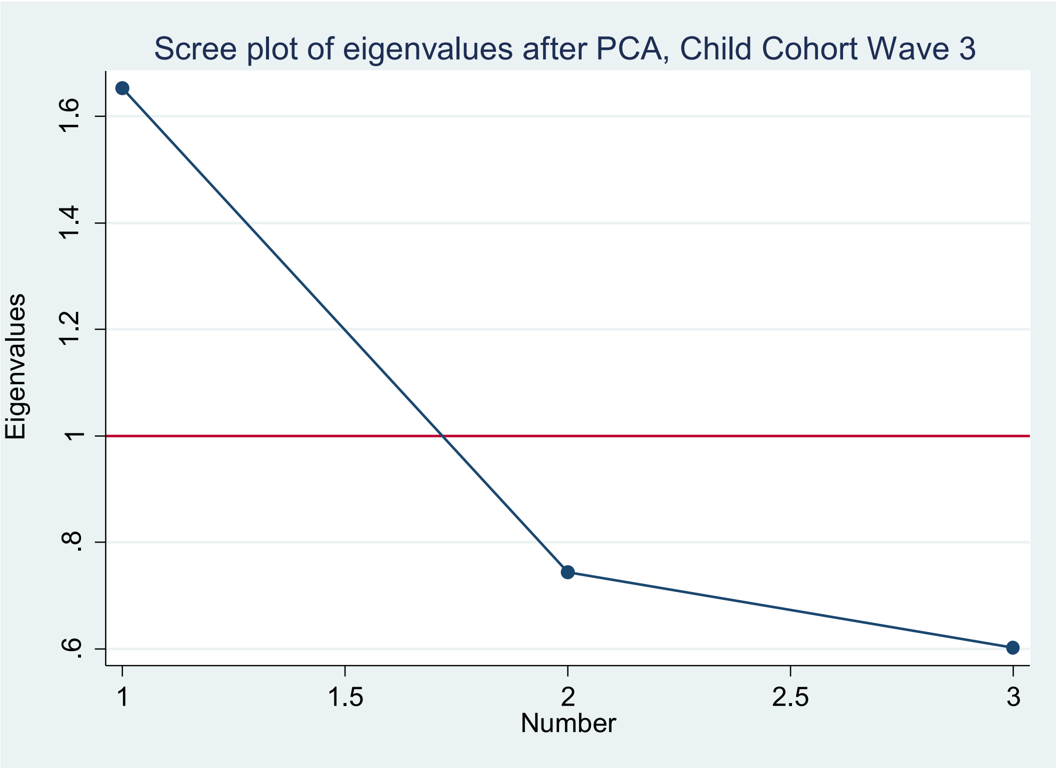





Scree plot of eigenvalues after PCA - Infant Cohort, Wave 3
Eigenvalues
0
1
2
3
1
2
3
4
5
6
Number


Scree plot of eigenvalues after PCA - Child Cohort, Wave 1
Eigenvalues
.5
1
1.5
1
1.2
1.4
1.6
1.8
2
Number


Scree plot of eigenvalues after PCA - Child Cohort, Wave 2
Eigenvalues
1.5
1
.5
1
1.2
1.4
1.6
1.8
2
Number


Scree plot of eigenvalues after PCA, Child Cohort Wave 3
Eigenvalues
1.6
1.4
1.2
1
.8
.6
1
1.5
2
2.5
3
Number

Appendix 2: Transition Matrices

Table 1a: Transition matrix, wave 1 to wave 2, Infant Cohort

	1	2	3	4	5
1	0.27	0.21	0.18	0.19	0.15
2	0.23	0.21	0.2	0.19	0.16
3	0.17	0.21	0.22	0.2	0.19
4	0.18	0.19	0.19	0.24	0.2
5	0.15	0.2	0.21	0.19	0.25

Table 1b: Transition matrix, wave 1 to wave 3, Infant Cohort

	1	2	3	4	5
1	0.24	0.2	0.2	0.18	0.16
2	0.21	0.2	0.21	0.19	0.19
3	0.19	0.19	0.2	0.21	0.2
4	0.19	0.21	0.19	0.21	0.21
5	0.16	0.2	0.2	0.2	0.24

Table 1c: Transition matrix, wave 1 to wave 5, Infant Cohort

	1	2	3	4	5
1	0.23	0.2	0.21	0.18	0.19
2	0.21	0.2	0.2	0.19	0.2
3	0.2	0.2	0.19	0.22	0.2
4	0.19	0.18	0.21	0.21	0.22
5	0.18	0.22	0.2	0.2	0.2

Table 1d: Transition matrix, wave 2 to wave 3, Infant Cohort

	1	2	3	4	5
1	0.38	0.2	0.19	0.14	0.08
2	0.21	0.25	0.21	0.2	0.14
3	0.18	0.21	0.22	0.19	0.2
4	0.13	0.18	0.2	0.25	0.24
5	0.1	0.16	0.18	0.23	0.34

Table 1e: Transition matrix, wave 2 to wave 5, Infant Cohort

	1	2	3	4	5
1	0.38	0.25	0.17	0.1	0.1
2	0.24	0.24	0.2	0.18	0.14
3	0.18	0.21	0.23	0.2	0.18
4	0.12	0.17	0.2	0.26	0.25
5	0.08	0.13	0.19	0.26	0.34

Table 1f: Transition matrix, wave 3 to wave 5, Infant Cohort

	1	2	3	4	5
1	0.46	0.23	0.13	0.1	0.07
2	0.24	0.25	0.21	0.17	0.12
3	0.16	0.22	0.24	0.21	0.17
4	0.09	0.18	0.22	0.27	0.24
5	0.05	0.12	0.19	0.25	0.4

Table 2a: Transition matrix, wave 1 to wave 2, Child Cohort

	1	2	3	4	5
1	0.58	0.26	0.11	0.04	0.01
2	0.25	0.3	0.25	0.16	0.05
3	0.12	0.28	0.26	0.24	0.1
4	0.04	0.12	0.26	0.32	0.26
5	0.01	0.04	0.12	0.25	0.58

Table 2b: Transition matrix, wave 1 to wave 3, Child Cohort

	1	2	3	4	5
1	0.49	0.25	0.16	0.08	0.02
2	0.28	0.27	0.2	0.16	0.1
3	0.15	0.24	0.26	0.22	0.14
4	0.07	0.17	0.22	0.27	0.27
5	0.02	0.08	0.17	0.27	0.47

Table 2c: Transition matrix, wave 2 to wave 3, Child Cohort

	1	2	3	4	5
1	0.61	0.25	0.11	0.02	0.01
2	0.28	0.3	0.23	0.15	0.05
3	0.08	0.25	0.29	0.27	0.11
4	0.03	0.14	0.26	0.29	0.28
5	0.01	0.06	0.11	0.28	0.54

UCD CENTRE FOR ECONOMIC RESEARCH – RECENT WORKING PAPERS

WP19/28 Judith M Delaney and Paul J Devereux: 'Math Matters! The Importance of Mathematical and Verbal Skills for Degree Performance' November 2019
WP19/29 Alan Fernihough and Cormac Ó Gráda: 'Across the Sea to Ireland: Return Atlantic Migration before the First World War' November 2019
WP19/30 Tamanna Adhikari and Karl Whelan: 'Do Business-Friendly Reforms Boost GDP?' December 2019
WP19/31 Judith M Delaney and Paul J Devereux: 'The Effect of High School Rank in English and Math on College Major Choice' December 2019
WP19/32 Tensay Hadush Meles, Lisa Ryan and Sanghamitra Mukherjee: 'Heterogeneity in Preferences for Renewable Home Heating Systems' December 2019
WP20/01 Cormac Ó Gráda: 'Economic History: 'An Isthmus Joining Two Great Continents'?' January 2020
WP20/02 Margaret Samahita and Håkan J Holm: 'Mining for Mood Effect in the Field' January 2020
WP20/03 Vincent Hogan and Patrick Massey: 'Professional Rugby on the Celtic Fringe' January 2020
WP20/04 Sanghamitra Chattopadhyay Mukherjee, Séin Healy, Tensay Hadush Meles, Lisa Ryan, Robert Mooney, Lindsay Sharpe and Paul Hayes: 'Renewable Energy Technology Uptake: Public Preferences and Policy Design in Early Adoption' February 2020
WP20/05 Ronald B Davies and James R Markusen: 'The Structure of Multinational Firms' International Activities' February 2020
WP20/06 Morgan Kelly: 'Direct Standard Errors for Regressions with Spatially Autocorrelated Residuals' March 2020
WP20/07 Anna Bayona and Oana Peia: 'Financial Contagion and the Wealth Effect: An Experimental Study' March 2020
WP20/08 Orla Doyle: 'Can Early Intervention have a Sustained Effect on Human Capital?' April 2020
WP20/09 Supriya Kapoor and Oana Peia: 'The Impact of Quantitative Easing on Liquidity Creation' April 2020
WP20/10 Judith M Delaney and Paul J Devereux: 'Choosing Differently? College Application Behaviour and the Persistence of Educational Advantage' May 2020
WP20/11 Jonathan Briody: 'Parental Unemployment During the Great Recession and Childhood Adiposity' May 2020
WP20/12 Tensay Meles, Lisa Ryan, and Joe Wheatley: 'COVID-19 and EU Climate Targets: Going Further with Less?' May 2020
WP20/13 Xidong Guo and Sarah Parlane: 'Addressing Private Practice in Public Hospitals' May 2020
WP20/14 Patrick Massey and Moore McDowell: 'EU Competition Law: An Unaffordable Luxury in Times of Crisis?' May 2020
WP20/15 Judith M Delaney and Paul J Devereux: 'How Gender and Prior Disadvantage predict performance in College' May 2020
WP20/16 Morgan Kelly, Joel Mokyr and Cormac Ó Gráda: 'The Mechanics of the Industrial Revolution' June 2020
WP20/17 Morgan Kelly and Cormac Ó Gráda: 'Connecting the Scientific and Industrial Revolutions: The Role of Practical Mathematics' June 2020
WP20/18 Ronald B Davies, Dieter F Kogler and Ryan Hynes: 'Patent Boxes and the Success Rate of Applications' June 2020
WP20/19 Ronald B Davies, Yutao Han, Kate Hynes, and Yong Wang: 'Competition in Taxes and IPR' June 2020

UCD Centre for Economic Research
Email economics@ucd.ie