Up and Down the Generational Income Ladder in Britain: Past Changes and Future Prospects

May 2008

Jo Blanden*

Stephen Machin**

* Department of Economics, University of Surrey and Centre for Economic Performance, London School of Economics.

** Department of Economics, University College London and Centre for Economic Performance, London School of Economics.

Abstract

This article seeks evidence on trends in intergenerational income for cohorts born after 1970. As many of these cohorts have not yet joined the labour market we must look at relationships between intermediate outcomes (degree attainment, test scores and non-cognitive abilities) and parental income to forecast forward from these to estimates of intergenerational earnings correlations. We find no evidence that the relationship between these intermediate outcomes and parental income have changed for more recent cohorts. Evidence from the earlier 1958 and 1970 cohorts shows that as mobility declined in the past the relationship between intermediate outcomes and parental income strengthened. We therefore conclude that under realistic assumptions, the decline in intergenerational mobility that occurred between the 1958 and 1970 is unlikely to continue for cohorts born from 1985 to 2000, and mobility is likely to remain at the low level observed for the 1970 cohort.

1. Introduction

Social mobility has become one of the high profile public policy issues of modern society. However, much of the evidence currently discussed is backward looking, owing to the very demanding data requirements and modelling issues needed to estimates changes in mobility over time. For income mobility, the extent to which people move up and down the generational ladder requires data on incomes of people and their parents which, by necessity, means they need to be old enough to earn an income.[1] This creates a problem for policy which needs to look forward and try to appraise how much (or little) social mobility there is likely to be for people not yet old enough to participate in the labour market.

Amongst people of labour market age, Blanden et al (2004) identified a fall in intergenerational income mobility in a comparison of a birth cohort born in 1970 compared to one born in 1958. More specifically, adult earnings of the second cohort were more closely linked to their parental income when they were aged 16 than was the case for the first cohort.[2] However, these results relate to individuals growing up in Britain in the 1970s and 1980s. In fact the 1958 cohort is, at the time of writing, almost aged 50 and the 1970 cohort nearing 40. Thus they tell us little (probably nothing) about children growing up in more recent policy environments. This is particularly the case if policy needs to be targeted toward early years as a number of influential authors have proposed (e.g. Carneiro and Heckman, 2003).

---

[1] The same requirement is obviously true for mobility measured in terms of social class as well.
[2] Using the British Household Panel Survey Ermisch and Nicoletti (2007) find evidence of a rise in the intergenerational elasticity over a similar period although less evidence of change based on the intergenerational correlation.

Seeking recent indicators of on changes in social mobility for more recent cohorts (who may not be of labour market age) is of paramount importance in the light of the Government's concerted policy focus on social mobility and young peoplesince 1997. The Government has directed additional funds into schools in inner-city areas through the Excellence in Cities Programme, provided pre-school services through Sure Start and has made substantial in-roads into reducing child poverty. Last year the then Secretary of State for Education anticipated the positive impact that these policies would have on intergenerational mobility.

"The progress we have made since 1997 – particularly at schools in deprived areas – means that there is every reason to expect that today's generation of poor children will have a much better chance to escape the limitations of their background."

Alan Johnson, Secretary of State for Education, 17th May 2007

In this paper we try to say something about the likely mobility patterns for younger cohorts than have been studied before. We do so by using a simple economic model which posits a link between a range of childhood outcomes and parental income which, under certain assumptions, provides a guide to subsequent levels of intergenerational income mobility. This is important since it enables us to study more recent changes in social mobility for the younger cohorts yet to reach adulthood. These, as already noted, are the most important for thinking about in terms of public policy design and implementation.

Our results show that the fall in mobility experienced across the 1958 and 1970 cohorts appears to have been an episode caused by the particular circumstances of the time such as the growth in income inequality and the uneven distribution of increased

educational opportunities. Social mobility worsened and took a step change downwards, leaving the UK near the bottom of the intergenerational league table of mobility, and on a different trajectory relative to other countries in the world where there is less evidence of changes over time (Blanden, 2008). This fall in mobility was accompanied by strong increases in educational inequalities (e.g. a very sharp rise in the association between educational attainment and family income and stronger links between test scores and behavioural measures and family income).

Looking at the changing connection between these earlier age intervening factors (education attainment, test scores, behavioural measures) and family income for more recent cohorts finds no evidence of change and thus it appears that the decline n social mobility may well have flattened out. However, at the same time, they have not reversed nor started to improve and, under reasonable assumptions, social mobility appears unlikely to improve for the current set of children when they do become of labour market age.

In the next section we present our analytical framework and describe the data. Section 3 revisits the results of our intergenerational mobility analysis for the 1958 and 1970 cohorts, and links these results into our new framework. Section 4 reports estimates of models for more recent cohorts and considers the implications of our findings for the evolution of intergenerational mobility. Section 5 concludes.

2. Analytical Framework and Data

Conceptual Framework

The extent of intergenerational income mobility is often measured by the following summary statistic, the coefficient β in the following statistical regression:

$$\ln Y_i^{children} = \beta \ln Y_i^{parents} + e_i \tag{1}$$

where $\ln Y_i^{children}$ is the log of some measure of earnings or income for children (when of adult age), and $\ln Y_i^{parents}$ is the log of the same measure of earnings or income of their parents, i identifies the family to which parents and children belong and e_i is an error term. β is then the elasticity of children's income with respect to their parents' income and $(1 - \beta)$ defines the extent of intergenerational mobility.

We want to estimate β consistently across time (over birth cohorts) to see if intergenerational mobility is getting better or worse. However, we cannot measure income for those cohorts not yet of labour market age, who are arguably the most important for considering policies to do with mobility.

We can, however, develop a model which studies earlier age outcomes and their relationship with parental income and examines how that maps into future intergenerational mobility. To see this, consider two life cycle stages, one which looks at how early age and childhood factors relate to parental income, and the other which looks at how income as an adult relates to these early age/childhood factors. In their simplest form, the two life cycle stages can be represented as:

Stage 1: The relationship between earlier age/childhood factors, Z, and family income: $Z_i = \theta \ln Y_i^{parents} + u_i$

Stage 2:   The relationship between child income (as an adult) and these earlier age factors $\ln Y_i^{child} = \lambda Z_i + v_i$

Here θ measures the sensitivity of the earlier age intervening factors Z to parental income and λ the income 'returns' to Z, (u_i and v_i are error terms). To be more concrete an obvious example would be if Z measures education, so that θ measures the sensitivity of education to parental income (stylistically 'how much more education children from rich backgrounds receive') and λ the income returns to education ('how much more the highly qualified earn'). Of course, Z may be a whole range of factors that correlate with parental income and/or yield an income return as adults and we consider different possibilities below.

Putting the two together by substituting the first stage into the second stage gives the intergenerational function:

$$\ln Y_i^{children} = \lambda\theta\ln Y_i^{parents} + \lambda u_i + v_i \qquad (2)$$

Comparing equation (2) with the standard intergenerational mobility function in equation (1) shows that $\beta = \lambda\theta$. However, it may be that v_i is also related to $\ln Y_i^{parents}$ so that $v_i = \alpha\ln Y_i^{parents} + w_i$. Substituting this into (2) gives:

$$\ln Y_i^{children} = (\lambda\theta+\alpha)\ln Y_i^{parents} + \varepsilon_i \qquad (3)$$

where the error term $\varepsilon_i = \lambda u_i + v_i$. In a regression context it is evident that $\alpha = \dfrac{Cov(v_i, \ln Y_i^{parents})}{Var(\ln Y_i^{parents})}$, which is the direct influence of parental income on children's earnings that does not come through Z_i.

Implementation

In the above model the intergenerational parameter is $\beta = \lambda\theta + \alpha$. For cohorts in adulthood we can estimate β and decompose the respective contributions of θ, λ and α. Importantly for our analysis we can study what happened to θ, λ and α in the period when β rose (i.e. mobility fell).[3] For younger cohorts we can only estimate θ but we can, under certain assumptions about how λ and α evolve over time, can estimate β even though the cohorts are too young to yet earn an income.

Our objective has similarities to the problem explored by Altonji, Bharadwaj and Lange (2007) in a recent paper. In their work Altonji et al wish to compare the wage distribution for the National Longitudinal Survey of Youth (NLSY) 1979 cohort in their forties with the likely wage distribution for the NLSY 1997 cohort (which like our later cohorts has not yet entered the labour market) at the same age.

Altonji et al observe w (wages) and z (characteristics correlated with wages) for the 1979 data but only z for the 1997 data. The wage distribution conditional on z for NLSY79 is estimated, this gives the relationship between z and w for the first cohort. By assuming that the conditional distribution of w given z does not change the data can be reweighted (drawing on the approach of DiNardo, Fortin and Lemieux (1996)) to replicate the distribution of z in the 1997 data and predict the future wage distribution for this cohort.

The two crucial assumptions implicit in Altonji et al's work are that the wage returns to z and the unmeasured influences on w do not change across the cohorts. The parallel assumption in our own work would be to assume that λ and α are unchanged. In our final section we experiment with these assumptions showing the impact on β of λ

---

[3] A similar decomposition approach is used in Blanden et al 2007, in that paper an extensive set of Zs are used to account as much of β as possible.

and α remaining constant and the effect of varying these parameters by a plausible amount.

Data

In the first part of our empirical work we review the evidence on changes in β and its components for older cohorts in the UK. In order to investigate intergenerational income mobility we need information on fathers' earnings or parental income as children are growing up and then information on the same children's earnings as adults. The 1958 (National Child Development Study or NCDS) and 1970 (British Cohort Study, or BCS) cohorts provide this information. These datasets selected all babies born in Britain in a single week in the springs of 1958 and 1970 respectively and rich information is obtained through childhood and into adult life. Data collection is ongoing, and data collected in 2004 has recently been released.

For both cohorts parental income data is obtained at age 16. It is more common in the literature to estimate intergenerational mobility using earnings from both generations, but information on the components of income is not available for the second cohort. In addition, using parental income is more representative of the household's resources and better reflects the contribution of women working outside the home.

We use adult earnings information for age 33 for the 1958 cohort and age 34 for the 1970 cohort (i.e. in 1991 and 2004). Blanden et al (2005) use earnings data for the BCS from age 30 but data from age 34 improves the comparability of data across the cohorts, and should lead to more accurate estimates of changes over time compared to previous work (see Haider and Solon, 2006, for a careful discussion of the possible impacts of lifecycle bias).

In order to compare the sensitivity of early age characteristics to parental income (θ) across the 1958 and 1970 cohorts and through to more recent data we need measures of Z that are comparable across a number of our data sources. We use degree attainment, early age test scores and measures of externalising behaviour. In almost all of our datasets mothers are asked a number of items from the Rutter A scale (this is the version of the Rutter behaviour scale which is asked of parents, see Rutter et al 1970). We combine selected items into an externalising behaviour measure by taking the first factor of a principal components analysis (as detailed in Table 1A and B of the appendix). A higher externalising score means children are more likely to 'act out' and misbehave. In the 1958 cohort we measure degree attainment at age 23[4], reading ability at age 11 and externalising behaviour at 7 and 11. In the 1970 cohort degree attainment is measured by age 23, reading at age 10 and externalising behaviour at 5 and 10.

To investigate the likely level of mobility for more recent cohorts we study cohorts born between 1970 and 2000; again using degree attainment, early age test scores and mother's reports of behavior as our Z variables. Our first source of information is the British Household Panel Survey. This survey began in 1991 and has collected evidence on 5000 households for all subsequent years. The longitudinal element of this data enables us to measure children's family income at age 16 and then to observe their educational achievements; here we consider whether they obtain a degree by age 23. As there are now 14 years worth of 16 year olds available we split these into younger and older pseudo cohorts.

As discussed above the 1958 and 1970 cohorts have made an important contribution to understanding mobility for those growing up in the 1970s and 1980s. The

---

[4] Degree attainment is measured at an early age as this can be replicated in the British Household Panel.

intergenerational story told by these data has been extended by collecting information on the children of the original cohort members. In 1991, data were collected about natural or adopted co-resident children for one third of the 1958 cohort members (those born in a week in 1958). 3000 children were included aged between 3 and 17 years old. Tests administered were the Peabody Individual Attainment test (for maths and reading) and the Peabody Picture Vocabulary test. The mother also answered a questionnaire providing more information on the behaviour and home environment of the children.

In 2004 a similar data collection exercise was conducted for the children of the 1970 cohort. In this case, data on children were collected for half of the cohort. Age-appropriate assessments of word and number skills from the British Ability Scales were carried out to gauge children's cognitive skills and attainment. Similar behavioural measures are taken from parents in both cohorts. Those that we use can be seen in Appendix Tables 1A and 1B.

Both datasets of the 'kids of' can be matched with information from the main surveys which provides details of parental education, family income and earnings, among numerous other characteristics. Information on family income is formed from adding together information on all of the cohort member's and their partner's sources of income (careful cleaning has been carried out here).

Another source of evidence which contains information on children's performance and their family background is the Millennium Cohort Study (MCS), which includes a large sample of children born in 2000 and 2001. The intention is to follow these children through life on a similar basis to the original cohorts. So far information is available at 9 months, 3 years and 5 years. Cognitive test scores and behavioural reports

are available at ages 3 and 5. We use percentiles of the vocabulary test as our main measure. Once again questions are asked about the children's behaviour which are highly comparable with those available from the other data sources we use and can be used to create a behavioural index.

In order to directly compare results from the 'kids of' datasets and the MCS we must consider a number of issues about the selection of the samples and the variables used. Among the issues we confronted were the selection of age groups within the 'kids of' data, the construction of an MCS sample to mimic the 'kids of' and the appropriate family income measure to use and the ethnic composition of the three datasets. There is more detail on these matters in the Appendix.

It should be noted that the 'kids of' datasets were not designed to be representative samples of all children in the age group and while we have tried our best to adapt the data in our youngest cohorts to be comparable the issue of representativeness introduces a caveat on our results. One useful aspect of the data we use is that the samples of the kids of BCS and the MCS were born only about one year apart. Therefore if we find similar results on these datasets we can be less worried about issues of comparability, The datasets we use in our comparative work are briefly summarised in the Table below. The Table emphasises why we need to look at earlier age intervening outcomes to say anything about more recent changes in social mobility. In almost all cases the N/A entries show that the most recent cohorts – the ones that are most relevant for contemporary policy debates – are characterised by missing information since these

cohorts are simply not old enough to have got to these stages in the life cycle.[5] Our two-stage framework allows us to discuss likely social mobility for these more recent cohorts.

Dataset	Year of birth	$\ln Y_i^{child}$ Earnings observed	Z Degree recorded	Tests taken	Behavioural questions answered
1958 Cohort	1958	1991	1981	1969	1965 and 1969
1970 Cohort	1970	2004	1993	1980	1975 and 1980
BHPS first pseudo cohort	1976 (average)	N/A	1999 (average)	N/A	N/A
BHPS second pseudo cohort	1980 (average)	N/A	2002 (average)	N/A	N/A
'Kids of' 1958 Cohort	1985 (average)	N/A	N/A	1991	1991
'Kids of' 1970 Cohort	1999 (average)	N/A	N/A	2004	2004
Millennium Cohort	2000-2001	N/A	N/A	2006	2006

* When making comparisons with the BHPS degree attainment in the cohorts is measured at age 23.

Measurement Issues

We take account of two important measurement issues in our estimations:

i) Changing income distributions:

Differences in the variance of $\ln Y$ between generations will distort estimates of β which is why inequality adjusted parameters have to be considered throughout (Solon, 1992). This is in fact the partial correlation between parents and children's status. This inequality adjusted measure of β is obtained simply by scaling β by the ratio of the

---

[5] Some of the older individuals used in our BHPS have earnings recorded in the data and could therefore be used to estimate intergenerational earnings mobility. We do not report these as intergenerational regressions on very young samples lead to βs that are strongly downward biased.

standard deviation of parents' income to the standard deviation of sons' earnings. [6] An alternative way of obtaining the correlation is to standardise income and earnings before estimating the intergenerational mobility regression (both variables are scaled to have a mean of 0 and a standard deviation of 1). We therefore use standardised parental income to account for changing inequality in our estimates of the relationship between intermediate outcomes and parental income.

ii) Permanent versus transitory income

Ideally we would seek to measure parents' and children's status with a measure of permanent incomes. A common approach to approximate this is to use income averaged over a number of periods (Solon, 1989, Mazumder, 2005). In cases where averaged income is not available we would be concerned that the measures of income available are not measured with equal accuracy over time. As shown by Solon (1992), and Zimmerman (1992) measurement error in parental income will lead to an attenuation of the estimated β and lead to difficulties in making correct inferences about changes over time. To avoid this we adopt a two stage procedure where supplement our measures of parental income with predicted income from a regression of income on more permanent 'income proxy' characteristics such as education, employment and housing status.

We thus estimate predicted income as $\ln \hat{Y}_i^{parents} = \hat{\delta} X_i$ where the $\hat{\delta}$ are coefficients from a first stage equation that relates family income to a range of income proxies, X_i, so that the two stage β_{2SLS} is estimated as

---

[6] $\text{Corr}_{\ln Y^{parents}, \ln Y^{son}} = \beta(\dfrac{SD^{\ln Y^{parents}}}{SD^{\ln Y^{son}}})$

$$\ln Y_i^{children} = \beta_{2SLS} \ln \hat{Y}_i^{parents} + \omega_i$$

This two-stage least squares (2SLS) approach has been shown to provide an estimate of β which is biased upwards compared to its true value if the characteristics used to predict income have an independent influence on children's outcomes (Solon, 1992). In this case we can think of β based on current income and β_{2SLS} as providing lower and upper bounds on the true estimate.

3. Recap of Existing Evidence on Changes in Mobility for Older Cohorts

The 1958 and 1970 Cohort Evidence on Changing Mobility

Tables 1a and 1b report information on intergenerational mobility for these cohorts in the form of transition matrices. The Tables show the proportion of sons in each parental income quartile that move into each quartile of their adult earnings distribution. We focus here on sons so that results are less directly influenced by women's labour market participation decisions. The extent of immobility can be summarised by an immobility index that computes the sum of the leading diagonal and its adjacent cells. The cells used to compute these are shaded in the Tables and the resulting index reported at the bottom. These numbers can be interpreted relative to the immobility index in the case of perfect mobility. If all individuals had an equal chance of experiencing an adult income in each quartile all cells would contain .25 and the immobility index (the sum of the diagonal band) would be 2.5.

The Tables show a fall in intergenerational mobility across cohorts. In all cases there is a higher probability of sons remaining in the same quartile as their parents in the 1970 cohort compared with the 1958 cohort. For example, for sons in the bottom quartile the proportion remaining in the bottom quartile is 30 percent for the 1958 cohort and 37

percent in the 1970 cohort. Equally, there is a smaller chance of large movements in the second cohort with 18 percent rising of those who start in the quintile bottom moving to the top quintile in the 1958 cohort and 13percent doing so in the 1970 cohort. The immobility indices reflect this fall in mobility, at 2.81 and 2.95 respectively.

In Table 2 we provide estimates of average mobility for sons and daughters, again using earnings data at age 33 for the first cohort and at age 34 for the second.[7] The average measure of intergenerational mobility shows a sharp (and statistically significant) fall, with β estimated as 0.21for the 1958 cohort of sons and 0.33for the 1970 cohort of sons. As shown in the second row estimates of β_{2SLS} are higher in both cohorts and also rise strongly across the cohorts. As discussed in Section 2 care must be taken in interpreting both β and β_{2SLS} when income inequality is changing. Columns 3 and 4 in each row show the partial correlation for each cohort with the final cohort showing the difference over time. This adjustment does not change our conclusions: the evidence indicates that intergenerational mobility has fallen across these cohorts for sons. In qualitative terms, Panel B shows similar qualitative results for women, although the changes are slightly smaller in magnitude.

The Influence of Intervening Factors

Table 3 summarises the association between family income at age 16 and several intermediate outcomes for the original cohort members. This reveals a sharp rise in the association between family income and degree attainment with the linear probability coefficient on standardised income rising from 0.05 to 0.12. This rise in the inequality of access to higher education has been explored in detail elsewhere (Blanden and Machin,

---

[7] Note that the figures presented here for the 1970 cohort differ from those in Blanden et al (2005) and Blanden et al (2007) because we are now using the most up-to-date earnings information.

2004) and shown to have a made a substantial contribution to the fall in intergenerational income mobility discussed above (Blanden, Gregg and Macmillan, 2007).

As shown in the second panel of Table 3 there is also a statistically significant increase in the association between percentile in reading test and parental income. The Ordinary Least Squares specification shows that in the 1958 cohort a one standard deviation increase in income is associated with a 5.65 point increase in test score percentile, while in the 1970 cohort a one standard deviation increase in income leads to an 8.72 percentile increase. The third panel reveals the association between externalising behaviour, using tobit models to account for the clustering at the lowest level of behavioural problems. As we would expect, as a higher score indicates worse behaviour, there is a negative association between family income and the externalising score (scaled to have a mean of zero and standard deviation of one). At both younger and older ages the association between behaviour and parental income has grown significantly across the cohorts. We know these 'non-cognitive' traits tend to be related to later labour market outcomes (Heckman, Stixrud and Urzua, 2006) and if this association has not declined over cohorts then this rising association will contribute to the change in intergenerational mobility.

As with the intergenerational income regressions in the previous section all our models are also estimated using the 2SLS framework. This confirms the findings of rising associations with income. It is reassuring that the findings are consistent across all the variables used. The difficulty with comparisons based on reading scores in these datasets is that the tests used are not identical, so it might be that they are picking up different skills. In this case we use also two variables that we know are comparable

across the cohorts (degree and externalising behaviour) meaning that it is not necessary to rely on test scores alone.

In our analysis of intergenerational income mobility in these two cohorts we explored both linear (regression) models and nonlinear (transition matrix) approaches. We can do the same for our intermediate outcomes. Table 4 presents the relationship between intermediate outcomes and age 16 income in an alternative form, by showing the mean of the outcome within the top and bottom income quintiles and then the difference between these, which we describe as 'inequality'.

Once again, there is a clear expansion of inequality by parental income across the cohorts. For the poorest 20 percent in terms of parental income 5 percent of the 1958 cohort achieved a degree, this compares with 20 percent for the richest fifth. Comparable figures for the 1970 cohort are 7 percent and 37 percent. Inequality in degree attainment has therefore widened from 15 percentage points to 30 percentage points.

The second set of results provides test score percentiles by income group. The gap between the richest and poorest groups here is 16 percentiles for 11 year olds in the 1958 cohort and 25 percentiles for 10 year olds in the 1970 cohorts. Again we find a substantial (and statistically significant) expansion in inequality. This pattern is replicated in the results for externalising behaviour where inequality at both young ages (7 and 5) and mid-childhood (11 and 10) increases by about one fifth of a standard deviation between the cohorts.

4. Evidence for More Recent Cohorts

More recent evidence on cross-cohort changes in the relationship between Z and family income

16

Table 5 shows a similar analysis to the earlier Table 3 for the more recent data. These estimates are based on the most comparable approaches to using the data for the 'kids of' and the MCS data. Estimations using alternative approaches are discussed in the appendix.[8]

The upper panel of the Table reports coefficients on income from linear probability models of obtaining a degree by age 23. These are shown for the BCS data (those born in 1970) and then for those in the BHPS born on average in 1976 and 1980. Unlike the previous cross-cohort comparison of Table 3, where the income coefficient rose steeply, there is no evidence of change for these cohorts.

The middle panel considers test scores, for the Ordinary Least Squares regression models there is no evidence that there has been a substantial change in the relationship between income and test scores, with a 1 standard deviation change in income leading to a 5-6 percentile change in reading/vocabulary score.[9][10] Notice that the estimates of the test-score income relationships in these recent cohorts around age 5 tends to be a little lower than for the 1970 cohort at age 10, this does not necessarily indicate that this relationship has declined, but more likely reflects the increasing influence of family background on attainment as children age (Feinstein, 2003 and Carneiro and Heckman, 2003). In the lower panel of Table 5 we show the relationship between family income and behavioural measures at around age 5, again there is no significant change.

---

[8] It should be noted that sample sizes for the 'kids of' and BHPS are rather smaller than those for the original cohorts. We are therefore not able to be as accurate in our estimates as is made clear by the larger reported standard errors.

[9] It is the case that the 2SLS for the MCS show a rise in the association between income and test scores compared to the 'kids of' data. We are reluctant to make too much of this as this appears to be an outlying estimate.

[10] If income is not standardised the test score income coefficients (and standard errors) are 8.44 (2.71) for the kids of the NCDS, 8.79 (2.05) for the kids of the BCS and 5.95 (0.84) for the MCS. The relative difference between the MCS and other two cohorts appears to be generated by the wider distribution of income in this data.

More recent distributional analysis

A distributional analysis is given in Table 6. There is much less evidence of a rise in the link between family income and degree attainment in more recent periods. The small increases in graduation rates that occurred for the (on average) 1975 cohorts compared with the 1970 group were evenly distributed across young people from different family income groups, consequently, there was no evidence of a strong widening of educational inequality. Comparing across the two BHPS samples there is a very slight widening of educational inequality with graduation rates among the poorest income groups dropping from 11 to 10 percent and graduation rates among the richest 20 percent growing by 4 percent. However, the small samples in the BHPS mean that we cannot draw strong conclusions from these small changes; the summary should be thought of as 'no evidence of change'.

This is also the case for the results based on test scores and externalising behaviour, in the original cohorts test score inequality grew by 9 percentage points from 1969 to 1980, from 1991 to 2006 the change was around 1 percentage point and not statistically significant. There was no change in the inequality of behaviour scores by parental income from 1991 to 2006.

Taking the linear and nonlinear results together it seems that all the large increases in educational inequality occurred between the 1958 and 1970 cohorts at the same time as the changes in intergenerational mobility. Unless changes in intergenerational mobility have been driven by very different forces in more recent years,

these results suggest that we might expect to observe little change in intergenerational income mobility for the cohorts born from around 1970[11] onwards.

Simulations to Predict Changes in β across More Recent Cohorts

In Section 2 we reviewed the relationship between θ, the sensitivity of Z to parental income, and β, the intergenerational parameter of real interest. In discussing our results so far we have assumed that future patterns in β for recent cohorts will mirror observed changes in θ. As noted in Section 2 this carries the implicit assumption that the returns to Z (λ) and the direct effect of parental income on earnings (α) will remain unchanged for our more recent cohorts.

The first panel of Table 7A lays out explicitly the relationship between inequality adjusted β, θ, λ and α for test scores for the older cohorts for whom all four parameters are observable. As noted in earlier results the substantial rise in β was accompanied by an increase in the sensitivity of test scores to parental income, θ. It is also clear that the direct relationship between parental income and sons' earnings (α) increased over this period. This is to be expected, it means that the relationship with parental income is rising with unmeasured attributes at the same time as it is rising with measured attributes (in this case test scores).

The second panel predicts the change in β between cohorts born around 1985 (the kids of the NCDS) and those born in 2000 (the MCS) using estimated θs and assuming that λ and α remain at the same level as they were for the BCS. Exactly as we would expect, under these assumptions no increase in persistence is predicted, with β falling slightly from .31 to .29.

---

[11] Note that our data does not allow us to pinpoint precisely when the fall in mobility stopped.

19

In the third and fourth panels we explore the changes in λ and α that would be necessary for there to be the same (annualised) change in adjusted β between 1985 and 2000/2001 as there was between the 1958 and 1970 cohorts. As the gap between these cohorts is a little larger, so is the equivalent change, at 0.17. For this change to be generated entirely by increases in returns to test scores these would have to rise four-fold, an extremely unlikely scenario. For α to generate these changes it would need to rise by 0.17, again this is unlikely, particularly over a period when there is no evidence that the relationship between parental income and test scores or behaviour increased.

Table 7B performs the same exercise considering degree attainment by age 23 as the driving mechanism behind intergenerational persistence. In this case the relevant cohorts to explore are the changes between the two BHPS periods, therefore focusing on sons born on average in 1975 and 1979. As before if λ and α are unchanged across the cohorts then β also remains unchanged.

The annualised increase equivalent to the 1958 and 1970 change is 0.04 over the four years between the two BHPS cohorts. In the third and fourth panels we once again explore the changes in λ and α that would be necessary to generate this change and find them to be implausibly large. In particular, the earnings differential associated with a degree would need to increase from 0.84 to 1.50, when in fact recent estimates show no rise in the returns to a degree among cohorts recently entering the labour market (O'Leary and Sloane, 2005).

5. Conclusion

In this paper we have presented evidence on changing patterns of intergenerational mobility for more recent cohorts of people than have previously been

considered in the literature on intergenerational mobility in Britain. We think this is important as these more recent cohorts (born after 1970) are of most relevance for contemporary (and future) discussions about public policy design and implementation.

Studying more recent cohorts presents some difficult modelling issues, not least the fact that many of the cohorts we study are not yet old enough to earn an income. We thus present a modelling framework where one can, under certain assumptions, say something about more recent patterns of changing mobility. We compare and contrast the findings from this framework for the post-1970 cohorts, with results from the two earlier birth cohorts (born in 1958 and 1970) for which we have earnings data.

Our results show that the widely quoted fall in mobility experienced across the 1958 and 1970 cohorts appears to have been an episode where social mobility worsened and took a step change downwards. However, results from our more up-to-date data show that this decline is not likely to have continued. However, at the same time, mobility patterns have not reversed nor started to improve and mobility appears to be set to remain at the low level seen for the 1970 cohort, at least for cohorts born up to 2000.

References

Altonji, J., P. Bharaduraj and F. Lange (2008) 'Changes in the Characteristics of American Youth: Implications for Adult Outcomes' NBER Working Paper No. 13883.

Blanden, J., A. Goodman, P. Gregg and S. Machin (2005), 'Changes in intergenerational income mobility in Britain', in M. Corak (ed.), *Generational income mobility in North America and Europe*, Cambridge, MA.: Cambridge University Press.

Blanden, J. and S. Machin (2004) 'Inequality in the Expansion of Higher Education' with Stephen Machin. Scottish Journal of Political Economy Special Issue on Education, 51, pp. 230-249.

Blanden, J., P. Gregg and L. Macmillan (2007) 'Accounting for Intergenerational Income Persistence: Noncognitive Skills, Ability and Education' Economic Journal, 117, C43-C60.

Blanden, J. (2008) 'Intergenerational Income Mobility in a Comparative Perspective' in P. Dolton , R. Asplund and E. Barth *Education and Inequality Across Europe*, Edward Edgar, Forthcoming.

Carneiro, P. and J. Heckman (2004) 'Human Capital Policy' in *Inequality in America: What Role for Human Capital Policies* ed. J Heckman and A. Krueger, MIT Press, Cambridge.

DiNardo, J., N. Fortin and T. Lemieux (1996) 'Labour Market Institutions and the Distributions of Wages, 1973-1992: A Semiparametric Approach' Econometrica, 64, pp. 1001-1044.

Ermisch, J. and C. Nicoletti (2007) 'Intergenerational Earnings Mobility: Changes Across Cohorts in Britain' B.E. Journal of Economic Analysis and Policy, Vol 7 Issue 2 (Contributions) Article 9.

Feinstein, L. (2003) 'Inequality in the Early Cognitive Development of British Children in the 1970 Cohort', Economica, 70, pp. 73-98.

Haider, S. and G. Solon (2006) 'Lifecycle Variation in the Association between Current and Lifetime Earnings', American Economic Review, 96, 1308-1320.

22

Heckman, J., J. Stixrud and S. Urzua (2006) 'The Effects of Cognitive and Noncognitive Abilities on Labour Market Outcomes and Social Behaviour' Journal of Labour Economics, 24, 411-482.

Mazumder, B. (2005) 'Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data.' The Review of Economics and Statistics, Volume 87, pp. 235-255,

Office for National Statistics (2004) *Birth Statistics* available at http://www.statistics.gov.uk/downloads/theme_population/FM1_33/FM1_33.pdf

O'Leary, N. and P. Sloane (2005) 'The Changing Wage Return to an Undergraduate Education' IZA Discussion Paper No. 1549.

Rutter, M., Tizard, J. and Whitmore, K. (1970) *Education, Health and Behaviour*, London: Longman.

Solon, G. (1989), 'Biases in the Estimation of Intergenerational Earnings Correlations', Review of Economics and Statistics, Vol. 71, pp. 172-174.

Solon, G. (1992) 'Intergenerational Income Mobility in the United States.' American Economic Review Vol. 82, pp. 383-408.

Zimmerman, D. (1992) 'Regression Toward Mediocrity in Economic Stature', American Economic Review, Vol. 82, pp. 409-429.

Table 1a : Intergenerational Income Mobility
Transition Matrix for the 1958 Cohort

Parental income	Sons' quartile			
	Lowest quartile	2nd quartile	3rd quartile	Top quartile
Lowest quartile	.30	.29	.24	.18
2nd quartile	.31	.27	.24	.19
3rd quartile	.22	.25	.25	.28
Top quartile	.18	.20	.27	.35

Notes: Sample size 2163; Immobility Index 2.806.

Table 1b :
Intergenerational Income Mobility Transition Matrix for the 1970 Cohort

Parental income	Sons' quartile			
	Lowest quartile	2nd quartile	3rd quartile	Top quartile
Lowest quartile	.37	.27	.22	.13
2nd quartile	.29	.30	.24	.17
3rd quartile	.22	.25	.28	.25
Top quartile	.13	.18	.24	.45

Notes: Sample size 1703; Immobility Index 2.955.

Table 2: Linear Estimates of Intergenerational Income Persistence

	Intergenerational Elasticities, β Parental income to child's earnings		Intergenerational Partial Correlations Parental income to child's earnings		
	1991 (1958 Cohort, Age 33)	2004 (1970 Cohort, Age 34)	1991 (1958 Cohort, Age 33)	2004 (1970 Cohort, Age 34)	Cross-Cohort Change in Inequality Adjusted β
A. Sons					
OLS	.21 (.03)	.33 (.03)	.17 (.02)	.30 (.02)	.13 (.03)
2SLS	.33 (.04)	.50 (.05)	.27 (.04)	.45 (.04)	.18 (.05)
B. Daughters					
OLS	.36 (.05)	.43 (.05)	.17 (.02)	.25 (.02)	.08 (.03)
2SLS	.55 (.08)	.63 (.07)	.26 (.04)	.37 (.04)	.11 (.06)

Notes: The OLS regressions include controls for parental age. Instrumental variables used are mother and fathers' education, employment status and housing tenure at age 16. Sample sizes are 2163 for the 1958 cohort and 1703 for the 1970 cohort.

Table 3: Associations between Intermediate Outcomes and Parental Income in the Cohorts

	Standardized Log(Income) Sensitivities, θ		
	1958 Cohort	1970 Cohort	Cross-Cohort Change in θ
Degree by Age 33/30, OLS	0.05	0.12	0.07
	(0.01)	(0.01)	(0.01)
Degree by Age 33/30, 2SLS	0.07	0.21	0.14
	(0.02)	(0.01)	(0.02)
Test Scores (Age 11, 1958 Cohort; Age 10, 1970 Cohort), OLS	5.65	8.72	3.07 (.49)
	(0.32)	(0.37)	
Test Scores (Age 11, 1958 Cohort; Age 10, 1970 Cohort), 2SLS	11.89	15.93	4.04 (.85)
	(0.56)	(0.64)	
Behavioural (Age 11, 1958 Cohort; Age 10, 1970 Cohort), OLS	-0.06	-0.12	-0.06
	(0.01)	(0.02)	(0.02)
Behavioural (Age 11, 1958 Cohort; Age 10, 1970 Cohort), 2SLS	-0.10	-0.18	-0.08
	(0.017)	(0.02)	(0.03)
Behavioural (Age 7, 1958 Cohort; Age 5, 1970 Cohort), OLS	-0.06	-0.15	-0.09
	(0.01)	(0.02)	(0.02)
Behavioural (Age 7, 1958 Cohort; Age 5, 1970 Cohort), 2SLS	-0.13	-0.22	-0.09
	(0.02)	(0.021)	(0.03)

Notes:
OLS estimates condition on parental age and the sex of the child. Instrumental variables used are mother and fathers' education, employment status and housing tenure at age 16. The degree coefficients are estimated using a linear probability model, although marginal effects from a probit model are almost identical. The behavioural score models are fitted using a tobit as in all cases around 15% of cases have the lowest score. Parental income data is standardised to have mean 0 and standard deviation 1. Sample sizes: Row 1: 7233; 5544. Row 2: 7766; 5983. Row 3: 7580; 6296. Row 4: 7709; 5616.

Table 4: Inequalities in Intermediate Outcomes by Parental Income in the Cohorts

		Lowest 20 Percent of Family Income	Middle 60 Percent of Family Income	Highest 20 Percent of Family Income	Inequality
Degree Acquisition	1981 (1958 Cohort Age 33)	0.07	0.10	0.24	0.17 (0.01)
	1993 (1970 Cohort Age 30)	0.11	0.15	0.44	0.33 (0.018)
Cross-Cohort Change					0.15 (0.02)
Test Score Percentile	1969 (1958 Cohort Age 11)	42.59	49.16	58.81	16.22 (1.05)
	1980 (1970 Cohort Age 10)	38.12	50.35	63.44	25.32 (1.10)
Cross-Cohort Change					9.10 (1.51)
Externalising Behaviour Score	1969 (1958 Cohort Age 11)	0.042	0.03	-0.11	-0.15 (0.03)
	1980 (1970 Cohort Age 10)	0.15	-0.033	-0.18	-0.33 (0.04)
Cross-Cohort Change					-0.19 (0.05)
Externalising Behaviour Score	1965 (1958 Cohort Age 7)	0.06	0.03	-0.11	-0.16 (0.04)
	1975 (1970 Cohort Age 5)	0.20	-0.04	-0.20	-0.40 (0.04)
Cross-Cohort Change					-0.24 (0.06)

Notes: Sample sizes are as for Table 3. Standard errors are in parentheses.

**Table 5: Associations between Intermediate Outcomes and
Parental Income in More Recent Cohorts**

	Standardised Log(Income) Sensitivities, θ			
	1970 Cohort (Age 23 in 1993)	BHPS (Age 23 in 1998)	BHPS (Age 23 in 2002)	Cross-Cohort Change in θ (1993-2002)
Degree by Age 23, OLS	0.11 (0.006)	0.10 (0.02)	0.09 (0.02)	-0.02 (0.02)
Degree by Age 23, 2SLS	0.19 (0.01)	0.23 (0.04)	0.19 (0.04)	0.00 (0.04)
	'Kids of' 1958 Cohort (Aged 5-7 in 1991)	'Kids of' 1970 Cohort (Aged 4-6 in 2004)	MCS (Aged 5 in 2006) Comparable	Cross-Cohort Change in θ, 1991-2006
Test Scores, OLS	5.17 (1.66)	5.63 (1.31)	5.66 (0.80)	0.49 (1.83)
Test Scores, 2SLS	7. 72 (1.98)	7.99 (1.65)	13.91 (1.28)	5.59 (2.55)
Behavioural, OLS	-0.21 (0.06)	-0.14 (0.04)	-0.18 (0.03)	0.03 (0.07)
Behavioural, 2SLS	-0.29 (0.07)	-0.28 (0.06)	-0.39 (0.05)	-0.10 (0.09)

Notes:
OLS estimates condition on parental age and the sex of the child. Instrumental variables used are mother and fathers' education, employment status and housing tenure at the time the income variable is observed.
Standard errors are displayed in parentheses. The degree coefficients are estimated using a linear probability model. As it is necessary to weight the MCS data we report regression models for the behavioural scores. Tobit models for the 'kids of' data can be found in the Appendix. 'Kids of' are restricted to the children of female cohort members. Sample sizes from left to right: Row 1: 4706; 725; 363. Row 2: 384, 541, 2661. Row 3: 366, 545, 2585.

Table 6: Inequality in Intermediate Outcomes by Parental Income in More Recent Cohorts

		Lowest 20 Percent of Family Income	Middle 60 Percent of Family Income	Highest 20 Percent of Family Income	Educational Inequality
Degree Acquisition	1993 (1970 Cohort Age 23)	0.07	0.15	0.37	0.30 (0.01)
	1998 (BHPS Age 23)	0.11	0.23	0.40	0.30 (0.05)
	2002 (BHPS Age 23)	0.10	0.21	0.44	0.34 (0.05)
Cross-Cohort Change (1993-2002)					0.04 (0.07)
Test Scores	1991 ('Kids of' 1958 Cohort Age 5-7)	38.39	52.84	52.74	14.35 (4.73)
	2004 ('Kids of' 1970 Cohort Aged 4-6)	40.76	50.86	56.00	15.24 (3.92)
	2006 (MCS Comparable Age 5)	38.96	48.75	55.61	16.65 (1.98)
Cross-Cohort Change (1991-2006)					2.30 (5.13)
Behavioural	1991 ('Kids of' 1958 Cohort Age 5-7)	0.290	0.06	-0.21	-0.50 (0.13)
	2004 ('Kids of' 1970 Cohort Aged 4-6)	0.21	-0.09	-0.14	-0.34 (0.14)
	2006 (MCS Comparable)	0.20	-0.07	-0.32	-0.52 (0.07)
Cross-Cohort Change (1991-2006)					-0.01 (0.15)

Notes: MCS data is weighted using longitudinal weights. Sample sizes are as for Table 5. Standard errors are provided in parentheses.

Table 7A: Simulations Based on Test Score Sensitivities for Sons

	Older Cohorts - Actual Changes When Can Estimate θ, λ and α		
	θ, λ and α	β	Cross-Cohort Change in β
NCDS 1958	θ = 3.29, λ =0.01, α = 0.13	0.17	
BCS 1970	θ = 6.86, λ = 0.01, α = 0.23	0.30	0.13
	Younger Cohorts - Simulate Changes When Can Estimate θ (Keep λ and α at BCS 1970 Values)		
Kids of NCDS, 1985	θ = 7.48, λ = 0.01, α = 0.23	0.31	
MCS, 2001	θ = 5.63, λ = 0.01, α = 0.23	0.29	-0.02
	Younger Cohorts - Simulate Changes When Can Estimate θ (Change λ so that β Rises by 0.17*)		
Kids of NCDS, 1985	θ = 7.48, λ = 0.01, α = 0.23	0.31	
MCS, 2001	θ = 5.63, λ = 0.04, α = 0.23	0.48	0.17
	Younger Cohorts - Simulate Changes When Can Estimate θ (Change α so that β Rises by 0.17*)		
Kids of NCDS, 1985	θ = 7.48, λ = 0.01, α = 0.23	0.31	
MCS, 2001	θ = 5.63, λ = 0.01, α = 0.40	0.48	0.17

* 0.17 is the annualised change between the 1985 and 2001 cohorts that would be equivalent to the 0.13 increase across the 1958 and 1960 cohorts.

Table 7B: Simulations Based on Degree Attainment Sensitivities for Sons

	Older Cohorts - Actual Changes When Can Estimate θ, λ and α		
	θ, λ and α	β	Cross-Cohort Change in β
NCDS 1958	θ = 0.04, λ = 0.72, α = 0.14	0.17	
BCS 1970	θ = 0.09, λ = 0.84, α = 0.22	0.30	0.13
	Younger Cohorts - Simulate Changes When Can Estimate θ (Keep λ and α at BCS 1970 Values)		
BHPS 1975	θ = 0.10, λ = 0.84, α = 0.22	0.31	
BHPS 1979	θ = 0.09, λ = 0.84, α = 0.22	0.30	-0.01
	Younger Cohorts - Simulate Changes When Can Estimate θ (Change α so that β Rises by 0.04*)		
BHPS 1975	θ = 0.10, λ = 0.84, α = 0.22	0.31	
BHPS 1979	θ = 0.09, λ = 0.84, α = 0.28	0.35	0.04
	Younger Cohorts - Simulate Changes When Can Estimate θ (Change λ so that β Rises by 0.04*)		
BHPS 1975	θ = 0.100, λ = 0.84, α = 0.22	0.31	
BHPS 1979	θ = 0.087, λ = 1.49, α = 0.22	0.35	0.04

* 0.04 is the annualised change between the 1974 and 1979 cohorts that would be equivalent to the 0.13 increase across the 1958 and 1960 cohorts.

Data Issues Appendix

'Kids of' Data

Our outcome measures is based on relative performance within age. For both datasets the impact of age in months within age in years is removed before converting reading/vocab test scores into percentiles within age in years. This relative measure is used in our analysis. This is useful as children of the 1970 cohort aged 3-5 are given a vocabulary test, whereas older children are tested in reading. By converting to percentiles we should have a comparable measure across all age groups. The PIAT reading score is available for all sampled children of the 1958 cohort from 5 onwards, and a small number of four-year-olds.

Our aim in this paper is to discover as much as possible about trends in intergenerational transmissions for recent cohorts of children. It is therefore essential to base our conclusions upon representative samples of children. The children of cohort members pose a difficulty in this regard, as while the initial sample of parents were representative of cohorts of births the children are not. In particular, the older children in the 'kids of' sample were born to younger parents, who are more likely to be more poorly educated or differ from other cohort members in other unobservable ways. In addition the pattern of this selection into fertility may be different for cohort members born in 1958 and 1970.

To evaluate this we can consider figures from Birth Statistics 2004 (ONS, 2004) which allow us to compare fertility rates of women born in 1958 and those born in 1970. By 1991 women equivalent to the 1958 cohort had given birth to 1711 children per 1000 women, while by 2004 the full cohort equivalent to the 1970 cohort had given birth to 1564 babies per 1000 women. Assuming that overall fertility is not declining across these cohorts this figures indicate that the 1970 cohort members are likely to be earlier in their child-bearing career than the 1958 cohort at the point when we observe them, despite being a year older. This indicates that it might be legitimate to compare slightly older children from the children of the 1958 cohort compared with the children of the 1970 cohort. Given that the 1970 cohort is one year older than the 1958 cohort at the time of the sample this means we would be sensible to compare 3-5 year olds from the BCS with 4-6 year olds in the NCDS and so on.

There is a further issue; as the sampling frame is co-resident natural or adoptive children of male cohort members are less likely to be included as they are more likely to live with their mothers. In fact 65 percent of the children of the 1958 cohort who are tested are in the sample because their mother is the cohort member, while in the 1970 cohort, 68 percent of the sample have cohort member mothers. This statistic also tends to vary by age, with the proportion of fathers much higher among younger children. We therefore concentrate on the children of female cohort members.

Appendix Table 2 shows the income association for different age groups from the 'kids of' data for a range of different models. The top-left panel gives the association between parental income and test scores using the OLS model. These comparisons indicate that the strength of the association between parental income and test score percentile is constant between the two 'kids of' datasets. 4-6 year old children of the 1970 cohort have a coefficient on family income in a test score regression of 5.63 compared with 5.17 for 5-7 year olds from the 1958 cohort. For slightly older children 5-7 year olds in the 1970 cohort have a coefficient of 6.05, compared with 6.38 for the next age group in the NCDS. This pattern of constancy continues up to the 6-8 age group in the 1970 cohort and the 7-9 year olds in the 1958 cohort. After this income coefficients in the BCS fall-off rapidly, with an insignificant association between family income and reading score among the 8-10 year olds. Fertility statistics indicate that children of women in the BCS over eight years old would make up just one third of the births by this age.

The top-right panel presents the coefficients for the same dependent variable from a 2SLS specification, as before income is predicted on the basis of parental education, employment and housing tenure. Once again, there is very little evidence of strong changes across the cohorts, with very similar coefficients across 1991 and 2004 for all age groups until we reach children of the 1970 cohort at age 7-9.

The lower panel shows the results for two alternative dependent variables, percentile in the maths/number tests and externalising behaviour score. The association between the maths test and parental income is lower overall than it was for reading but there is no strong evidence of changes over time. For the externalising score there is evidence of a slight decline in the association with parental income between 1991 and

32

2004 with the coefficient from the tobit regression at -.2 to -.3 in 1991 and at -.1 to -.2 in 2004.[12]

It therefore seems that comparisons of children of the 1958 cohort aged 5-7 and children of the 1970 cohort aged 4-6 might be appropriate for comparison with the MCS at age 5. We now discuss the best way of using the MCS data to make these comparisons legitimate.

A further problem resulting from the design of the 'kids of' datasets is that they tend to under-represent children from ethnic minority backgrounds compared with the population. Of the NCDS cohort members whose children were tested in 1991 98 percent were of white British ethnic origin, of the children of the BCS 93 percent were perceived by the interviewer to be of white British descent. As we shall see below the sampling frame for the MCS is considerably more diverse.

MCS Data

The MCS data provides information on a sample if children born from 2000-2001. We weight the data throughout to achieve a representative sample of children born in this period. We do not additional weights to adjust for attrition as these are not available for the cohort datasets. As noted above, the 'kids of' are representative of children of mothers born in certain years. In order to check what difference this makes we also limit the MCS data to those children with mothers aged 33-35.

There are several ways of constructing parental income from the questions available in the MCS. The most straight-forward is to use the categorical variables. Parents are asked to indicate the category that their total take-home income falls into (they are coded into weekly, monthly and annual amounts for the respondents' ease). The categories offered for the respondent to choose from vary depending on whether the child lives in a one or two parent family. We convert the categorical information into a continuous measure by treating income as the midpoint of the category stated. In addition to the categorical total income questions parents are also asked in detail about all the sources of income they have and the period to which each applies. Using these

---

[12] Additional results experimenting with non-standardised income, using children of all cohort members and income equivalising are available on request.

questions it is therefore possible to code up continuous net income measures (more similar to what is available in the 1991 and 2004 birth cohort surveys). We compare our results using different income variables.

The top left panel of Appendix Table 3 shows the test score – standardised income relationships for all in the MCS, with the panel below showing these estimates for the restricted sample. The estimates for the full sample show that a 1 standard deviation increase in parental income is associated with a 6 percentile increase in vocab score. Estimates for the smaller sample of age 33-35 mothers are slightly lower. For the behavioural scores the opposite is true with slightly stronger effects found among the subsample. Within these comparisons we also show the effect of using income based on parents' reported category of take home income and a constructed net income measure. These alternatives make little difference to the estimated coefficient.

As noted previously the Millennium Cohort Study was designed to reflect the UK's ethnic diversity. Indeed the over-sampling of wards with high populations from minority ethnic groups is an important reason why the data must be weighted. In the full sample 82 percent of the sample are white while in the weighted data this rises to 87 percent. It is clearly the case that the ethnic composition of the MCS and 'kids of' data are dissimilar and it is difficult to prove with certainty that this does not have an influence on the results. Experimenting with estimations using white children only for the MCS tends to indicate that inequalities are slightly narrowed when a less diverse group is considered. However this effect is not large enough to alter our substantive conclusions.

Appendix Table 1A: Elements Used to Make Externalising Scores

in Original Cohorts

1958 Cohort Age 7	1970 Cohort Age 5	1958 Cohort Age 11	1970 Cohort Age 10
Generally destructive	Destroys belongings (d027)	Destroys own, others, things	Destroys belongings (m45)
Squirmy, fidgety	Squirmy, fidgety (d026)	Squirmy, fidgety	Squirmy, fidgety (m44)
Irritable	Irritable (d032)	Irritable, quick tempered	Irritable (m50)
Fights other children	Fights with other children (d028)	Fights other children	Fights other children (m46)
Disobedient	Disobedient (d038)	Disobedient at home	Often disobedient (m56)
Temper tantrums	Child has temper tantrums (d009) Never in last year Not in last month Not in last week More than once a week	Not available	Not available
Never, sometimes, frequently	Does not apply, applies, certainly applies	Never, sometimes, Frequently	Answers are given on a 1-100 scale, amd recoded to give proportions in 'frequently, sometimes, never' to match NCDS at 11.

Appendix Table 1B: Elements Used to Make Externalising Scores
in More Recent Cohorts

Kids of 1958 Cohort aged 4-6	Kids of 1958 cohort aged 7+	Kids of 1970 cohort	MCS Age 5
Restless or overly active (n518219)	Restless, has difficulty staying seated long (n518345)	Restless, overactive over past 6 months (q1b)	Restless, overly active, cannot stay still (cmsdro)
Stubborn, sullen or irritable (n518330)	Irritable and is quick to 'fly off the handle' (n518352)	Temper tantrums in last 2 months (q1e)	Often has temper tantrums (cmsdtt)
Bullies or is cruel to other children (n518321)	Bullies other children (n518362)	Child often had fights or bullied over children in last 6 month (q1l)	Fights with or bullies over children (cmsdfb)
Disobedient at home (n518329)	Often disobedient (n518357)	Child has been generally obedient over last 6 months (q1g)	Child is generally obedient (cmsdor)
Not true, sometimes true, often true	Does not apply, applies somewhat, certainly applies	Not true, somewhat true, certainly true	Not true, somewhat true, certainly true

Appendix Table 2: Relationship between Test Scores and Income for Different Age Groups of 'Kids of' Female Cohort Members

	Kids of Cohort Members			
	Income coefficients from regression of percentile of reading/vocab tests – OLS		Income coefficients from regression of percentile of reading/vocab tests – 2SLS	
	1991	2004	1991	2004
Age group	1958 Cohort	1970 Cohort	1958 Cohort	1970 Cohort
3-5		5.69 (1.37) [567]		7.70 (1.86) [567]
4-6	1.93 (1.97) [277]	5.63 (1.31) [541]	5.90 (2.42) [277]	7.99 (1.65) [541]
5-7	5.17 (1.66) [384]	6.05 (1.31) [579]	7.72 (1.98) [384]	7.60 (1.83) [579]
6-8	6.38 (1.65) [355]	3.74 (1.38) [517]	8.24 (1.92) [355]	6.30 (1.99) [517]
7-9	5.95 (1.79) [328]	1.00 (1.46) [493]	9.04 (2.08) [329]	3.12 (2.09) [493]
8-10	5.58 (1.73) [335]	-.34 (1.54) [430]	7.58 (2.02) [335]	4.06 (2.17) [430]
	Income coefficients from regression of percentile of number test		Income coefficients from tobit model of externalising behaviour	
Age group	1958 Cohort	1970 Cohort	1958 Cohort	1970 Cohort
3-5		4.15 (1.38) [566]		-0.11 (0.05) [571]
4-6	3.13 (1.95) [281]	3.09 (1.33) [541]	-0.14 (0.07) [346]	-0.16 (0.06) [545]
5-7	5.12 (1.65) [387]	3.06 (1.33) [579]	-0.26 (0.07) [366]	-0.17 (0.06) [574]
6-8	3.28 (1.66) [358]	1.86 (1.40) [515]	-0.37 (0.07) [335]	-0.16 (0.06) [510]
7-9	2.06 (1.82) [329]	0.90 (1.47) [493]	-0.39 (0.07) [307]	-0.14 (0.06) [486]
8-10	1.99 (1.75) [335]	2.10 (1.55) [428]	-0.20 (0.07) [301]	-0.20 (0.07) [430]

Notes: Standard errors in parenthesis. Sample sizes in square brackets.
All regressions include controls for the child's sex, the cohort member's partner's age and a polynomial in the child's age in days at testing.

Appendix Table 3: Relationships between Test Scores and Parental Income in MCS

	ln(income) coefficient in regression of vocab score percentile		ln(income) coefficient in regression of externalising score	
Year	2006	2006	2006	2006
Sample	MCS kids Aged 5	MCS kids Aged 5	MCS kids Aged 5	MCS kids Aged 5
Income measure	Midpoint of income category	Continuous net income	Midpoint of income category	Continuous net income
OLS coefficient	6.03 (0.31)	5.87 (0.32)	-0.18 (0.01)	-0.16 (0.01)
2SLS coefficient	12.33 (0.46)	13.83 (0.54)	-0.38 (0.02)	-0.42 (0.02)
Sample	13448	13448	13003	13003

	Mums 33-35 years old			
	ln(income) coefficient in regression of vocab score percentile		ln(income) coefficient in regression of externalising score	
Year	2006	2006	2006	2006
Sample	MCS kids Aged 5	MCS kids Aged 5	MCS kids Aged 5	MCS kids Aged 5
Income measure	Midpoint of income category	Continuous net income	Midpoint of income category	Continuous net income
OLS coefficient	6.83 (0.71)	5.66 (0.80)	-0.20 (0.03)	-0.18 (0.03)
2SLS coefficient	12.93 (1.15)	13.91 (1.28)	-0.34 (0.04)	-0.39 (0.05)
Sample	2661	2661	2585	2585

Notes:
Standard errors in parenthesis.
All regressions include controls for the child's sex, parents' age and a polynomial in the child's age in days at testing.
Income variables are standardised.
The results for the MCS sample are limited to those who have valid observations for both income variables. Results change only slightly if this restriction is lifted.
As weights are required we use a regression model for the behavioural models rather than the preferred tobit model.