The University Of Sheffield.

Department Of Economics.

Sheffield Economic Research Paper Series.

Intergenerational Income Mobility in the UK: New evidence using the BHPS and Understanding Society

Authors: Bertha Rohenkohl

ISSN 1749-8368

SERPS no. 2019017

July 2019

Intergenerational Income Mobility in the UK:

New evidence using the BHPS and Understanding Society

Bertha Rohenkohl*

July 29, 2019

Abstract

Using a new dataset combining the British Household Panel Survey and Understanding Society, I estimate the intergenerational income elasticity in the UK for individuals born between 1973 and 1991. Employing the traditional OLS approach as well as an alternative two-stage residual method that better controls for life-cycle effects, my results indicate that the intergenerational income elasticity is approximately 0.25. This means that around one quarter of every additional 1% of income advantage enjoyed by parents is passed on to their children. I also estimate income rank coefficients, which are a measure of positional mobility in the income distribution and these results corroborate the analysis of elasticities. These main results are largely robust to changes in the specifications of the model, sample restrictions and to the use of different measures of income. I also obtain regional estimates of mobility, and find large differences between the North and South of England.

Key Words: intergenerational mobility, income dynamics, social mobility
JEL Classification: J62

* *Address for correspondence*: Department of Economics, The University of Sheffield, 9 Mappin Street, Sheffield, S1 4DT, UK. E-mail: brohenkohlcruz1@sheffield.ac.uk

I would like to thank my PhD supervisors Andrew Dickerson, Nicolas Van de Sijpe and Anita Ratcliffe for comments and continuous support, and participants at the SSSI 2019 Bergen and 8th ECINEQ Meeting 2019 for their feedback and suggestions.

1 Introduction

The rise in inequality in recent years has exacerbated social divisions and increased the perceived importance of social mobility for modern societies. Income mobility across generations relates to the extent to which individuals with different initial conditions are offered equal opportunities to succeed in life, and this is seen as an indicator of a fair and equal society. As a result, there has been a resurgence of interest in social mobility as a topic in academia, policy spheres and public domain.

In the UK, the current common belief is that the class society is quite rigid and that opportunities are unequally distributed. In one of her first speeches as UK Prime Minister in 2016, Theresa May said "I want Britain to be a place where advantage is based on merit not privilege; where it's your talent and hard work that matter, not where you were born, who your parents are or what your accent sounds like" (May, 2016). This statement reflects the importance of one specific dimension of social mobility, the intergenerational mobility, which refers to the extent to which an individual's economic or social success is determined by the socio-economic position of their parents. If there is limited intergenerational mobility, coming from a relatively more advantaged or disadvantaged family may define children's opportunities as adults and have a significant impact throughout their lives.

Previous studies have provided empirical evidence of the extent of intergenerational mobility for a select group of developed countries (reviewed in Solon (2002); Black and Devereux (2011); Jäntti and Jenkins (2015)). In fact, for a number of countries, there is a growing perception that relative earnings and income mobility across generations has declined in recent years (OECD, 2018a), with family background having an increasing importance in determining an individual's socio-economic status. However, generally speaking, conducting such studies can be a challenging task due to the extensive data requirements, as information is needed for at least two generations and most surveys do not persist over a long enough period to allow for this.

In spite of the clear importance of intergenerational income mobility to understanding the income dynamics of a society and its relevance to inform public policy, the empirical evidence for the UK is still relatively scarce. This paper aims to contribute to this small literature, by providing updated and robust estimates of intergenerational mobility in the UK.

Previous efforts to measure intergenerational mobility of income and earnings in the UK mainly used two cohort studies: the National Child Development Study (NCDS, born in 1958) and the British Cohort Study (BCS, born in 1970)(Dearden et al., 1997; Blanden et al., 2004; Gregg et al., 2017; Belfield et al., 2017). However, the presence of data limitations and the use of various methodologies to estimate intergenerational mobility of income and earnings mean that most UK studies have produced wide-ranging estimates and fairly inconclusive results until now.

In this paper, using a recent data set that spans over 26 years, I investigate the extent of intergenerational income mobility for adults born between 1973-1991. While previous UK studies were mostly focused on the NCDS and BCS cohorts, the intergenerational mobility for this younger generation has not yet been studied. The study of these younger cohorts is relevant for current and future policy design and implementation. The availability of new data with the British Household Panel Survey (BHPS), spanning from 1991-2008, and with the harmonised continuation of this data with the Understanding Society (UKHLS), from 2009-2016, allows me to measure the extent of intergenerational income mobility in the UK and test the sensitivity of the estimates to several methodological and empirical choices. To my knowledge, this is the first study to use this harmonised data set to estimate intergenerational mobility in the UK.

Employing the traditional OLS method from this literature and also an alternative two-stage residual approach, which allows for a more flexible control of the age-income profile of young individuals, I estimate the intergenerational income elasticity (IGE), for the UK. The main results suggest that the intergenerational income elasticity is in the range 0.25-0.27 and precisely estimated. This coefficient represents the fraction (0.25-

0.27%) of every additional 1% of parental income advantage (or disadvantage) that will be passed on to their descendants. These estimates are robust to changes in the sample's age and coresidency restrictions, to the use of different income measures in the estimation and to the treatment of outliers. In addition to estimating IGEs, following recent developments in the international literature (Dahl and DeLeire, 2008; Chetty et al., 2014), I also estimate rank coefficients, which aim to characterise intergenerational mobility from a positional perspective, analysing the persistence of percentile ranks of parents and children in the income distribution. The rank coefficient is estimated to be around 0.25-0.30, depending on the definition of income used. These estimates are also robust according to a number of tests.

Finally, this paper also contributes to the growing literature on the geography of opportunity initiated by Chetty et al. (2014), by examining differences in the intergenerational income mobility estimates across regions in the UK. I find a significant divide between the North and South of England, with the South being more mobile according to all obtained measures of relative mobility.

The remainder of the paper is structured as follows. In the next section I briefly review the relevant literature on intergenerational income mobility. Section 3 presents the existing evidence for the UK. Section 4 describes the data sources and main variables. Section 5 explains the estimation methods and contains the main results, robustness checks and discussion. Section 5.2 presents the regional estimates of intergenerational income mobility. The paper ends with a summary and conclusion in Section 6.

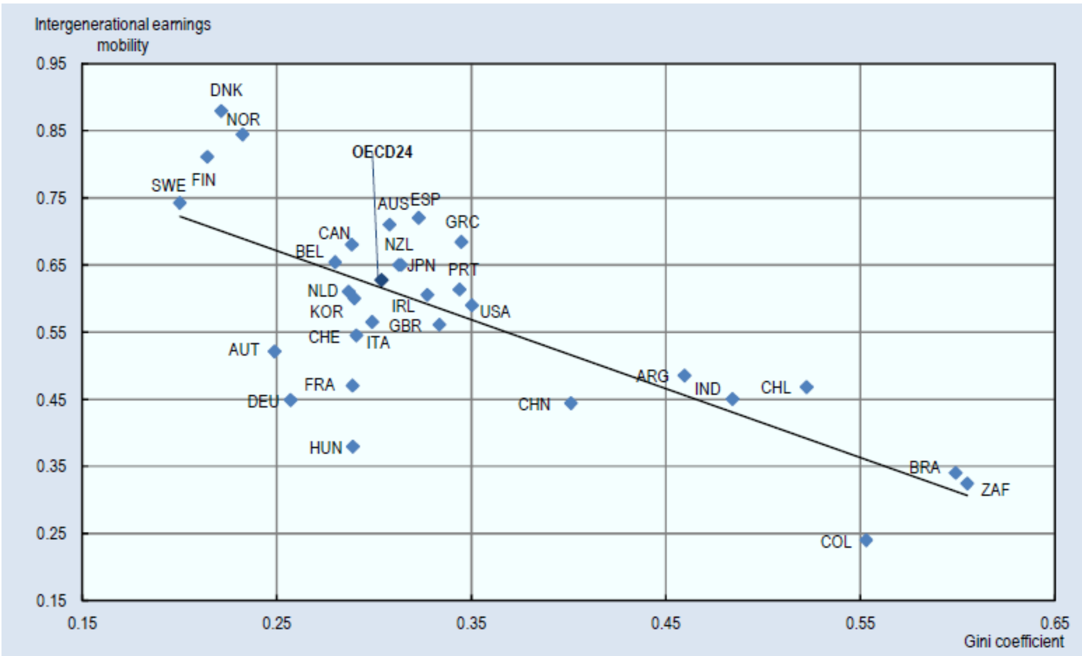
2 Intergenerational Mobility

2.1 Background

The study of intergenerational social mobility has been of interest to academics, economists, sociologists and policy-makers, as it seems to be a common goal of politicians and societies to ensure its citizens have equal opportunities for social advancement (OECD, 2018a) and to understand how this can be achieved. With the rising inequality in recent years, fears that inequalities will persist into future generations have increased. If a society is highly unequal, with low levels of mobility, inequalities will likely persist over time.

The relationship between inequality and mobility has been widely discussed in the literature and is commonly represented with the "Great Gatsby Curve"[1], shown in Figure 1. The figure is obtained by plotting the Gini coefficient against the intergenerational income/earnings elasticity. Although the theoretical link between inequality and income mobility across generations at a point in time is not clear (OECD, 2018a), the empirical evidence suggests a negative relationship between inequality and intergenerational earnings mobility across countries.

Figure 1: The Great Gatsby Curve



Notes: Intergenerational mobility is equal to 1 minus the intergenerational earnings elasticity. Income inequality measured by the Gini coefficient of the mid-1980s to early 1990s. Source: OECD (2018a)

Besides its link to the persistence of economic disparities over time, intergenerational mobility is also important for other reasons (OECD, 2018a; Narayan et al., 2018). Firstly, having some degree of intergenerational mobility is often considered one of the pillars of a

[1]This term was first used by Alan Krueger in a speech to the Center for American Progress on January 12, 2012.

fair and equal society, where fairness and equality of opportunities means that hard work and talent (or ability) are rewarded rather than purely inherited financial advantages. Secondly, low levels of intergenerational mobility are harmful to economic growth because if the future of individuals is strongly determined by their family background, this means that that human potential (especially from poorer backgrounds) is wasted and under-developed (OECD, 2018a; Narayan et al., 2018). The general perception that a society is 'unfair', and chances are unequal among individuals can reduce individuals' motivation to do their best and lead to further underinvestment in human capital and reduced productivity.

When discussing intergenerational social mobility, the role of earnings and income as a broad representation of socio-economic status is often highlighted in the literature. There is growing evidence that the income of the parents during childhood years has a significant influence on their offspring's socio-economic outcomes as adults.

Since the first studies of intergenerational mobility from the late 1970s, the extent to which socio-economic status (SES) is transmitted between generations and how this happens have been widely investigated (Solon, 1999; Black and Devereux, 2011; Jäntti and Jenkins, 2015). The theoretical explanation for intergenerational income mobility comes from the model by Becker and Tomes (1979, 1986). This model assumes that each family maximises a utility function across several generations and that parents can influence the future earnings capacity of their children directly, through investing time and money to provide them with human capital and non-human capital. This will, of course, depend on parents' own preferences and on their budget and credit constraints. In addition, the future income of children will also be (indirectly) influenced by other family-related 'endowments', such as genetic disposition, ability, culture, values, family connections and other skills, knowledge and goals provided by the family environment (Becker and Tomes, 1979, p. 1153). This mixture of human and non-human capital and other family-related endowments will be generically called 'family resources'.

While economists tend to focus on income, earnings or education as representative measures of permanent socio-economic status (Solon, 2002; Black and Devereux, 2011; Jäntti and Jenkins, 2015), sociologists have used social class and occupational status (Erikson and Goldthorpe, 2002). The first works on intergenerational mobility by economists focused on estimating intergenerational mobility of earnings and wages for pairs of fathers and sons, mainly because this was the data available at the time. Later on, other studies started focusing on distinct outcomes that are closely related to wages, such as educational and occupational mobility. Comprehensive reviews of these studies are provided by Solon (1999), Black and Devereux (2011) and Jäntti and Jenkins (2015).

Based on the idea that income may both influence and reflect many of the other factors directly (health, occupation, education, neighbourhood, etc) and may also be regarded as a broader proxy for long-run economic status, the focus of more recent studies has shifted to measuring intergenerational income mobility.[2] Parental income is often considered a good predictor of children's socio-economic status as they become adults (Lee and Solon, 2009; Chadwick and Solon, 2002; Chetty et al., 2014).

The main idea behind estimating intergenerational income mobility relates to the understanding that parents invest time and money into their children and the availability of family resources during childhood is an important determinant of children's outcomes as an adult. If children have less or more access to resources when they are growing up (at least relatively) this could have implications for their future economic success or failure. If this is the case, the income distribution of children would be more closely tied to those of their parents. Therefore, the analysis of income mobility encompasses the idea that family resources are important for a child's development.

[2] Income measures (rather than earnings or occupation) are not only a broader representation of economic status, but also more adequate to study mobility among women - for mothers and daughters - as this avoids issues of selection into labour force participation and part time work. Women are now increasingly important in the labour force, but were frequently excluded from intergenerational mobility studies due to a lack of information on their wages or earnings (Jäntti and Jenkins, 2015).

2.2 Estimation

The common empirical strategy to estimate intergenerational mobility consists of relating, across generations, a proxy of a measure of individual, family or household permanent income. The standard approach is to estimate the generational association between parent's and children's income with intergenerational elasticities (IGEs). The typical formulation of the model used to estimate IGEs is:

$$log(Y_i^{child}) = \alpha + \beta \; log(Y_i^{parent}) + \epsilon_i$$

Where Y_i^{child} is the measure of permanent income of individuals (children) and Y_i^{parent} is the measure of parental permanent income.

The coefficient of interest (β), called intergenerational income elasticity, is a measure of intergenerational persistence and represents the extent to which parental income is transmitted to the next generation. A larger coefficient indicates more persistence in incomes, or less mobility. On the other hand, $1-\beta$ is a measure of mobility. Referring back to the model by Becker and Tomes (1979, 1986), this elasticity includes a combination of factors: the direct investments in human and non-human capital, and also the broader inherited family 'endowments'. In addition, the intergenerational income elasticity is a measure of *relative*[3] immobility, or relative persistence (OECD, 2018a; Chetty et al., 2014). It compares the socio-economic outcomes (in our case, income) of children in families in different points of the social ladder, e.g. of children from richer families relative to children of poorer families.

Two common issues related to the estimation of IGEs have been emphasised in the literature: the transitory variation in observed income measures in the data and the life-cycle bias. The first issue relates to biases from inaccurately measuring 'permanent' income (Solon, 1989, 1992; Grawe, 2006). The information on 'permanent' (or long-run)

[3]This differs from the concept of *absolute* mobility, which indicates how much living standards (health, education, income) have improved or deteriorated across generations (i.e. between parents and their children).

income or earnings must be derived from the observed income variables present in the data sets - which are usually a measure of annual or monthly income or earnings.

Most early IGE estimates relied on single year measures of earnings or income due to data availability issues (Solon, 1992). However, the use of a short-run proxy for long-run status implies that 'permanent' income will be measured with error due to transitory fluctuations. As a result, there will be an errors-in-variables attenuation bias, leading to a lower estimated intergenerational income or earnings elasticity (Grawe, 2006; Solon, 1989, 1992), giving the impression of more mobility.

Typical corrections to this classical errors-in-variables problem in the literature include using a multi-year average of parental income observations[4] in order to reduce the transitory variation (Solon, 1992; Chadwick and Solon, 2002; Mazumder, 2005; Chetty et al., 2014; Gregg et al., 2017), or use an IV approach to predict 'permanent' parental income based on other parental characteristics (Solon, 1992; Dearden et al., 1997; Nicoletti and Ermisch, 2008). [5] In this paper, I use a rich longitudinal data set that allows me to calculate multi-year averages of parental income during childhood years and also of individual income during adulthood, in order to reduce the bias from transitory shocks to income.

The second issue to be aware of when estimating IGEs relates to the ages at which current incomes of parents and children are observed. The literature suggests that the relationship between current and lifetime ('permanent') income changes over the life-cycle (Haider and Solon, 2006). Thus, similarly to what has been described in the first issue regarding transitory shocks to permanent income, current earnings (income) might not be a good proxy for lifetime earnings (income), depending on the age at which they are

[4]Nonetheless, the results in Mazumder (2005) and Haider and Solon (2006) suggest that even estimates based on five-year averages of the earnings variable for fathers are subject to some attenuation bias. In fact, Mazumder (2005) shows with a simulation exercise that one would need approximately 20 to 25 years of income data in order to calculate a multi-year average that would be a very good proxy for the permanent component of earnings and obtain a high reliability rate, or in other words, have almost no attenuation bias.

[5]Examples of instruments used in the literature are indices of father's socioeconomic status (social class, occupation) and father's education. However, it would be important that the instruments do not explain children's earnings (i.e. the exclusion restriction is valid), otherwise the IV results would be biased upwards (Dearden et al., 1997; Solon, 2002), generating an amplification bias.

observed. Individual annual incomes tend to grow considerably between the ages of 20 and 30, get to a maximum and flatten between the ages of 40 and 50 and decline thereafter (Corak, 2004). This pattern of growth might be heterogeneous across individuals due to differences in the income/earnings profile and in human capital investments (Haider and Solon, 2006; Grawe, 2006; Nybom and Stuhler, 2016). As emphasised by Black and Devereux (2011) this is an important issue, as in practice it is likely that the current income data will be observed relatively late for parents[6] and relatively early for children.[7] If this is the case, this implies that the β would be underestimated.

Because the relationship between current earnings and lifetime earnings evolves over the life cycle and is age dependent, estimates of intergenerational elasticities will be sensitive to the age at which both children's and parental current incomes are observed. Fortunately, it seems that the life-cycle bias varies predictably across age (Grawe, 2006). The results from Haider and Solon (2006) and Nybom and Stuhler (2016) suggest that it is possible to mitigate the life-cycle bias when children's and parents' earnings and income are measured around mid-life.[8] Thus, in order to reduce the influence of measurement error from the life-cycle effects, incomes for both generations should be observed when they are most representative of permanent income and at similar point in the life-cycle.

In this paper, I use several income observations for parents measured around mid-life and child income is measured no earlier than the age of 25. I also use a two-stage residual approach that controls more flexibly for age effects and test the sensitivity of the main results to changes the minimum age at which child income is observed.

In addition to being subject to these two estimation issues, more recent studies have

[6]Grawe (2006) highlights the importance of father's age when income is observed. The author shows that intergenerational earnings persistence is negatively associated with the age at which father's earnings is observed. Assuming that sons are observed at some point in mid-life, we would observe a lower persistence (lower β) if parental income is observed at older ages. Grawe argues that 20% of the variance in IGE estimates among studies using similar methodologies and data can be attributed to differences in fathers ages when income is observed.

[7]If we observe earnings/income for all individuals in early-career years, before they had the chance to experience different growth rates, we will underestimate the gap between low and high earners relative to what it will be in mid-life and thus underestimate the degree of intergenerational persistence (β).

[8]Böhlmark and Lindquist (2006) apply the Haider and Solon model to Swedish data and find that total income and earnings have a similar evolution over the life-cycle.

argued IGEs might be limited estimates of relative mobility. This is because the relationship between child and parental income could be non-linear (or non-log linear) and because IGEs could be sensitive to the treatment of children with zero income (Bratsberg et al., 2007; Chetty et al., 2014) and to the use of different measures of income (Landersø and Heckman, 2017).

An alternative rank approach to estimate intergenerational mobility was proposed in Dahl and DeLeire (2008) and became well-known in this field after the influential paper by Chetty et al. (2014) that uses the rank estimates to compare income mobility across geographical areas in the United States. The use of rank-based measures focuses on the analysis of the correlation between parents' and children's rank position in the distribution of income, instead of looking at the values of income variables directly. The ranks for parents and children are constructed separately, based on their respective income distributions. The rank-rank slope, or rank coefficient represents the probability that a child's rank in the income distribution is higher (or lower) than their parents' income rank. It is also a measure of relative mobility, but it captures solely the extent of re-ranking across generations.[9]

The rank approach is a way of estimating mobility without assuming a (log)linear relationship between parental and children's incomes (Dahl and DeLeire, 2008). In addition, evidence from other studies shows that the rank-rank relationship is almost perfectly linear for a sample of countries[10], that it already allows for changes in inequality across generations (Bratberg et al., 2017) and that it is much less sensitive to specifications of the model and to attenuation and life-cycle bias, as it is scale invariant (Chetty et al., 2014; Mazumder, 2016; Gregg et al., 2017). In this paper, I also estimate rank coefficients in order to get a more complete picture of intergenerational income mobility in the UK, and subject these estimates to the same robustness checks as the intergenerational

[9]While IGEs combine both marginal and joint distributions capturing the extent of re-ranking across generations and the spread of the income distributions, rank measures focus solely on the re-ranking. As explained by Gregg et al. (2017), if the income distribution is represented by a ladder, re-ranking describes people switching rungs on the ladder and inequality describes how far apart the rungs of the ladder are.

[10]Germany, Sweden, Norway and the US.

elasticities.

3 Intergenerational Mobility in the UK

A number of studies have examined intergenerational mobility for a set of countries for which longitudinal data is available, such as the United States, Denmark, Sweden and Germany and the United Kingdom. Most previous research has suggested a variation in intergenerational earnings and income mobility across countries, and the UK is often ranked as a country with relatively low mobility (Solon, 2002; Björklund and Jäntti, 2011; Corak, 2004, 2006; Jäntti et al., 2006; Raaum et al., 2008).

Comparing the results and the extent of intergenerational mobility across independent studies and different countries constitutes a challenging task, as there are often several methodological choices and data limitations involved. A few studies have taken up on the challenge to provide reliable comparisons of intergenerational mobility across countries. Most have focused on comparisons of intergenerational elasticities (Solon, 2002; Corak, 2004, 2006; Jäntti et al., 2006; Raaum et al., 2008) and recently the OECD (2018a) has provided an updated ranking of IGEs (Figure 2).

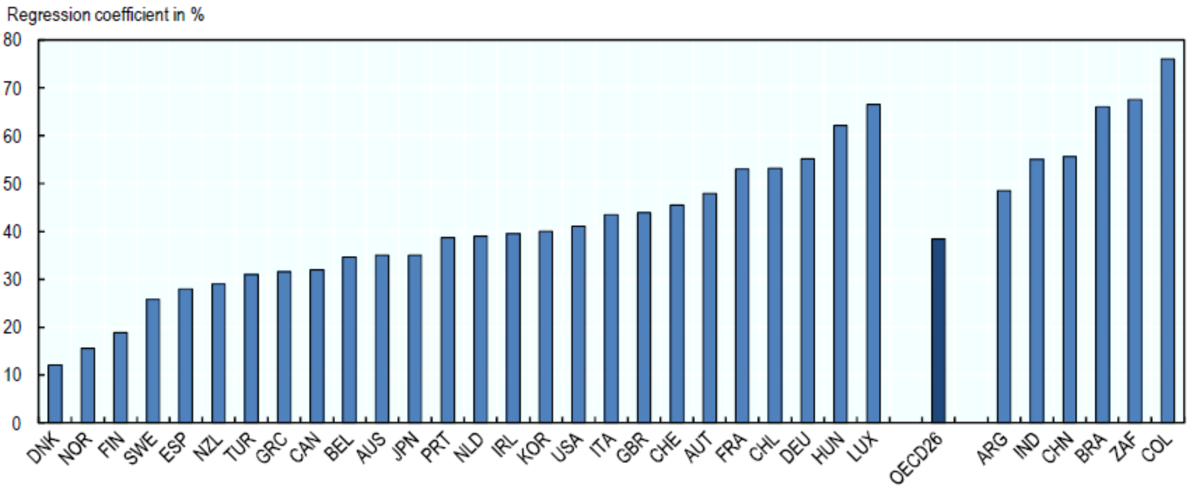
Overall, these studies characterise the UK as having relatively low intergenerational mobility, especially in comparison to the Nordic countries. Corak (2006) reports an IGE of son's earnings that lies in the interval 0.43-0.55, Blanden (2009) suggests it is around 0.37, while Jäntti et al. (2006) and Raaum et al. (2008) estimate it at 0.30 and 0.41, respectively. However, it is noteworthy that the existing evidence on intergenerational mobility for the UK is still very thin and "there is a lot of uncertainty for the UK" (Björklund and Jäntti, 2011, p.507). The UK position in the cross-national rank is based on a couple of selected studies[11] that use cohort surveys data when, in reality, the results produced by the literature are diverse and inconclusive.

The first study of intergenerational mobility in Britain[12] using nationally representa-

[11] Dearden et al. (1997) and Blanden et al. (2004).

[12] The very first evidence on intergenerational income mobility in Britain was presented by Atkinson

Figure 2: Intergenerational earnings elasticity between fathers and sons



Note: Each bar represents the point estimate of the intergenerational earnings elasticity for each country. A higher estimate means a higher persistence of earnings across generations, and lower intergenerational earnings mobility. Source: OECD (2018a). Estimates are OECD calculations based on different datasets using the two-sample two-stage least squares estimator for Austria, Belgium, Ireland, Luxembourg, the Netherlands, Portugal, Spain, Greece, Italy, the United Kingdom, Hungary and Chile. For other countries, estimates come from various studies.

tive data was undertaken by Dearden et al. (1997). Using the National Child Development Survey (NCDS), a longitudinal dataset that follows a cohort of individuals born in Britain in a particular week in March 1958, their paper examines intergenerational mobility in terms of labour earnings and years of schooling, measured at the age of 33. Dearden and colleagues find a limited degree of earnings mobility, with IGEs in the wide range of 0.24-0.59 for sons and 0.35-0.70 for daughters.

Following this study, Blanden et al. (2004) estimate the earnings elasticities also for the NCDS and compare it with the cohort from the British Cohort Survey (BCS), born in 1970, with the aim of understanding how intergenerational mobility has changed over time. They observe that for the younger cohort (BCS) intergenerational mobility is generally lower. For sons, the estimated IGEs are around 0.17 for NCDS and 0.26 for BCS. For daughters, they obtain an IGE of 0.17 for NCDS and 0.23 for the BCS. They hypothesise that part of the fall in earnings mobility between the 1958 and 1970 cohorts could be related to the unequal increase in educational attainment over this period,

using an original data set created from a household survey in York (Atkinson, 1980).

which mostly benefited children from richer parents. Using the same data, Blanden et al. (2013) estimate the IGE of income for sons obtain 0.21 for the NCDS and 0.28 for the BCS cohort.

The findings from Blanden et al. (2004, 2013) have been particularly contested (Jäntti and Jenkins, 2015), partly due to controversies around the comparability of the NCDS and BCS data sets. One reason is the use of different measures of permanent economic status in the two studies, with separate father's and mother's earnings being used in the NCDS and combined parental income in the BCS.[13] In spite of these criticisms, these results have been used widely in the UK public policy debates about social mobility - though not always with the appropriate interpretation. As Goldthorpe (2013) argues, a 'consensus view' that social mobility has been in decline over recent decades has emerged, solely based on these two studies, and is open to question.

Most early studies focused on measuring earnings mobility and the shift in the literature to study mobility of family income is only recent. The emphasis on capturing all available childhood resources makes sense, if the goal is to capture the relationship between living standards of parents and children. The different estimates obtained by Blanden et al. (2004) and Blanden et al. (2013) using the NCDS and BCS and the comparability issues involving the observed outcomes also illustrate the importance of observing the income measure used. In the international literature, this has also been discussed by Landersø and Heckman (2017). For the UK, only a handful of studies measure intergenerational income mobility with focus on household income, and this paper contributes to this small literature. I also examine the robustness of my results to changes in the children's income variables used.

A recent UK study by Belfield et al. (2017) emphasises the importance of understanding what exactly is the nature of association being measured and that IGE estimates

[13] In the NCDS, there is only a single measure of father's and mother's earnings, when children were 16. In addition, these earnings are measured by net weekly earnings (wages) and only reported in bands, with no exact value being observed. Children's earnings were observed at one point, when they are aged 33 years. In the BCS, there is only parental earnings combined, measured when children are aged 10 and 16, and children's earnings is observed at age 30.

might be sensitive to how child income is observed in the data. They show how their estimates change when using different definitions of son's income and earnings: sons' individual gross earnings, gross private income[14] and net family income, while holding constant parental income[15] as net family income. For the NCDS cohort, authors obtain an IGE of 0.22 using gross earnings, 0.20 using gross private income and 0.17 using net family income. For the BCS cohort, they estimate an IGE of 0.36 for gross earnings, 0.37 for gross private income and 0.31 for net family income.

Moving away from the cohort data, other recent studies have alternatively used the BHPS data. Ermisch and Francesconi (2004) use a father-child matched sample and are able to measure intergenerational mobility for individuals from different cohorts[16] and backgrounds. This was the first intergenerational mobility study with the BHPS data, using the first 8 waves available, from 1991-1999. Using a similar approach to my own to construct the sample, they create a matched sample of fathers and children who were interviewed in the survey. Employing the traditional OLS method, their estimates of IGE for monthly earnings and annual income for sons are around 0.05. Their results suffer from the short period of data available and young age at which children's earnings is observed, at 16, which is extremely early in terms of working life and likely to attenuate considerably the IGE estimates due to life-cycle effects.

In order to reduce this downward bias, they do an IV estimation with four different sets of instruments for parental income[17]: parental education, HG-index[18], family structure and local unemployment rate. The IV regressions yield IGEs of around 0.10 for the first two instruments and 0.20 for the third and fourth instruments. A similar approach is

[14] For the authors, this is a similar measure to gross household income

[15] In this model, they use a one-point parental income observation when children were 16 and sons' income or earnings is captured when they were 42 years old.

[16] They focus on pairs of fathers and sons born between 1970 and 1983.

[17] Some of these instruments, such as parental education and occupation are very likely to be correlated with sons' earnings, what probably creates an upward bias on these estimates, providing an upper bound of IGE.

[18] The Hope-Goldthorpe index is based on a ranking of occupations obtained from a random sample of individuals interviewed in England and Wales in 1972. Ermisch and Francesconi (2004) argue that the HG-index is likely to be a good measure of permanent socio-economic status, as it is highly correlated with earnings and thought to be relatively stable over the working life.

also used by Nicoletti and Ermisch (2008), who extend this earnings mobility analysis to sons born between 1950-1972. Since there is no data with information on both fathers' and sons' earnings for these cohorts, they attempt to overcome this by combining two samples from the BHPS and estimating the elasticities by two-sample two-stage least squares (TS2SLS).[19] The reported IGEs by cohort for single year earnings are estimated to be in the range 0.20-0.30.

These studies highlight important issues. Due to the nature of the instruments available for parental income, it is likely that they are positively correlated with children's earnings indirectly, through fathers' earnings, but also directly, which would mean that the IV and TS2SLS estimates are positively biased and would suggest an upper bound of IGEs. In addition, Ermisch and Francesconi (2004) looked at children's incomes at a very young age (16), which would suggest that their IGEs are largely downward biased due to life-cycle effects.

As discussed previously in the background literature, another common issue around the estimation of intergenerational elasticities is possible attenuation bias related to the use of short-run income variables as a proxy to permanent income. Another recent paper, by Gregg et al. (2017) presents revised IGE estimates for the NCDS and BCS cohorts accounting for both life-cycle and attenuation bias, by using the two measures of parental income available in the BCS study (at 10 and 16) and estimating intergenerational mobility at various points along the life-cycle. They show that the use of two points of parental income rather than one reduces the attenuation bias for the BCS estimates.[20] They also show that for both cohorts the IGEs start very low during the early twenties and rise constantly until the mid-forties.[21]

[19] The TS2SLS approach is often used when there is no information on parental earnings/income in the data set (Jäntti and Jenkins, 2015; OECD, 2018a). First used in the context of intergenerational mobility by Björklund and Jäntti (1997), the method consists of using a second sample to predict earnings/income for the parental generation. Thus, it is based on a sample of children including information on their income distribution and key predictors of parental income and another sample, of parents, which contains information on the unconditional distribution of income in that generation.

[20] Due to lack of multiple observations of parental income in the NCDS they could not present comparable estimates for this cohort.

[21] For the NCDS cohort, IGEs go from 0.042 at age 23 to 0.259 at age 46. For the BCS, they go from 0.203 at age 26 to 0.397 at age 42.

Gregg et al. (2017) is, to my knowledge, the first to calculate rank coefficients for the UK as a complement to IGE measures. They observe that the rank-coefficients follow a similar pattern to the IGEs across the life-cycle. However, their results suggest that rank coefficients seem less attenuated than IGEs at lower ages and less affected by attenuation bias from measurement error or transitory shocks in the parental income variable. Based on this, the authors suggest that rank estimates might be more adequate when income is observed at early ages. Overall, their best revised IGE estimates[22] for the UK are 0.43 for the BCS cohort and 0.25 for the NCDS.[23]

In this paper, I have access to 26 waves of data with the BHPS + UKHLS, which allow me to look at children's outcomes when they are older, from the age of 25. In addition, I employ a two-stage residual approach alongside the usual OLS to calculate intergenerational elasticities, which allows me to better control for the ages at which income is observed. This large data set also contains multiple observations of income for parents and children, which allows us to construct a better proxy of permanent income less subject to transitory shocks, and reduce the magnitude of the attenuation bias from measurement error. Finally, following the recent development of this literature, I am also able to estimate rank coefficients.

It becomes clearer after this more detailed examination that there is a great deal of uncertainty around the degree of intergenerational mobility in the UK. This literature has certainly developed in the recent years, following new methods and estimation strategies, but many areas remain unexploited. As Jäntti and Jenkins point out, "a key conclusion that we draw about the UK debate [...] is that much richer data than those provided by the NCDS and BCS is needed to draw firm conclusions about the level trend in UK income mobility." (Jäntti and Jenkins, 2015, p. 911).

[22]The correspondent rank coefficients are 0.195 for the NCDS and 0.298 for the BCS. These rank coefficients suffer less from attenuation and workless spells bias and are very stable to the adjustments made to IGEs.

[23]Although the NCDS IGE probably still suffers from attenuation bias as it is only based on one observation of parental income.

4 Data and Methodology

4.1 Data

As discussed before, the study of intergenerational mobility involves very strict data requirements. In this paper, I use a combined data set of two household longitudinal surveys: the British Household Panel Survey (BHPS), from 1991-2008, and Understanding Society, from 2009-2016.

The BHPS is a panel survey of households in the UK. Having started in 1991, the BHPS interviewed a nationally representative sample of 5,500 households and 10,300 individual respondents. Every year, the same individuals have been re-interviewed. Even though there is some attrition, the sample is replenished with the addition of new households and individuals that join participant households. If individuals leave the original household to form new households, all adult members (aged 16+) in this household are also added to the survey. In addition, children of the original households are also interviewed using the adult questionnaire once they are aged 16. In total, the BHPS is comprised of 18 waves of data, spanning from 1991 to 2008.

The Understanding Society, also referred to as UKHLS for UK Household Longitudinal Study, started in 2008 and provides a larger and more comprehensive continuation of the BHPS. After its second wave, the Understanding Society main study includes information collected for continuing participants of the BHPS. Of around 8,000 BHPS participants[24] invited to join, almost 6,700 accepted the offer and are being interviewed in Understanding Society every year since 2009. The two studies have been recently adapted and harmonised, and as such they provide an opportunity to increase the main BHPS sample as well as to extend the analysis to more recent years (2009-2016).

This harmonised data is particularly suited to the estimation of intergenerational mobility. Firstly, differently to the NCDS and BCS data, children and parents come from multiple heterogeneous cohorts. Secondly, the longitudinal nature of this dataset makes

[24]Participants on the last wave (18) of the BHPS.

it possible to link children to their parents and obtain the relevant variables directly from these individuals, as opposed to relying on variables originated from retrospective questions and could be affected by recall error. Another advantage of this data is that the information on income for parents and children comes from multiple years and is not restricted to a single observation of income. Considering all 26 available waves of data, the harmonised data set is sufficiently long to calculate an approximation of parental permanent income during childhood for a sub-sample of young individuals who can be linked to at least one parent within the survey. Lastly, the harmonised data allows us to observe directly the income of sons and daughters as they become adults.

The use of the harmonised BHPS and Understanding Society allows me to complement and expand the analysis of the intergenerational mobility of income in the UK started by researchers who used the British cohort studies. To my knowledge, this is the first study using the harmonised data for this purpose. Using this rich data, I can test the sensitivity of the main results to a series of changes in the model specifications, sample restrictions and income measures, and it represents an addition of at least 10 years of data since the last related studies with the BHPS, by Ermisch and Francesconi (2004) and Nicoletti and Ermisch (2008).

4.2 Main Variables

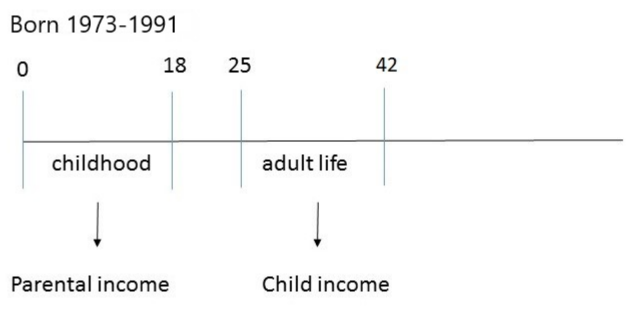
Following the approach by Lee and Solon (2009) and Chetty et al. (2014), the main variable of interest throughout this paper is gross (pre-tax) household income in the month before the interview. The use of family or household income has been discussed in the literature and justified by this variable being a good indicator of the parental living standards, or socio-economic situation during childhood (Lee and Solon, 2009; Chadwick and Solon, 2002; Chetty et al., 2014). Hence, the use of household income as opposed to parental individual income or earnings is supported by the attempt to measure the influence of all resources available during childhood to the outcomes of young adults.

Parental income is gross household income and is captured for parents during their children's childhood years (when the children are aged 0-18). When the child lives with just one biological parent, parental household income equals that of the single parent only.[25] A restriction is imposed so that to be in the matched sample, children must have lived in the same household with at least one of their biological parents, for at least one of the childhood years. This restriction means that children must be no older than 18 years old in 1991, that is, they have to be born on/after 1973.

Child income is the (gross) household income of the child when they become adults, measured at age 25 or above. The choice of 25 as the cut-off minimum age relates to this being the age at which most young individuals will have already left education and entered the labour market. This classification is also used by the OECD labour markets statistics (OECD, 2018b), which considers individuals to be 'young' until the age of 24, and adults thereafter. In order to have at least one observation for income when children are 25+, a restriction is imposed on the sample: only individuals born on/before 1991 are included.

Figure 3 illustrates how the main variables are collected over the life time of individuals (the children) in my sample.

Figure 3: Collection of variables: timeline



Note: This Figure shows how the main variables for this analysis are collected over the life-cycle. Source: Own elaboration.

The income variables are measured in the same way in the BHPS and Understanding Society, and collected every year from individual respondents using the same question.

[25]The coresidency status during childhood is obtained using the household situation after the child individual respondent first entered the survey as a respondent, at the age of 16. I assume that this cohabitation status at 16 is the same as it was in the previous childhood years. Because I am looking at the income of the household, this does not affect the collection of the parental income variable.

All income variables are in GBP and deflated using the CPI of every interview year, with 2015 as the base year. More specific information about these income variables can be found in the Appendix A.

Finally, the age variables used as controls in some of the estimation models are always captured as the age at which the relevant income variable is observed. When a child lives with both parents, parental age is the father's age, as fathers are heads of 96% of these households. When the child lives with only one parent, the age of this parent is considered.

4.3 The Matched Sample

Overall, individuals need to meet all the following conditions (C) in order to be retained in the final matched sample:

- C1) need to be children born between 1973 and 1991;
- C2) can be matched to at least one of their biological parents within the BHPS;
- C3) there is at least one income observation for parents available during the childhood years, and when parents are aged 25-60;
- C4) there is at least one income observation after they are aged 25+.

The final sample of individuals after imposing C1-C4 is comprised of 2102 children, corresponding to 78% of the age-eligible individuals who were interviewed at 25+. This sample will be used throughout the remainder of this paper, unless stated otherwise. The issue of sample selection does not seem to be cause for major concern, and more details on this sample and its representativeness are provided in Appendix B.

4.4 Descriptive Statistics

Table 1 presents the summary statistics for the main variables used in this paper. Panel A refers to parental characteristics and Panel B refers to the characteristics of their sons and daughters in the matched sample.

Table 1: Descriptive Statistics

Panel A: Parents (N=2102)	Mean	SD	Min	Max
Parental household income (£/month)	3623.09	(2761.81)	0	102426.2
Parental age	41.2	(7.62)	17	75
Parental age when income observed	42.1	(6.63)	25	60
Child age when parental income observed	13.1	(4.16)	0	18
Panel B: Children (N=2102)	Mean	SD	Min	Max
Adult household income (25+) (£/month)	4142.54	(2388.72)	0	24941.51
Adult personal income (25+) (£/month)	1953.44	(1376.44)	0	18337.41
Adult labour income (25+) (£/month)	1724.05	(1431.17)	0	18337.41
Age when household income observed	29.9	(4.31)	25	43

Notes: Summary statistics based on the main matched sample of 2102 individuals for whom all of the variables are observed. Each pair of parent-child has multiple income observations. In total, there are 15220 observations of household income for parents and 12773 for children. Adult personal income and labour income are available for N=2090 pairs and for these variables there are 12771 and 12704 observations, respectively.

This data reveals that, on average, children have higher real household income than their parents, which could suggest the improvement of living standards across generations in absolute terms. According to recent reports, upward mobility in absolute terms has been occurring in most OECD and emerging countries (OECD, 2018a). The Table also shows that the ages at which parental and child income are observed - parents are on average older when their income is observed. Finally, average personal income[26] for the children is, by definition, higher than (or at least equal to) labour income because it includes also income from non-labour sources.

5 Estimates of Intergenerational Income Mobility in the UK: Estimation Strategy and Main Results

Initially, I employ the traditional Ordinary Least Squares (OLS) approach to estimate the intergenerational income elasticity and to offer comparability with previous studies. Then, an alternative two-step residual approach (TSRA) is proposed to estimate IGEs, which better accounts for the presence of life-cycle effects. Thirdly, I estimate the per-

[26]More information about the income variables is included in the Appendix B.

centile rank coefficients. These three estimation methods and main results obtained are presented and discussed in this section.

5.0.1 Ordinary Least Squares (OLS) Approach

Based on the traditional approach used in the literature (Solon, 2002; Black and Devereux, 2011; Jäntti and Jenkins, 2015), I estimate the intergenerational income elasticity using a version of the general model presented in Section 2, described in the following equation:

$$log\ (\overline{Y_i^{child}}) = \alpha + \beta\ log(\overline{Y_i^{parent}}) + X_i'\theta + \epsilon_i$$

Where $\overline{Y_i^{child}}$ is the multi-year average[27] of income for individual i as an adult (after the age of 25), and $\overline{Y_i^{parent}}$ is the multi-year average of parental income observations during the childhood of individual i. The estimated β is the intergenerational income elasticity. X is a vector of control variables at the individual level, such as the average age of children and of the main parent[28] when income is observed, as well as birth year dummies of children and parents, which capture cohort effects.

The estimates from this model give us a comparative benchmark to the findings of previous studies that use a similar methodology. As discussed previously, there are two common issues related to the estimation of IGEs in this way. Firstly, the errors-in-variables bias from calculating permanent income because of transitory shocks that might affect the observed income. Secondly, the life-cycle bias from measuring incomes of children and parents at different points of the life-cycle.

To deal with the attenuation bias, parental income is taken as the multi-year average of observed household incomes during childhood, that is, average household income when children are 0-18. On average, I observe parental income during 7.2 years. Hence, permanent parental income is proxied by an average of multiple observations of income rather than being a single point in time measure. This is one way to reduce the attenuation bias

[27] Using the log of multi-year averages, not multi-year averages of log incomes.

[28] When the individual lives with two biological parents, the age of the father is considered. This is because for 96% of the households with two parents, the father was the head of household

from transitory shocks to permanent income and one of the advantages of using this data as compared to the NCDS and BCS cohort studies.[29] Simulation results by Mazumder (2005)[30] suggest that using 7 years of observations would not be enough to eliminate completely this source of attenuation bias, but could reduce it considerably.

To deal with the life-cycle bias, I impose two selection criteria; one that drops all children without incomes observed after/on the age of 25 and one that drops parental income observations for parents who are younger than 25 and higher than 60 at the time their income is observed. In addition to this, I control flexibly for ages at which incomes are observed using TSRA.

Using a multi-year average of child income is also useful to reduce the year-to-year variability of income and the influence of episodes of very low income (Nybom and Stuhler, 2016) and limit the life-cycle bias. Arguably, the cut-off age of 25 could still be considered young in terms of obtaining a good proxy for lifetime income and could be a cause of underestimation of IGEs in this model. In the second part of this paper, I test whether these results are robust to changing the age restrictions in the sample.

Table 2 presents the estimates of intergenerational income elasticity obtained by using the traditional OLS model and regressing the log of averaged child income on the log of the averaged parental income (plus controls, when mentioned).

Column 1 shows the results of the estimation with no additional controls. In the second column, I also control for the average age of parents and children at the time that income is observed.[31] In column 3, I add a set of dummies to control for the birth year of parents and children, in order to account for the fact that, for each generation, individuals come from multiple cohorts. These results suggest that the intergenerational

[29]Previous studies that make use of the NCDS, for example, rely on a single measure of parental income during childhood. Studies that use the BCS have only two observations of parental income during childhood available.

[30]Mazumder (2005) estimates that one would need 20 to 25 years of income data to calculate a multi year average that would serve as a good proxy for the permanent component of earnings and eliminate completely the attenuation bias from this source.

[31]The results were identical adding average age and average age squared for parents and children to models (2) and (3).

Table 2: OLS Estimates of intergenerational income elasticity in the UK

Log Child Income	(1)	(2)	(3)
Log Parental Income	0.28***	0.27***	0.27***
	(0.027)	(0.027)	(0.028)
Average Child Age	-	0.016***	0.015***
		(0.0042)	(0.0058)
Average Parental Age	-	0.0042*	-0.011
		(0.0022)	(0.0076)
Child birth year	No	No	Yes
Parent birth year	No	No	Yes
Constant	5.91***	5.35***	6.22***
	(0.21)	(0.23)	(0.55)
Observations	2102	2102	2102

Clustered standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Notes: IGE of income estimated by OLS. In (1) the model is estimated without additional controls. In (2) I control for the average age of parents and children. In (3) I add birth year dummies for parents and children.

elasticity of income for the UK is around 0.27 and precisely estimated.[32] This means that an additional 10% in parental income during childhood would give their children a 2.7% income advantage as adults.

5.0.2 Two Stage Residual Approach (TSRA)

Complementing the OLS approach, I also propose the use of a two-stage residual approach to estimate IGEs.

Due to the panel structure of this data, for each pair of parent-child in the sample, I observe income multiple times. Using the OLS method, the multiple observations of incomes of parents and children are averaged, and the same is done with the ages at which income is observed. However, this way aggregating the necessary information on income and age means that specific information on the age-income profiles of individuals gets lost in the process.

The relationship between the current observed income and 'permanent' income changes over the life-cycle (Haider and Solon, 2006). In addition, there is evidence of heterogene-

[32] All standard errors reported are robust and clustered on parents. The idea behind this cluster analysis is that if parents have more than one child, the error term for these children would be cross-sectionally dependent.

ity in the individuals' age-income profile (Nybom and Stuhler, 2016), which reflects the development of income at different ages. Because of the existence of such age-income profiles, the accuracy and the meaningfulness of the averaged current income to represent lifetime income will depend on the ages at which current income is observed. For example, if current income is observed and averaged during the period of lower levels and fast growth, such as in early-career years, it is probable that the obtained average is understating lifetime income. In my data set, adult income is first observed when individuals are 25 years old, and so this might be a relevant issue.

I propose an alternative approach that considers a more flexible way of collapsing the multiple observations of current income, by which I make use of all the information available in the data on income and age at which it is observed. This alternative method is the two-stage residual approach (TSRA). It consists of estimating the intergenerational elasticity of income in two steps. In the first stage, I use auxiliary regressions of parental income and child income and estimate a measure of income adjusted by age and time effects, which can serve as a better proxy of permanent income. In the second stage, I use the residualised (adjusted) income to calculate the IGEs.

A similar approach has been employed by studies that examine *intragenerational* mobility trends, more specifically, research on the volatility and instability of individual (especially men's) earnings. They tend to first run regressions of earnings controlling for differences in age, education and work experience and then work with these earnings residuals (Jäntti and Jenkins, 2015). For example, Shin and Solon (2011) use a residualised measure of the change in log earnings by regressing, in the first stage, log earnings on a quadratic in age. Moffitt and Gottschalk (2012) use a similar method by regressing log earnings on education, age polynomials and interactions between them. In the second stage, they compare the earnings residuals for the same individuals at different points in time. Nonetheless, to my knowledge, this is the first time that this two-stage approach has been used to estimate *intergenerational* income mobility in the UK.

The first stage auxiliary regressions for parents and children take the form:

$$log\ Y_{it}^{parent} = f(age_{it}, time) + w_{it}$$

$$log\ Y_{it}^{child} = f(age_{it}, time) + v_{it}$$

In these auxiliary regressions, I control for age and its square and year dummies.An example of first-stage auxiliary regressions is shown in Appendix C.

In the second stage, the residuals obtained in the auxiliary regressions i.e. the age and time adjusted log income are averaged for each individual. Then, they are used in the second stage main regression. Parents' averaged residuals will be called adjusted parental income and, similarly, children's averaged residuals will be called adjusted child income. The main idea behind this model is to examine the intergenerational association between the age- and time-adjusted incomes of parents and children. This should help reduce the bias from observing incomes of parents and children at different stages of life.

The second stage main equation can be written as:

$$\overline{\hat{Y}}_{res_i}^{child} = \alpha + \beta\ \overline{\hat{Y}}_{res_i}^{parent} + e_i$$

$$(\text{where } \overline{\hat{Y}}_{res_i}^{child} = \overline{\hat{v}_i} \text{ and } \overline{\hat{Y}}_{res_i}^{parent} = \overline{\hat{w}_i})$$

Table 3 presents the main results from the second stage for the intergenerational income elasticity using TSRA.

It is also noteworthy that when employing the TSRA method, the main variables in the second stage are the residuals (age- and time-adjusted income), which are generated regressors that come from the first step auxiliary regressions. Because of this, instead of reporting the usual clustered (by parents) standard errors from Stata, which are likely to be underestimating the amount of variation that exists in the data, I report the standard errors that have been adjusted with bootstrapping techniques in the two stages.

Combining the two sets of results obtained by OLS and TSRA, the intergenerational

Table 3: TSRA Estimates of intergenerational income elasticity in the UK

Adjusted Child Income	(1)	(2)	(3)
Adjusted Parental Income	0.24***	0.25***	0.25***
	(0.026)	(0.026)	(0.026)
Bootstrapped SEs	(0.020)	(0.020)	(0.020)
Age	Yes	Yes	Yes
Age squared	No	Yes	Yes
Year dummies	No	No	Yes
Constant	0.023*	0.023*	0.027**
	(0.013)	(0.013)	(0.013)
Observations	2102	2102	2102

Clustered standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Notes: In (1) the only control is age in the auxiliary regressions for children and parents. In (2), I control for age and age squared and in (3) for age, age squared and dummies for every year. The bootstrapped standard errors are based on 1999 replications.

income elasticity for the UK is estimated to be between 0.25-0.27. The TSRA method allows for a better control for age and cohort effects because this is done based on the point at which income is observed and using the full information available in the data and not only considering the average age and birth year effects like in OLS.

5.0.3 Rank Coefficient

The standard measure of intergenerational mobility, the intergenerational elasticity (IGE) combines both marginal and joint distributions of parents' and children's incomes, capturing the extent of re-ranking across generations and the spread of the distributions.[33] When estimating the rank coefficients, we focus solely on the re-ranking across generations, as the spread of the distribution is standardised.

The rank-rank slope (γ), or rank coefficient, is obtained from

$$Rank_i^{child} = \alpha + \gamma \ Rank_i^{parent} + X_i'\theta + u_i$$

where the percentile rank for each individual i is obtained from the averaged child income over multiple periods. The same is done for parents, using the averaged parental

[33]The relationship between IGEs and rank-based measures is further discussed by Mazumder (2016).

income in childhood years. Similarly to the OLS model, X is a vector that contains individual controls for the average age and birth year dummies of parents and children The ranks are obtained by ranking parents with respect to the income distribution of parents in the sample and children with respect to the income distribution of children.

Table 4 presents the rank coefficients of income obtained by ranking children and parents according to their position in the income distribution of each generation and then regressing the percentile rank of the children on the percentile rank of the parents (plus controls, when mentioned).[34]

Table 4: Estimates of income rank coefficient in the UK

Rank Child Income	(1)	(2)	(3)
Rank Parental Income	0.31***	0.31***	0.30***
	(0.023)	(0.023)	(0.024)
Average Child Age	No	0.0082***	0.0081***
		(0.0024)	(0.0031)
Average Parental Age	No	0.0014	-0.0059
		(0.0012)	(0.0041)
Child birth year	No	No	Yes
Parent birth year	No	No	Yes
Constant	0.34***	0.058	0.39
	(0.013)	(0.081)	(0.27)
Observations	2102	2102	2102

Clustered standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Notes: Income rank coefficient estimated by OLS. Individuals were ranked separately for each generation among the sample of 2102 pairs. In (1), the model is estimated without any additional controls. In (2), I control for the average age of parents and children. In (3), I add birth year dummies for parents and children.

Overall, Table 4 shows a rank coefficient of 0.30 when controlling for average age of children and parents and birth year dummies.[35] This means that, on average, a 10 percentile point increase in the rank of the parents would translate in a 3 percentile point increase in the rank of income for children. That is, we still see a positive intergenerational

[34]For Table 4, I use the percentile ranks created based on the final sample of 2102 matched pairs of parents and children. These results are identical to those obtained using a larger sample of individuals with income information but who could not be matched to an eligible individual from the other generation. In the larger sample, parents were ranked among the 4630 age-eligible parents of individuals born between 1973 and 1991, while their children were ranked among all 2692 age-eligible individuals with some income information available at age 25 or older. For consistency, all results presented in this paper will be based on the ranking of parents and children among the 2102 individuals in the final sample.

[35]The results were identical adding average age and average age squared for parents and children to models (2) and (3).

persistence of income, and of similar magnitude to the estimated IGEs.

5.1 Robustness Analysis

The results so far suggest that the estimated IGEs in the UK are in the range 0.25-0.27, and significantly different from zero. The rank coefficient is slightly larger, estimated at 0.30. In this section, I examine the robustness of the estimated IGEs and rank coefficients along a number of dimensions. First, I investigate the sensitivity of the results to changes in the definition of the main income variable. Second, I investigate whether the restrictions imposed for children regarding the age at which income is observed affect my estimations. Third, I look at differences in the coresidency status of children and how that might affect the estimated coefficients. Finally, I treat and exclude outliers to check if they influence my results.

5.1.1 Using alternative definitions of income

In many cases, the choice of the income measures in intergenerational studies are restricted by data availability. This is further aggravated by the data-intensive nature of intergenerational analyses, which by definition require income information across generations. In theory, however, this choice should reflect the purpose of the study and the research question being investigated (Björklund and Jäntti, 2011). Other empirical studies have shown that estimates of mobility might be affected by the choice of income measure (Landersø and Heckman, 2017; Belfield et al., 2017).

The household income variable reflects the resources of the household in a broader sense. As a measure of parental income, it captures the resources of the parents when their children were being raised. When used as a measure of child income, it represents the current resources of their new households highlighting the importance of the partner's income to the well-being of the whole household. In this section, I check the robustness of my main estimates to using alternative measures of child income, keeping parental income constant. These results are presented in Table 5.

The first row summarises the main results obtained in the first part of this paper using the household income variable for parents and children. In the second row, the dependent variable is total labour income. This variable represents the usual pre-tax monthly labour earnings from the main job, self-employed profit and second and occasional jobs, as appropriate. Several studies[36] use labour earnings to analyse earnings mobility, as information on wages is commonly available in most data sets. This concept captures people's earnings power in the labour market. The main criticism of using this type of data relates to the fact that it excludes all individuals who were not working at the time of interview[37], which can also complicate the inclusion of women in the analysis (Chadwick and Solon, 2002; Raaum et al., 2008). In addition, in some data sets self-employment earnings are missing or measured less accurately (Björklund and Jäntti, 2011).

Another possible income measure is total personal income. This can be also called total factor income, and is a broader measure of income-generating power (Björklund and Jäntti, 2011), including inherited capital income and other non-labour income sources (benefits, pensions, transfers), as well as labour income. The results obtained using this measure of income are shown in the third row of Table 5 and are very similar to those obtained with the household income measure.

The fourth row contains the results for pooling personal income for spouses, reflecting the income generating power of the couple. In this case, it is important to keep in mind how the intergenerational associations of family resources might be affected by assortative mating.[38] This is an interesting topic in itself and deserves being investigated further.

Overall, the IGEs seem to be very stable to changes in the income variable, except for the estimates using labour income, which might be affected by individuals being unemployed at the time of the survey. Surprisingly, the rank coefficients appear to be a bit less stable to the use of different income measures.

[36] For the UK, Dearden et al. (1997), Blanden et al. (2004), Blanden et al. (2013), Ermisch and Francesconi (2004) and Nicoletti and Ermisch (2008) are examples of studies looking at earnings mobility.

[37] Gregg et al. (2017) studies the bias in IGEs from not considering spells out of work.

[38] Assortative mating is defined as "any nonrandomness in the process of who mates with whom" (Chadwick and Solon, 2002, p.336). In the context of marriage, assortative mating can also be called marital sorting.

Table 5: IGE and rank coefficient based on different measures of income

Child income	Parental income	N	IGE (OLS)	IGE (TSRA)	Rank
Household income	Household income	2102	0.27***	0.25***	0.30***
			(0.028)	(0.026)	(0.024)
Total labour income	Household income	1907	0.32***	0.25***	0.27***
			(0.040)	(0.030)	(0.024)
Total personal income	Household income	2077	0.27***	0.25***	0.25***
			(0.031)	(0.030)	(0.023)
Total personal income + spouse	Household income	2079	0.27***	0.25***	0.22***
			(0.032)	(0.030)	(0.022)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: In the OLS and rank models, controls for the average age of parents and children and birth year dummies are used. In the TSRA model, I control for age and age squared of parents and children, as well as year dummies. For the rank model, due to the use of income in levels (not log), the sample size is 2090 for the measures of individual income.

5.1.2 Relaxing age restrictions in the sample

As discussed before, it is a consensus in this literature that the age at which income is observed (for parents and children) is important because of the life-cycle bias (Haider and Solon, 2006; Grawe, 2006). For the UK, Gregg et al. (2017) show how IGEs and rank coefficients vary for the NCDS and BCS cohorts based on when sons' income is observed. They are underestimated when sons are in their early twenties and peak when sons are in their mid-forties. These results are similar to what has been observed for other countries (Grawe, 2006; Nybom and Stuhler, 2016; Haider and Solon, 2006)

The choice of when to observe income is often dictated by data availability. However, the evidence suggests that if income is observed when individuals are too young, they would have not reached their earnings potential yet, leading to an attenuation of the estimated coefficients and overestimation of mobility. This is seen, for example, in the results obtained by Ermisch and Francesconi (2004), which measure earnings mobility for children at the age of 16 and obtain IGEs close to zero.

In this paper, the main models already control flexibly for the ages of children and parents when income is observed and for their birth years, which should be sufficient to control for incomes being observed at different ages and for children and parents from different cohorts. However, due to the heterogeneous income profile across individuals and to possible differences between the two generations, it is probable that income observed in mid-thirties or forties would still be a better representation of lifetime income. Here, I investigate whether IGEs and rank coefficients are affected by varying the sample restrictions on the age at which child income is observed. These results by age groups are presented in Tables 6 and 7.

Table 6 contains the results for the household income variable. The IGEs and rank coefficient based on the household income variable decrease gradually as the minimum age increases from 20 to 25, but become more stable afterwards. This is related to the fact that when looking at household income as a measure of socio-economic status one extra confounding factor manifests itself. Young individuals probably still mainly live with their parents and, for the group of these so called ‘coresidents’, the estimated elasticity would be more a measure of persistence of income within the same household than of intergenerational mobility. The coresidency issue will be further analysed in the next section, but for now I look into how these estimates for different ages change if I use the personal income variable.

Table 6: IGE and rank coefficient when household income is measured at different ages (children born 1973-1991)

Child income	Parental income	N	Age	IGE (OLS)	IGE (TSRA)	Rank
Household income	Household income	3328	20-43	0.40***	0.35***	0.43***
				(0.021)	(0.022)	(0.023)
		2819	22-43	0.33***	0.30***	0.38***
				(0.025)	(0.024)	(0.024)
		2332	24-43	0.29***	0.26***	0.32***
				(0.027)	(0.025)	(0.024)
		2102	25-43	0.27***	0.25***	0.30***
				(0.028)	(0.026)	(0.024)
		1852	26-43	0.29***	0.26***	0.31***
				(0.032)	(0.030)	(0.025)
		1455	28-43	0.26***	0.24***	0.28***
				(0.033)	(0.033)	(0.028)
		1066	30-43	0.26***	0.25***	0.30***
				(0.034)	(0.036)	(0.032)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Notes: In the OLS and rank models, controls for the average age of parents and children and birth year dummies are used. In the TSRA model, I control for age and age squared of parents and children, as well as year dummies. For the rank model, due to the use of income in levels (not log), the sample size (N) increases by 2.

Table 7 presents the results for the personal income variable. These results are con-

Table 7: IGE and rank coefficient when personal income is measured at different ages (children born 1973-1991)

Child income	Parental income	N	Age	IGE (OLS)	IGE (TSRA)	Rank
Personal income	Household income	3230	20-43	0.19***	0.11***	0.21***
				(0.026)	(0.024)	(0.022)
		2762	22-43	0.20***	0.18***	0.23***
				(0.027)	(0.025)	(0.022)
		2297	24-43	0.25***	0.23***	0.25***
				(0.030)	(0.028)	(0.023)
		2077	25-43	0.27***	0.25***	0.25***
				(0.031)	(0.030)	(0.023)
		1831	26-43	0.30***	0.27***	0.25***
				(0.035)	(0.033)	(0.024)
		1442	28-43	0.29***	0.26***	0.25***
				(0.042)	(0.040)	(0.028)
		1062	30-43	0.30***	0.29***	0.25***
				(0.047)	(0.046)	(0.033)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Notes: Notes: In the OLS and rank models, controls for the average age of parents and children and birth year dummies are used. In the TSRA model, I control for age and age squared of parents and children, as well as year dummies. For the rank model, due to the use of income in levels (not log), the sample size (N) increases for all groups (by a maximum of 40, at ages 20-43 and less for other ages).

sistent with the hypothesis that the life cycle bias would attenuate the intergenerational elasticity of earnings coefficient. Indeed, when considering income at 20 years old, the estimated coefficients are much lower. Both IGEs and rank coefficients increase with age, and rank coefficients become very stable around the ages of 25 and 26.

Finally, looking only at the cohorts born between 1973-1986[39] (i.e. able to reach the age of 30 in the period studied) the same pattern is observed. The estimated coefficients when child income is observed from their early 20s were large using the household income measure and small using the personal income measure (Table 17 in Appendix D).

5.1.3 Dividing the sample by coresidency status

In order to investigate to which extent the estimated IGEs and rank coefficients are being affected by individuals still living with their parents at young ages, I also separate the

[39] In order to be able to reach the age of 30 in this data set, the individuals need to be born between 1973 and 1986. Restricting the estimation of the main coefficients to these cohorts only is useful in order to separate between the potentially confounding issue of sample attrition and the changes in the estimated coefficients related to the ages at which income is observed.

main sample by family situation. There are two groups, individuals who live with at least one of their biological parents (coresidents) and individuals who do not (non-coresidents) when their income is observed. In practice, this is done by looking at household identifiers.

Tables 8 and 9 present the distribution of the sample by coresidency status at the ages of 20, 25 and 30. As expected, the number of individuals who live with their parents decreases with age. At the age of 20, only around 27% of individuals in the sample live in a separate address but this number almost doubles at the age of 25 (around 57%). Furthermore, in Table 9, looking at the people in my sample who can reach the age of 30 (individuals born between 1973-1986) it is possible to see that this trend continues. For these cohorts, at the age of 20 30% of individuals lived in a separate household, but this number rises to almost 81% at the age of 30.

Table 8: Cohabitation status at ages 20 and 25 (children born 1973-1991)

Status	Frequency (%) at age 20	Frequency (%) at age 25
Lives with father only	3.5	2.6
Lives with mother only	23.9	13.9
Lives with both parents	45	26.5
Lives in a separate household	27.6	57
Total	100	100
Coresidents	72.4	43
Non-coresidents	27.6	57
Total	100	100
N	2048	2045

Note: This Table shows the proportion of the sample with each cohabitation status for children born from 1973-1991 at the ages of 20 and 25. Individuals with this variable missing were not interviewed at these ages.

Table 9: Cohabitation status at ages 20, 25 and 30 (children born 1973-1986)

Status	At age 20	At age 25	At age 30
Coresidents (%)	70	38	19.3
Non-coresidents (%)	30	62	80.7
N	1544	1539	1132

Note: This Table shows the proportion of the sample with each cohabitation status for children born from 1973-1986 at the ages of 20, 25 and 30. Individuals with this variable missing were not interviewed at these ages.

In order to check whether the IGEs and rank coefficients vary according to children's coresidency status when their income is observed, I estimate them again for coresidents and non-coresidents separately.[40] The results are shown in Table 10.

Table 10: IGE and rank coefficient by cohabitation status at ages 25 and 30

	N	IGE(OLS)	IGE (TSRA)	Rank
Status at 25	2045	0.27***	0.25***	0.31***
		(0.029)	(0.026)	(0.024)
coresidents	880	0.41***	0.36***	0.42***
		(0.036)	(0.038)	(0.035)
non-coresidents	1165	0.21***	0.21***	0.25***
		(0.037)	(0.033)	(0.028)
Status at 30†	1043	0.26***	0.24***	0.30***
		(0.034)	(0.036)	(0.032)
coresidents	183	0.48***	0.33***	0.44***
		(0.082)	(0.078)	(0.090)
non-coresidents	860	0.24***	0.23***	0.29***
		(0.038)	(0.039)	(0.035)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Notes: In the OLS and rank models, controls for the average age of parents and children and birth year dummies are used. In the TSRA model, I control for age and age squared of parents and children, as well as year dummies. † The status at the age of 30 is only available for the subsample of individuals born between 1973-1986.

Table 10 shows the IGE and rank coefficient obtained using the household income variable for children. Based on the status at the age of 25, the IGEs for coresidents are much larger than for non-coresidents. This is partly because when using the household income measure the IGE and rank for coresidents will capture not only the intergenerational persistence but also the persistence in parental income over time. For non-coresidents, they only capture the intergenerational persistence. In reality, we could expect the representative point estimates to be between these two values, as the population is comprised of people in both groups. When looking at the results based on coresidency status at the age of 30, the aggregate estimates for both groups together are similar to the results obtained in the first part of this paper - 0.27 (OLS), 0.25 (TSRA) and 0.30 (Rank) - and also similar to what is observed for the group of non-coresidents (that represent 80% of the sample at this age), for which these coefficients represent strictly intergenerational

[40] In the data, it is not possible to tell whether certain individuals live or not with their parents for sure if the parental household ID is missing. At the age of 20 this is true for 3% of the sample and at 25 for around 10%. It increases monotonically with age. These are cases in which the parents leave the sample or die, or they are not traced by Understanding Society. In these cases, they have been included as non-coresidents.

persistence of incomes, and not intrahousehold persistence of income over time. This suggests that the estimates obtained for the full sample are quite robust and not particularly affected by the coresidency issue related to using a household income measure.

5.1.4 Treating outliers in the income data

It has also been emphasised in the literature that IGE estimates might be sensitive to the treatment of extreme and missing values (Nybom and Stuhler, 2016; Dahl and DeLeire, 2008). Even though the income variables are top-coded in the data set, I test the robustness of my results to the treatment of outliers at the bottom and top of the income distribution. This is presented in Table 11.

Table 11: IGE and rank coefficient estimates after the treatment of outliers

Child income variable	N	IGE (OLS)	IGE (TSRA)	Rank
Truncating sample between 5-95%				
Household income	1989	0.26***	0.25***	0.27***
		(0.023)	(0.023)	(0.024)
Personal income	2014	0.27***	0.27***	0.23***
		(0.030)	(0.030)	(0.023)
Winsorising 1-99%				
Household income	2102	0.29***	0.27***	0.30***
		(0.025)	(0.025)	(0.023)
Personal income	2076	0.28***	0.27***	0.25***
		(0.030)	(0.029)	(0.023)
Winsorising 5-95%				
Household income	2102	0.30***	0.28***	0.30***
		(0.024)	(0.024)	(0.023)
Personal income	2090	0.27***	0.27***	0.25***
		(0.030)	(0.031)	(0.023)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: In the OLS and rank models, controls for the average age of parents and children and birth year dummies are used. In the TSRA model, I control for age and age squared of parents and children, as well as year dummies.

Here, the outliers are treated in three different ways. Firstly, I truncate the sample and restrict it to individuals with income between the 5th and 95th percentiles of children's and parental income distributions. This has very little effect on the results using household income apart from a slight decrease in the rank coefficient and increases slightly the IGEs estimated using personal income. Then, I Winsorise[41] the top and bottom incomes

[41]Winsorising is done by limiting the extreme values in the data. In practice, I substitute the lowest

between 1-99% and 5-95%. Winsorising has an effect of increasing the estimated IGEs slightly for both income measures. This would suggest that part of the persistence in IGEs is due to individuals at the very top and very bottom of the distribution and is an issue worth investigating further. The rank coefficients, on the other hand, remain very stable to this treatment.

5.2 Regional Estimates

Socio-economic disparities between regions are very present in the UK. In the conventional wisdom, the various regions are well-known for their unique characteristics, and the persistence of many of these features has been observed in practice in relation to several indicators. In terms of social mobility, a recent report by the Social Mobility Commission reinforces the government's idea that "[...]Britain's social mobility problem is not just one of income or class background. It is increasingly one of geography. A stark social mobility postcode lottery exists today, where the chances of someone from a disadvantaged background getting on in life is closely linked to where they grow up and choose to make a life for themselves."(Social Mobility Commission, 2017, p. 2).

The recent work by the Social Mobility Commission (SMC) has emphasised this regional divide in terms of social mobility indicators related to educational attainment, labour market outcomes and home ownership (Social Mobility Commission, 2016b, 2017). With the goal of examining geographical inequalities in the country, the Commission created the Social Mobility Index in 2016, which reflects differences in opportunities across local authority districts by looking at a range of educational outcomes for children and youth from disadvantaged backgrounds, and the situation of the local job and housing market in these areas. Using this Index, the local authority districts are classified in coldspots and hotspots. These results highlight great regional disparities in opportunities, with areas in London performing very well when compared to the rest of the country,

income values for the value at 1% and 5% and the highest income values for the value at 99% and 95%, respectively.

and the Midlands providing the worst social progress opportunities for individuals from disadvantaged backgrounds.

The work by the SMC, especially the Social Mobility Index, is surely an invaluable tool for economic analysis and points to the existence of regional disparities across the country in terms of social mobility. However, it focuses exclusively on the *intragenerational* aspect does not consider directly the *intergenerational* side of mobility. Empirical regional differences in intergenerational income mobility in the UK remain largely unresearched.

Internationally, a number of studies has emphasised the importance of regional differences in mobility. The influential work by Raj Chetty and colleagues emphasises great regional differences in intergenerational income mobiity across the United States, raising questions about this country being seen as a 'land of opportunities' (Chetty et al., 2014). Using a large administrative data set, their study found considerable heterogeneity across areas, with children who grow up in certain states and cities having much better odds of experiencing upwards mobility than similar children elsewhere.

Following this paper, other studies have examined regional differences in intergenerational mobility for other countries. Bratberg et al. (2017) look briefly at rank coefficients by (big) regions in Germany, Norway, Sweden and the US. Other papers by Heidrich (2017), Corak (2019) and Acciari et al. (2019) use administrative data to estimate income IGEs and rank coefficients across regions in Sweden, Canada and Italy, respectively.

For the UK, the body of evidence on regional differences in intergenerational mobility is very thin. A recent paper by Bell et al. (2018) has investigated the regional variation in intergenerational occupational, educational and housing mobility. Using the decennial census data contained in the Longitudinal Study of England and Wales (LS), and looking at the percentile rank coefficient, their results suggest the highest occupational mobility in London, and the lowest in the North (Yorkshire and Humberside). In terms of educational mobility, this is also lower in the North, especially in Yorkshire, and higher in London. Finally, when looking at housing mobility, however, they observe London with the lowest mobility, while Wales was the region with the highest mobility, which indicates that home ownership mobility is negatively related to house prices.

In this section, I examine whether the role of family background varies depending on where you grow up, by looking at regional estimates of intergenerational mobility using the matched sample from the BHPS + UKHLS data. In the data, I observe the years in which parents have been interviewed (and have their income observed) and assign to each pair of parent-child the region in which the parents lived for the longest time before their child was aged 18. A map of the UK and regions in England is provided in Appendix E.

The results for the North and South[42] of England are presented in Table 12. The intergenerational income elasticities obtained by both OLS and TSRA and the rank coefficients[43] are significantly higher for the North, meaning that there is more intergenerational persistence there, and more opportunities for relative mobility in the South. This is similar to what is found by Bell et al. (2018) for occupational and educational mobility, and reflects what was expected from the differences in income levels and opportunities between these regions. The North and South coefficients are statistically different from each other ($p\ value = 0.000$).

As highlighted by the SMC, there is often a division between London and the rest of the country. Looking at the results from the Social Mobility Index in 2017, London accounts for more than two thirds of the social mobility hotspots in the country (Social Mobility Commission, 2017). Bell et al. (2018) also find the highest occupational and educational mobility in London. Table 12 shows results for the South of England excluding London, and London separately, in order to examine the so called 'London effect'. In terms of intergenerational income mobility, however, it does not seem that the lower estimates obtained for the South are being driven exclusively by London.

Finally, examining intergenerational income mobility at a more disaggregated regional

[42]The division of England in North and South is not official nor straight forward. In this paper, the South is comprised of the East of England, London, South East and South West, and the North of East and West Midlands, Yorkshire and Humber, North West and North East. The map of regions in England is provided in Appendix E.

[43]Following the approach of Chetty et al. (2014) and to facilitate the intuition of comparing estimates across regions, for the rank analysis, parents and children are ranked based on their respective national income distribution (rather than the distribution within each region).

Table 12: IGE and rank coefficients: North and South of England

	England	North	South	South(-)London	London
OLS					
Log Parental Income	0.29***	0.36***	0.19***	0.18***	0.18
	(0.031)	(0.045)	(0.040)	(0.042)	(0.13)
TSRA					
Adjusted Parental Income	0.26***	0.37***	0.15***	0.14***	0.23**
	(0.032)	(0.045)	(0.041)	(0.043)	(0.11)
Rank					
Rank Parental Income	0.29***	0.35***	0.22***	0.22***	0.20*
	(0.031)	(0.044)	(0.041)	(0.044)	(0.12)
Observations	1381	686	695	558	137

Clustered standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Ranks were determined using the national income distribution in each generation.

level, I look at the IGEs and rank coefficients for different regions. Table 13 presents the results obtained using separate regressions for each region. These results suggest the presence of some regional differences, with particularly high elasticities and rank coefficients for the North East and Yorkshire and Humber, indicating low intergenerational mobility, while much lower point estimates are obtained for the East of England and the South West of England. These numbers do provide some additional insight into regional differences, but unfortunately the large standard errors do not allow us to make strong claims about differences between individual regions (or from the national UK average).

Table 13: Regional estimates of IGE and rank coefficients

Region	IGE(OLS)	IGE(TSRA)	Rank	Obs.
North East	0.56***	0.40**	0.67***	66
	(0.20)	(0.17)	(0.23)	
North West	0.33***	0.31***	0.35***	201
	(0.067)	(0.065)	(0.075)	
Yorkshire and Humber	0.42***	0.42***	0.48***	146
	(0.12)	(0.094)	(0.11)	
East Midlands	0.47***	0.44***	0.47***	149
	(0.14)	(0.12)	(0.14)	
West Midlands	0.36***	0.33***	0.31***	124
	(0.11)	(0.12)	(0.12)	
East of England	0.19*	0.082	0.20*	156
	(0.10)	(0.093)	(0.10)	
London	0.18	0.23**	0.20*	137
	(0.13)	(0.11)	(0.12)	
South East	0.20***	0.15***	0.18**	261
	(0.066)	(0.059)	(0.073)	
South West	0.11	0.16**	0.18*	141
	(0.087)	(0.073)	(0.11)	
Wales	0.14	0.20***	0.24***	288
	(0.11)	(0.070)	(0.076)	
Scotland	0.26***	0.24***	0.33***	261
	(0.066)	(0.072)	(0.067)	
NI	0.28***	0.25***	0.26***	170
	(0.093)	(0.085)	(0.091)	

Clustered standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Ranks were determined using the national income distribution in each generation.

6 Conclusions

There is a consensus in the literature that family background can have a significant influence on individual's socio-economic outcomes as adults. This paper provides updated estimates of intergenerational income mobility in the UK. Using the harmonised longitudinal dataset comprised of the BHPS and Understanding Society, I estimate the intergenerational income elasticities and rank coefficients for individuals born between 1973 and 1991.

I estimate intergenerational income elasticities (IGEs) using the traditional OLS approach in the literature, and also propose the use of a two-stage residual approach

(TSRA), which allows for a more flexible control for parents' and children's ages at which income is observed. Controlling adequately for age is very important to minimise the effects of life-cycle bias that might affect these estimates. In addition to adding age controls in the models and using TSRA, I also restrict the sample to include only observations of income for children and parents when they are aged 25 and over. In order to minimise the possible attenuation bias related to the measurement error of permanent income, I use multi-year income observations for both parents and children.

To complement this mobility analysis, I also estimate income rank coefficients, which are an alternative measure of intergenerational mobility that aim to capture the positional mobility of children in relation to their parent's rank in the income distribution.

Overall, my results suggest that the intergenerational elasticity of income in the UK is in the range 0.25-0.27 and statistically significant. This corresponds to the average effect that a small relative difference in parental income will have on their children's income as adults. For example, for every additional (reduced) 10% of parental income advantage (or disadvantage) 2.5-2.7% will be passed on to the next generation. The rank coefficients suggest a similar level of mobility. Using household income measures, I obtain a rank coefficient of 0.30 and using personal and labour income, 0.25-0.27. This means that an increase of 10 percentile points in the rank of the parents would mean an increase of 2.5 to 3.1 percentile points in the rank of children.

Using the rich BHPS+UKHLS harmonised data, I examine the robustness of these main results to changes in the sample restrictions, model specifications and main variables used. While the main results use household income for parents and children with the objective of capturing the general living standard of the household, I find that IGEs using alternative measures of income (i.e. labour income and personal income) are very similar. I also find that the IGEs and rank coefficients are robust to changing age restrictions in the model and to the treatment of outliers. The levels of income mobility for the group of individuals who do not live with their parents and are at least 30 years old further reinforces the robustness of the initial results.

Both IGEs and rank coefficients lie in the (wide) range provided by previous studies. As these other studies are largely based on cohort data from the BCS and NCDS, any direct comparison must be done with care. In general, the estimated IGEs in this paper are lower than the previous findings by Blanden et al. (2004) for the BCS cohort (born in 1970), using children's earnings and family income, which were around 0.30 for men and 0.40 for women. Similar results were obtained in a more recent study by Belfield et al. (2017) for the BCS, but Gregg et al. (2017) estimate the elasticity for sons to be even higher (0.43) considering multi-year earnings observations for children and adjusting for workless spells. In relation to other studies that used the BHPS, the paper by Ermisch and Francesconi (2004) estimated an IGE of around 0.05-0.10 using a matched sample of individuals born between 1970-1983. However, because they were restricted to the first eight waves of the BHPS (until 1999) they observe sons' income at extremely young ages (at 16), which means that these numbers are likely attenuated by the life-cycle bias. Nicoletti and Ermisch (2008) estimate the IGE to be around 0.20-0.30 for a cohort of sons born from 1950-1972 using the TS2SLS approach with the BHPS data. To my knowledge, the work by Gregg et al. (2017) is the only other study to estimate rank coefficients for the UK and they estimate it to be around 0.30 for the BCS cohort, similar to the results I obtained.

Even though my results can be only compared to the previous findings very carefully because of different data sets, variables and model specifications, the fact that they are very robust to various sensitivity tests could indicate some of these choices do not matter too much if we believe that the common estimation issues of transitory shocks in income and life-cycle bias are being addressed properly. One possible explanation for some of these considerable differences from other studies is that they could be driven by the different cohorts being studied.My analysis contributes with a step in the right direction, yet it is still possible that these results are still a lower bound of the true IGEs and rank coefficients, due to residual attenuation bias.

Finally, in the last part of the paper, I also present regional estimates of elasticities and rank coefficients. My results indicate that there is a North and South divide in terms of intergenerational income mobility, with the South being much more mobile than the North.

This paper represents an advancement in the literature of intergenerational income mobility in the UK, yet many questions remain unanswered. This study will act as a platform for further investigation of the heterogeneity of intergenerational income mobility across several dimensions highlighted in the international literature, such as gender and differences across the income distribution. I plan to address some of these questions in future research.

References

Acciari, P., A. Polo, and G. L. Violante (2019). *'And Yet, it Moves': Intergenerational Mobility in Italy.* IZA Discussion Paper 12273, Institute for the Study of Labor (IZA).

Atkinson, A. B. (1980). On intergenerational income mobility in Britain. *Journal of Post Keynesian Economics 3*(2), 194–218.

Becker, G. S. and N. Tomes (1979). An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy 87*(6), 1153–1189.

Becker, G. S. and N. Tomes (1986). Human capital and the rise and fall of families. *Journal of Labor Economics 4*(3, Part 2), 1–39.

Belfield, C., C. Crawford, E. Greaves, P. Gregg, and L. Macmillan (2017). *Intergenerational income persistence within families.* IFS Working Paper W17/11.

Bell, B., J. Blundell, and S. Machin (2018). The changing geography of intergenerational mobility. *CEP Discussion Papers (CEPDP1591).*

Björklund, A. and M. Jäntti (1997). Intergenerational income mobility in Sweden compared to the United States. *American Economic Review 87*(5), 1009–1018.

Björklund, A. and M. Jäntti (2011). Intergenerational income mobility and the role of family background. In W. Salverda, B. Nolan, and T. Smeeding (Eds.), *The Oxford Handbook of Economic Inequality*, pp. 491–521. Oxford University Press.

Black, S. E. and P. J. Devereux (2011). Recent developments in intergenerational mobility. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 4, Chapter 16, pp. 1487–1541. Elsevier.

Blanden, J. (2009). *How much can we learn from international comparisons of intergenerational mobility?* CEE Discussion Papers 0111, Centre for the Economics of Education, LSE.

Blanden, J., A. Goodman, P. Gregg, and S. Machin (2004). Changes in intergenerational mobility in Britain. In M. Corak (Ed.), *Generational income mobility in North America and Europe*, Chapter 4, pp. 122–46. Cambridge: Cambridge University Press.

Blanden, J., P. Gregg, and L. Macmillan (2013). Intergenerational persistence in income and social class: the effect of within-group inequality. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 176*(2), 541–563.

Böhlmark, A. and M. J. Lindquist (2006). Life-cycle variations in the association between current and lifetime income: Replication and extension for Sweden. *Journal of Labor Economics 24*(4), 879–896.

Bratberg, E., J. Davis, B. Mazumder, M. Nybom, D. D. Schnitzlein, and K. Vaage (2017). A comparison of intergenerational mobility curves in Germany, Norway, Sweden, and the US. *The Scandinavian Journal of Economics 119*(1), 72–101.

Bratsberg, B., K. Red, O. Raaum, R. Naylor, M. Jantti, T. Eriksson, and E. Osterbacka (2007). Nonlinearities in Intergenerational Earnings Mobility: Consequences for Cross-Country Comparisons. *Economic Journal 117*(519), 72–92.

Chadwick, L. and G. Solon (2002). Intergenerational income mobility among daughters. *American Economic Review 92*(1), 335–344.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014). Where is the land of opportunity? the geography of intergenerational mobility in the United States. *Quarterly Journal of Economics 129*(4), 1553–1623.

Corak, M. (2004). Generational income mobility in North America and Europe: an introduction. In M. Corak (Ed.), *Generational Income Mobility in North America and Europe*, Chapter 1, pp. 1–37. Cambridge: Cambridge University Press.

Corak, M. (2006). Do poor children become poor adults? Lessons from a cross-country comparison of generational earnings mobility. *Research on Economic Inequality 13*(1), 143–188.

Corak, M. (2019, 05). The Canadian Geography of Intergenerational Income Mobility. *The Economic Journal*.

Dahl, M. W. and T. DeLeire (2008). *The association between children's earnings and fathers' lifetime earnings: estimates using administrative data*. Discussion Paper Number 1342- 08. University of Wisconsin-Madison, Institute for Research on Poverty.

Dearden, L., S. Machin, and H. Reed (1997). Intergenerational mobility in Britain. *Economic Journal 107*(440), 47–66.

Erikson, R. and J. H. Goldthorpe (2002). Intergenerational inequality: A sociological perspective. *Journal of Economic Perspectives 16*(3), 31–44.

Ermisch, J. and M. Francesconi (2004). Intergenerational mobility in Britain: new evidence from the British Household Panel Survey. In M. Corak (Ed.), *Generational Income Mobility in North America and Europe*, Chapter 7, pp. 147–189. Cambridge: Cambridge University Press.

Goldthorpe, J. H. (2013). Understanding–and misunderstanding–social mobility in Britain: The entry of the economists, the confusion of politicians and the limits of educational policy. *Journal of Social Policy 42*(3), 431–450.

Grawe, N. D. (2006). Lifecycle bias in estimates of intergenerational earnings persistence. *Labour Economics 13*(5), 551–570.

Gregg, P., L. Macmillan, and C. Vittori (2017). Moving towards estimating sons' lifetime intergenerational economic mobility in the UK. *Oxford Bulletin of Economics and Statistics 79*(1), 79–100.

Haider, S. and G. Solon (2006). Life-cycle variation in the association between current and lifetime earnings. *American Economic Review 96*(4), 1308–1320.

Heidrich, S. (2017). Intergenerational mobility in Sweden: a regional perspective. *Journal of Population Economics 30*(4), 1241–1280.

Jäntti, M., B. Bratsberg, K. Røed, O. Raaum, R. Naylor, E. Österbacka, A. Björklund, and T. Eriksson (2006). *American Exceptionalism in a New Light: A Comparison of Intergenerational Earnings Mobility in the Nordic Countries, the United Kingdom and the United States*. IZA Discussion Papers 1938, Institute for the Study of Labor (IZA).

Jäntti, M. and S. P. Jenkins (2015). Income mobility. In A. Atkinson and F. Bourguignon (Eds.), *Handbook of Income Distribution*, Volume 2A, Chapter 10, pp. 807–935. Elsevier.

Landersø, R. and J. J. Heckman (2017). The Scandinavian fantasy: The sources of intergenerational mobility in Denmark and the US. *The Scandinavian Journal of Economics 119*(1), 178–230.

Lee, C.-I. and G. Solon (2009). Trends in intergenerational income mobility. *The Review of Economics and Statistics 91*(4), 766–772.

May, T. (2016). Britain, the great meritocracy: Prime minister's speech. Delivered on 09/09/2016.

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 87*(2), 235–255.

Mazumder, B. (2016). Estimating the intergenerational elasticity and rank association in the United States: Overcoming the current limitations of tax data. In L. Cappellari, S. W. Polachek, and K. Tatsiramos (Eds.), *Inequality: Causes and Consequences*, Chapter 3, pp. 83–129. Emerald Group Publishing Limited.

Moffitt, R. A. and P. Gottschalk (2012). Trends in the transitory variance of male earnings methods and evidence. *Journal of Human Resources 47*(1), 204–236.

Narayan, A., R. V. der Weide, A. Cojocaru, C. Lakner, S. Redaelli, D. G. Mahler, R. G. N. Ramasubbaiah, and S. Thewissen (2018). *Fair Progress? Economic Mobility across Generations around the World*. Washington, DC: World Bank.

Nicoletti, C. and J. Ermisch (2008). Intergenerational earnings mobility: Changes across cohorts in Britain. *The B.E. Journal of Economic Analysis & Policy 7*(2), 1–38.

Nybom, M. and J. Stuhler (2016). Heterogeneous income profiles and lifecycle bias in intergenerational mobility estimation. *Journal of Human Resources 51*(1), 239–268.

OECD (2018a). *A Broken Social Elevator? How to Promote Social Mobility*. Paris, OECD Publishing.

OECD (2018b). *OECD Labour Force Statistics 2017*. Paris, OECD Publishing.

Office for National Statistics (2017). *Statistical Bulletin: Regional gross disposable household income (GDHI): 1997 to 2015.* Available at: https://www.ons.gov.uk/economy/regionalaccounts/grossdisposablehouseholdincome /bulletins/regionalgrossdisposablehouseholdincomegdhi/2015/. Accessed: 01 Nov 2017.

Raaum, O., B. Bratsberg, K. Røed, E. Österbacka, T. Eriksson, M. Jäntti, and R. A. Naylor. (2008). Marital sorting, household labor supply, and intergenerational earnings mobility across countries. *The B.E. Journal of Economic Analysis & Policy 7*(2), 1–49.

Shin, D. and G. Solon (2011). Trends in men's earnings volatility: What does the panel study of income dynamics show? *Journal of Public Economics 95*(7-8), 973–982.

Social Mobility Commission (2016a). *The Social Mobility Index.*

Social Mobility Commission (2016b). *State of Nation 2016: Social Mobility in Great Britain.* London: Stationary Office.

Social Mobility Commission (2017). *State of Nation 2017: Social Mobility in Great Britain.* London: Stationary Office.

Social Mobility Commission (2018). *Social Mobility Barometer: Public attitudes to social mobility in the UK.*

Solon, G. (1989). Biases in the estimation of intergenerational earnings correlations. *The Review of Economics and Statistics 71*(1), 172–174.

Solon, G. (1992). Intergenerational income mobility in the United States. *American Economic Review 82*(3), 393–408.

Solon, G. (1999). Intergenerational mobility in the labor market. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3, Chapter 29, pp. 1761–1800. Elsevier.

Solon, G. (2002). Cross-country differences in intergenerational earnings mobility. *Journal of Economic Perspectives 16*(3), 59–66.

Taylor, M. F., J. Brice, N. Buck, and E. Prentice-Lane (2010). *British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices.* Colchester: University of Essex.

Appendices

Appendix A: Income variables

The **household income variable** "sums the values of total income in the month before interview for individuals in the household" (Taylor et al., 2010, App2-5). This variable includes the sum of non-labour income and labour income for all individuals in the household. It is comprised of the following components:

- Household gross labour earnings (taken from wPAYGU This measures usual monthly wage or salary payment before tax and other deductions in current main job for employees, wJSPROF This computes a monthly self-employed profit variable for self-employed respondents who draw up profit and loss accounts, wJSPAYG This converts employees' last wage or salary payment before tax and other deductions in current main job (wPAYGL) to a monthly amount, as appropriate. Income from second and occasional jobs is also added if non-missing.

- And non-labour income:

  - Household investment income : This variable totals the estimated income from savings and investments, and receipts from rented property, received in the month before interview.
  - Household benefit income : This variable totals all receipts from state benefits (including NI retirement pension), received in the month before interview.
  - Household pension income : this variable totals all receipts from non-state pension sources, recieved in the month before interview.
  - Household transfer income : This variable totals all receipts from other transfers, (including education grants, sickness insurance, maintenance, foster allowance and payments from TU/Friendly societies, from absent family members), received in the month before interview

The **labour income** variable is the pre-tax labour income in month before interview. It is a derived variable comprised of all labour income sources (main job, second job, self-employed). In addition, the **personal income** variable is the sum of all the above mentioned non-labour and labour income sources at the individual level.

All these income variables include imputed data. The imputation flag variable takes a value 0 if there was no imputation, 1 if some component of an individual household member's income was imputed, and 2 if the whole income of one or more members was imputed. The BHPS guide recommends always using imputed data (Taylor et al., 2010, A5-22), in order to reduce the potential bias that would be caused by the elimination of observations with missing data. In this paper I use the imputed data.

Appendix B: Sample Representativeness

As described in section 4.3, various restrictions are adopted to get the main sample used throughout this paper. Firstly, individuals must be born between 1973-1991 in order to be eligible in terms of age. Secondly, they need to be matched to biological parents. Thirdly, there needs to be at least one observation of parental income during childhood years, and at the time of this observation parents need to be aged 25-60. Finally, individuals have to be interviewed on/after the age of 25 and have at least one observation of income.

After imposing these restrictions it is not possible to guarantee that the sample remains nationally representative. In this section, I provide some descriptive statistics of the sample used and also of the sample not used (after the two first conditions were met[44]). This can be explored from the perspective of both parents' and children's generations.

Table 14: Parental characteristics: main variables

	In sample [N=2102] Mean	Out of sample [N=2528] Mean
Parental household income (£/month)	3623.09	3494.67
	(2761.81)	(2448.17)
Parental age when income observed	42.1	41.0
Child age when parental income observed	13.1	12.6

Notes: Sample sizes in brackets. Standard deviation in parenthesis. The combination of both samples adds up to the number of individuals whose parents have at least one income observation available during childhood (N=4630).

Table 14 contains the means of income and age for parents in and out of the sample. It is possible to see that the parents in the sample, i.e. whose children are also in the sample with at least one observation of income after the age of 25, are on average richer (+ £130 per month). They are also one year older when this income is observed.

Table 15 presents some background data on employment work and education for mothers and fathers in and out of the sample. It shows that the proportion of parents with no qualifications is higher for those out of the sample. In addition, the fraction of parents working when their child is aged 14 is smaller. All this suggests that the pattern of attrition in the sample affects slightly more individuals from disadvantaged socio-economic backgrounds (i.e. with lower parental income and less qualified parents), who are not included in the final matched sample.

[44]It only makes sense to compare individuals from the same cohorts and who can also be matched to biological parents.

Table 15: Parental characteristics: employment and education

	In sample [N=2102]	Out of sample [N=2528]	Out of sample [N=3492]
Panel A: Fathers	[1542]	[1632]	[2324]
Highest Education	[1466]	[1505]	[2040]
Degree (%)	14.26	14.22	12.79
No qualifications (%)	17.94	23.72	27.11
Retrospective information: employment	[1542]	[1536]	[2324]
Father working when child aged 14 (%)	76.20	67.90	69.23
Father not working when child aged 14 (%)	9.99	11.46	10.11
Missing info at age 14 (%)	13.81	20.64	20.66
Panel B: Mothers	[2043]	[2369]	[3273]
Highest Education	[2021]	[2332]	[3135]
Degree (%)	12.12	10.07	9.06
No qualifications (%)	20.44	24.57	25.07
Retrospective information: employment	[2043]	[2246]	[3273]
Mother working when child aged 14 (%)	65.20	59.13	56.28
Mother not working when child aged 14 (%)	25.45	27.16	27.50
Missing info at age 14 (%)	9.35	13.71	16.12

Notes: Sample sizes in brackets. The combination of the samples in the first two columns adds up to the number of individuals whose parents have at least one income observation available during childhood (N=4630). The combination of individuals in the sample plus the individuals in the out of sample in the last column adds up to all individuals born between 1973-1991 matched to their biological parents (N=5594).

Appendix C: TSRA First Stage Regressions

Table 16: TSRA First Stage Auxiliary Regressions

	(1) Log Parental Income	(2) Log Child Income
Parental age	0.15***	
	(0.015)	
Parental age squared	-0.0015***	
	(0.00018)	
Child age		-0.0014
		(0.023)
Child age squared		0.00050
		(0.00037)
Year dummies	Yes	Yes
Constant	4.30***	8.01***
	(0.32)	(0.35)
Observations	15216	12744

Clustered standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Notes: This Table shows the first stage auxiliary regressions of the TSRA approach. Parental and child income are taken as household income and the first stage regressions control for age, age squared and year dummies.

Appendix D: Relaxing age restrictions (children born 1973-1986)

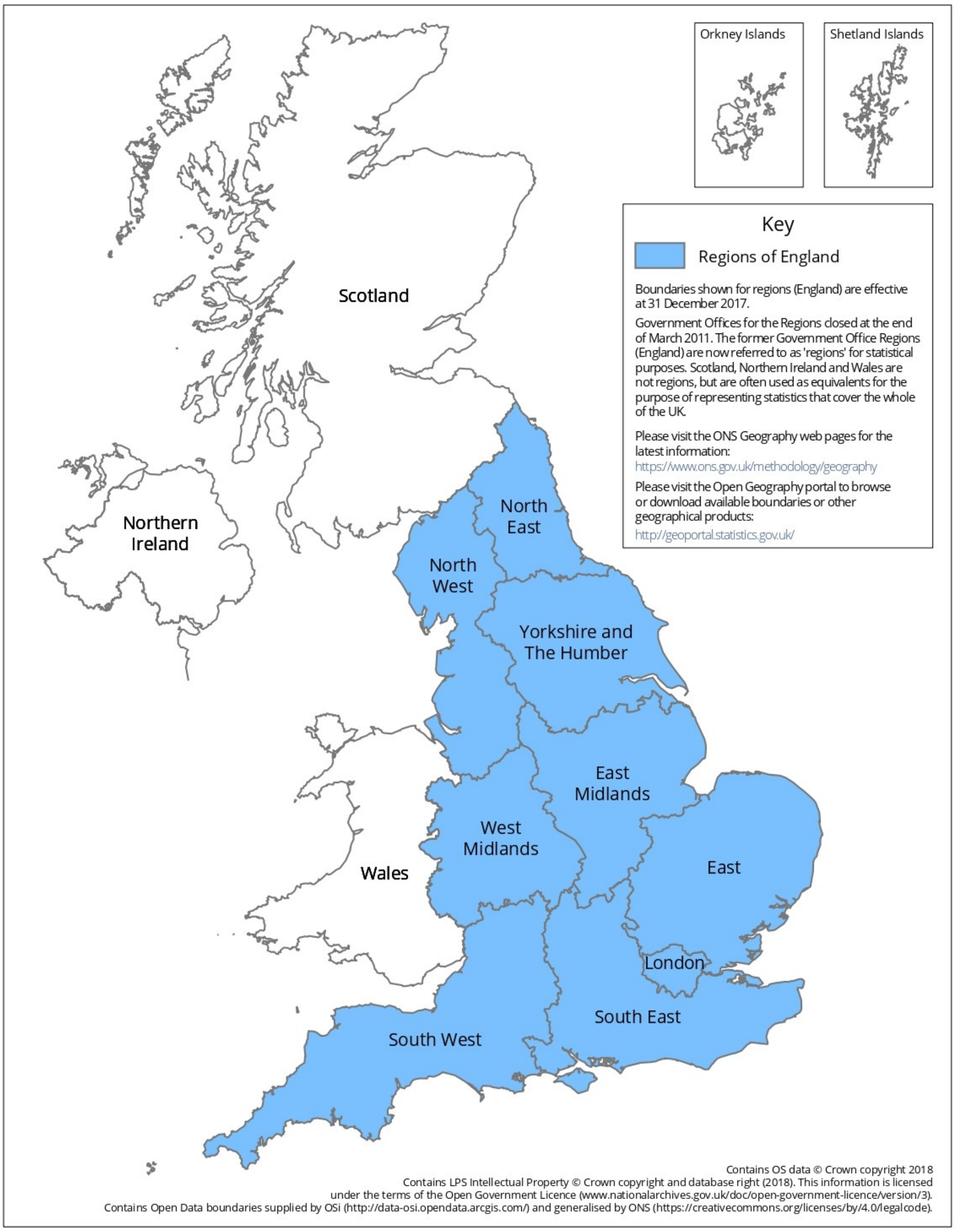
Table 17: IGE and rank coefficient when household income and personal income are measured at different ages (children born 1973-1986)

Child income	Parental income	N	Age	IGE (OLS)	IGE (TSRA)	Rank
Household income	Household income	2347	20-43	0.36***	0.32***	0.42***
				(0.024)	(0.024)	(0.027)
		1713	24-43	0.29***	0.28***	0.38***
				(0.031)	(0.027)	(0.027)
		1484	26-43	0.29***	0.27***	0.34***
				(0.033)	(0.031)	(0.027)
		1066	30-43	0.26***	0.25***	0.30***
				(0.034)	(0.036)	(0.032)
		819	32-43	0.31***	0.29***	0.31***
				(0.038)	(0.038)	(0.037)
		620	34-43	0.24***	0.24***	0.26***
				(0.041)	(0.043)	(0.043)
Personal income	Household income	2275	20-43	0.22***	0.15***	0.24***
				(0.027)	(0.026)	(0.027)
		1689	24-43	0.28***	0.25***	0.27***
				(0.034)	(0.032)	(0.027)
		1470	26-43	0.32***	0.29***	0.27***
				(0.037)	(0.035)	(0.027)
		1062	30-43	0.30***	0.29***	0.25***
				(0.047)	(0.046)	(0.033)
		813	32-43	0.33***	0.35***	0.26***
				(0.052)	(0.052)	(0.040)
		617	34-43	0.27***	0.28***	0.21***
				(0.067)	(0.068)	(0.046)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Notes: In the OLS and rank models, controls for the average age of parents and children and birth year dummies are used. In the TSRA model, I control for age and age squared of parents and children, as well as year dummies. For the rank model, due to the use of income in levels (not log), the sample size (N) increases for all groups (by a maximum of 23 pairs, at ages 20-43 and smaller for other ages).

Appendix E: Map of the UK and Regions of England



Source: ONS Geography Open Data