THE CENTRE FOR MARKET AND PUBLIC ORGANISATION

Measuring the intergenerational correlation of worklessness

Lindsey Macmillan

December 2011

Working Paper No. 11/278

Centre for Market and Public Organisation
Bristol Institute of Public Affairs
University of Bristol
2 Priory Road
Bristol BS8 1TX
http://www.bristol.ac.uk/cmpo/

Tel: (0117) 33 10952
Fax: (0117) 33 10705
E-mail: cmpo-admin@bristol.ac.uk

The Centre for Market and Public Organisation (CMPO) is a leading research centre, combining expertise in economics, geography and law. Our objective is to study the intersection between the public and private sectors of the economy, and in particular to understand the right way to organise and deliver public services. The Centre aims to develop research, contribute to the public debate and inform policy-making.

CMPO, now an ESRC Research Centre was established in 1998 with two large grants from The Leverhulme Trust. In 2004 we were awarded ESRC Research Centre status, and CMPO now combines core funding from both the ESRC and the Trust.

ISSN 1473-625X

CMPO Working Paper Series No. 11/278

Measuring the intergenerational correlation of worklessness

Lindsey Macmillan

Dept of Economics and CMPO, University of Bristol

December 2011

Abstract
This research uses the vast developments in the measurement of the intergenerational earnings mobility correlation over the past twenty years to explore the issues surrounding the measurement of the intergenerational correlation of worklessness. The correlation is estimated for a range of data sources. The role of conventional biases, measurement error and life-cycle bias, are considered in this context. An additional bias driven by local labour market conditions is introduced. For the UK, this correlation is moderate with large economic implications. Measurement error takes a different form to that commonly observed in the mobility literature but does not appear to play a substantial role in this story. In contrast to the mobility literature, life-cycle bias may not be playing a role either. Instead, there appears to be an additional bias driven by local labour market conditions at the time of measurement that should be considered when measuring intergenerational worklessness.

Keywords: Intergenerational mobility, unemployment

JEL Classification: J62, J64

Electronic version: www.bristol.ac.uk/cmpo/publications/papers/2011/wp278.pdf

Acknowledgements
I would like to thank Paul Gregg, Christopher Muller, Paul Devereux, Kjell Salvanes and Bhashkar Mazumder for their insightful comments, as well as comments from the UCD economics workshop and the CMPO intergenerational mobility conference. This work is funded by the ESRC as part of an ESRC studentship.

Address for correspondence
CMPO, Bristol Institute of Public Affairs
University of Bristol
2 Priory Road
Bristol BS8 1TX
Lindsey.MacMillan@bristol.ac.uk
www.bristol.ac.uk/cmpo/

1. Introduction

Research into the correlation of workless spells across generations has been largely overlooked in the growing body of literature on intergenerational transmissions over the past two decades (see Black and Devereux (2011) and Solon (1999) for comprehensive reviews). This is despite the fact that many of the intergenerational earnings mobility measures within economics only capture the intergenerational correlation of the employed. These measures are typically based on a regression of the log earnings of the 2nd generation measured at a certain point in time on the log earnings of the 1st generation. This puts constraints on the sample we are able to measure mobility for as to have earnings you must be in employment. This research aims to analyse the other side to this story by considering the intergenerational correlation of worklessness.

Worklessness is considered here instead of unemployment as the definition of unemployment is very narrow, often defined as those who are actively seeking work. This therefore only captures a transitory state of being out of work but trying to get back into work. This is difficult to measure if we only observe individuals at a point in time and only captures part of the story for those who are out of work. Worklessness on the other hand captures a wider group of individuals; both those who are out of work and are not seeking to be in work as well as those in transitory unemployment.

Individuals' who experience workless spells will often be found at the bottom of any income distribution. Johnson and Reed (1996) argue that the 'exclusion from society' of individuals who are from the poorest groups makes them of interest for a variety of economic, social and political reasons. An intergenerational correlation of worklessness may be of far more concern to policy makers than any movement or lack of movement around the middle or top parts of the mobility distribution. Despite this there has been little evidence to date on this relationship. As in the intergenerational mobility literature, one reason for this lack of evidence may be that the measurement of such a correlation leads to many problems. This research utilises the major developments in the measurement of the intergenerational elasticity of earnings and income over the past twenty years to explore the many issues that arise when trying to capture intergenerational correlations including measurement error and life-cycle bias.

Solon (1992) and Zimmerman (1992) first drew attention to the issue of measurement error within the intergenerational mobility literature when attempting to estimate correlations in income across generations for the United States. The issue is that we wish to measure the correlations of the permanent or lifetime state for both generations, be it income or in this case spells spent out of work. Therefore to capture lifetime intergenerational workless spells, the researcher requires information on the work spells of both generations throughout their adult life. As is the case with income, this kind of longitudinal data is rare. Instead what we commonly observe are snap-shots of the income or work experiences of each generation at certain points in time. This can lead to attenuation bias in our estimates as our 1st generation measures only proxy their lifetime equivalents and therefore suffer from errors-in-variables bias.

More recently, further work by Haider and Solon (2006) and Grawe (2006) has drawn attention to a separate form of errors-in-variables bias often found in these types of measurements; life-cycle bias. Unlike the more straightforward measurement error discussed above, this bias effects both the 1st and 2nd generation measures as illustrated in Grawe (2006). This is because individuals' earnings trajectories, or in this case their propensity to experience workless spells, are not stable across the life-cycle and more specifically can vary by family background. Therefore the window in which we view the snapshot of data for each generation can bias our estimates of the intergenerational elasticity. The direction of this bias will depend on when in the life-cycle the individual is observed and the trajectory of the intergenerational coefficient across the life-cycle.

In addition to the existing biases noted in the intergenerational literature that will lead to a mis-measurement of the intergenerational correlation, there may be a further bias to consider when measuring intergenerational worklessness driven by trends in regional employment patterns and business cycle shocks. Employment is typically more responsive to negative shocks than real wages leading to larger employment shocks than wage shocks (Romer, 2006). Therefore when considering intergenerational worklessness, these regional and business cycle shocks may matter more than in the intergenerational earnings mobility literature. Individuals' living in areas of higher unemployment or experiencing periods of higher unemployment at the snap-shot of time that they are observed may face higher probabilities of being unemployed. If this differentially impacts those with workless fathers' compared to those with employed fathers' this will lead to a bias driven by the outside local labour market conditions.

Alongside these advancements in the intergenerational research, this research is particularly relevant as we enter a time of increasing interest in the public domain as to the extent of the problem of generations of families that have never worked. To date, there is very little empirical evidence to inform the debate. In particular, the extent to which individuals' from the same families experience spells of worklessness, or never work has not been measured for any recent data.

To give a sense of scale at the population level, estimates from the April to June quarter of the Labour Force Survey in 2010 indicate that there are 3.668 million households[1] in the UK where nobody of working age is working (see table 1). This number is non-trivial representing over 18% of all working age households in the UK. Restricting the analysis to households with two or more generations co-residing 4% of multi-generational households are in a position where both generations are workless. For these households, intergenerational worklessness is presenting a real problem in terms of the duration of periods that generations are spending workless. However, contrary to some commentary on this subject, there are very few households where both generations have never worked. Only 15,350 households in the UK have two or more generation who report to have 'never worked' and of these, many of the younger generation have only been out of education for less than a year (Table 1, panel B). Of course these figures are restricted to both generations co-residing in the same household. The analysis in this research uses the British Household Panel Survey and two longitudinal birth cohort studies to relax this restriction to allow us to observe the 2nd generation in their own adult households.

This research provides an important contribution to the intergenerational literature, quantifying the scale of the intergenerational correlation of worklessness and exploring the issues in measuring such correlations. It introduces a new element to consider when thinking about measurement error in this setting where binary variables are often used. It also introduces an additional bias to consider in this context driven by local labour market conditions. The next section reviews the previous literature on the measurement of the intergenerational earnings elasticity and intergenerational workless spells in the UK. Section three discusses the

[1] Defined as households with at least one working age adult and excluding student households.

methodological issues in measuring the correlation accounting for measurement error, life-cycle bias and any biases resulting from local labour market conditions. Section four discusses the data sources used, the measurement issues associated with each source and describes some summary statistics. The main results are presented in section five while section six concludes.

2. Literature review

The intergenerational correlation in income has been extensively researched over the past thirty five years. As noted, by contrast little has been done to date on the intergenerational correlation of workless spells. Work dating back to 1975 by Sewell and Hauser through to Becker and Tomes (1986) attempted to measure the elasticity of intergenerational mobility in the US and found correlations of around 0.2. Since then there has been a large push within the literature to improve the measurement of the intergenerational correlation with a large number of influential researchers contributing to the issue (Solon, 1999). Black and Devereux (2011) illustrate that more recently the mobility literature has evolved to shift the focus to investigating the potential causal mechanisms behind such intergenerational transmissions. While this is an important advancement, the first step in the mobility literature and here within the intergenerational worklessness research is to pin down the precise measurement of the intergenerational correlation before mechanisms or causality can be considered.

Solon (1992) and Zimmerman (1992) were the first pieces of research within the intergenerational mobility literature to draw attention to the potential biases arising from measurement error. They pointed out that basing the estimates for the correlation of intergenerational mobility on data on income at a point in time led to a downward bias in the estimated coefficient. In addition, previous work that had not used representative samples, for example, Behrman and Taubman (1985), were also likely to be underestimating the magnitude of the intergenerational correlation. Using the Panel Study of Income Dynamics (PSID) and the log earnings of the fathers averaged across a range of periods from one year to five years the intergenerational correlation of father's earnings and son's earnings for the US between 1967 and 1971 was 0.413, double the correlation found in previous studies. Dearden, Machin and Reed (1997) further illustrated the bias arising from measurement error in intergenerational income mobility using UK data.

More recently Mazumder (2005), Haider and Solon (2006) and Grawe (2006) drew attention to an additional bias within the intergenerational mobility literature, noted but not brought out by previous measurement papers in the area. This is driven by the age at which both generations are observed in the data. Haider and Solon (2006) state that much of the research within the intergenerational mobility literature ‘devoted considerable attention’ to classical measurement error attempting to proxy long-run 1st generation measures with shorter term measures available in the data. Much less attention was given, however to the non-classical measurement error which affects both the 1st and the 2nd generation measures. Grawe (2006) shows that the structure of this non-classical error can be modelled by considering age-earnings profiles in both generations. This is because individuals with higher education do not reach their full potential in the labour market until later in their lives compared to individuals with lower education. The returns to higher education are often not realised in full until the individual reaches the age of 40 (Lee and Solon, 2009). Mazumder (2005) points out that for earnings, the transitory component follows a U-shape across the life-cycle meaning that while this component is lowest during an individuals’ 40s, later measures as well as early measures could bias the relationship. He suggests that when considering both biases, the correlation in the US could be as large as 0.6.

An additional bias that may affect the intergenerational correlation of worklessness is the impact of local labour market conditions. Romer (2006) illustrates that employment is highly pro-cyclical whereas real wages are at best only mildly pro-cyclical (Stadler, 1994). Regardless of the type of recession faced, the employment shock is usually larger than the wage shock (Gregg and Wadsworth, 2011). The intergenerational income literature therefore may not need to pay as much attention to the labour market conditions at the time that a snap-shot of income is observed. When considering intergenerational worklessness however the observed employment status is likely to be highly correlated with the local labour market conditions at the time of observation.

The intergenerational correlation in workless spells has been given far less attention. To date, there are only a couple of studies in the UK that touch on the issue, dating back to Johnson and Reed (1996) and O’Neill and Sweetman (1998). These studies use only the first UK birth cohort study available, the NCDS, a cohort born in 1958. More recently, Ekhaugen (2009) uses Norwegian data to estimate the intergenerational correlation in unemployment and there are related studies on the intergenerational correlation in welfare dependency using Canadian and Swedish data (Corak, Gustafsson and Osterberg, 2000) and US data (Gottschalk, 1996, Levine and Zimmerman, 1996). Page (2004) discusses the issues with measurement in the context of the intergenerational welfare dependency literature drawing attention to only viewing short windows of welfare receipt. These studies however often focus on mothers and daughters and therefore implicitly are considering the issue of intergenerational lone parenthood. This work focuses on fathers and sons, in line with earnings mobility research.

3. Methodology

i) Measurement error

In order to capture the correlation in worklessness across generations we would ideally want to measure the coefficient beta from a reduced form regression of the son's work history throughout their entire working adult life, $w_i^{son^*}$, on their father's work history throughout their entire working adult life, $w_i^{father^*}$.

$$w_i^{son^*} = \alpha + \beta w_i^{father^*} + e_i \tag{1}$$

As in the earnings mobility literature the aim is to capture as close to a lifetime estimation of the intergenerational coefficient as is possible.

The main potential source of measurement error therefore arises if the 1st generation are only observed for a small window of time, represented by w_i^{father} from equation (2) rather than their entire working adult life, $w_i^{father^*}$. This introduces error, ε_i, to our measurement of $w_i^{parent^*}$.

$$w_i^{father} = w_i^{father^*} + \varepsilon_i \tag{2}$$

When using continuous variables such as earnings, the error term has an expectation of zero and therefore there is classical measurement error leading to attenuation bias in the estimation. It can be shown that this will bias down the estimate of the intergenerational coefficient as seen in (3).

$$plim\hat{\beta} = \beta \frac{\sigma_w^2}{\sigma_w^2+\sigma_\varepsilon^2} \tag{3}$$

If the 1st generation are only observed for a short window, for example at a point in time, in this context we are often forced to place them into a binary category of employed or workless. Aigner (1973) shows that measurement error for binary explanatory variables differs from conventional classical measurement error as the expected value of the error term is no longer zero. To illustrate this, consider the joint frequency distribution of w_i^{father} and ε_i in table (4). If fathers were to be observed for a longer window, it would be optimal to use a continuous distribution of the proportion of time spent out of work throughout the period to measure their lifetime work history.

$\boldsymbol{\varepsilon_i}$ \ $\boldsymbol{w_i^{father}}$	**0**	**1**	$\boldsymbol{f(\varepsilon_i)}$
$\boldsymbol{x_i}$	Ew	0	**Ew**
0	(1-w)E	(1-e)W	**(1-w)E+(1-e)W**
$\boldsymbol{y_i}$	0	We	**We**
$\boldsymbol{f(w_i^{father})}$	**E**	**W**	**1**

(4)

For the dichotomous explanatory variable of the father being employed or workless, W represents the proportion of fathers who are observed as workless at a point in time. Of these, a proportion ‘e’ would be observed in work for at least some of the time in the continuous, longer window, setting. Likewise, (1-W) or E fathers are observed employed at a point in time whereas ‘w’ of these would be observed as workless for some of the time in a continuous setting. The amount of time that the proportions ‘e’ and ‘w’ would be observed employed or workless in the longer window is given by the distributions x_i and y_i. These may not be symmetrical as the persistence in worklessness may vary from persistence in employment. The expected value of the error term is therefore

$$E(\varepsilon_i) = Ewx_i + Wey_i \quad (5)$$

with the variance $Var(w_i^{father})$ and covariance$Cov(w_i^{father}, \varepsilon_i)$ given by

$$Var\left(w_i^{father}\right) = WE \quad (6)$$

$$Cov\left(w_i^{father}, \varepsilon_i\right) = (ey_i - wx_i)WE \quad (7)$$

The probability limit of $\hat{\beta}$ in this context is therefore

$$plim\hat{\beta} = \beta(1 - ey_i + wx_i) \quad (8)$$

A longer window of data could therefore be used to assess the values of ‘w’ and ‘e’ and the distributions of x_i and y_i. This could give an indication of the size of the issue. The measures available in each of the data sources are described in the next section.

In addition to this main potential source of error, measurement error of this type can also arise for other reasons such as reporting error in the employment status, unrepresentative samples or recall bias when responding to questions about a distant period of time. The likely presence of each type of error in this analysis will be discussed in section 4ii). It is important to note that this specific bias only affects the explanatory variable. Any similar error in the dependent variable will not affect the estimate of β. As will be discussed in the next section this is not true for life-cycle bias.

ii) Life-cycle bias

Biases resulting from life-cycle effects are not only affected by *if* the individual is only observed at a snap-shot of time but more importantly *when* that individual is observed. As with the life-cycle bias in incomes illustrated by Mazumder (2005), Haider and Solon (2006) and Grawe (2006), this is because different individuals are likely to face differing probabilities of experiencing workless spells dependent on both how old they are and their background. While on average individuals tend to age out of worklessness; someone straight out of school will likely face a higher probability of experiencing a workless spell than someone with more experience in the labour market[2], the rate individuals' age out of worklessness may vary by fathers' workless experiences. Unlike simple attenuation bias, the impact of errors-in-variable bias due to life-cycle effects can also bias the measurement of beta upwards. In addition, unlike measurement error, the error from life-cycle bias is not restricted to the 1st generation. The 2nd generation are affected by the bias as well.

As stated previously, in the intergenerational mobility literature, the trajectory of the intergenerational coefficient can be modelled by analysing age-earnings profiles across time (Grawe, 2006). It is possible therefore to assert that if the 1st generation is aged 40 but the 2nd generation is younger, it is likely that the intergenerational coefficient will be downward biased. When considering life-cycle bias in workless spells it is not necessarily the case that the intergenerational trajectories in worklessness follow the same pattern as age-earnings profiles.

By considering the age profiles of the proportion of time spent out of work for sons with workless fathers compared to sons with employed fathers we can get a sense of any pattern emerging in the difference between trajectories as individuals' age. This can also be estimated more explicitly by considering the estimated interaction term, $\hat{\theta}$, between the age of the son and the fathers' workless status from equation (9). This shifts the focus to person time observations by regressing the proportion of time the 2nd generation individual i at time t, w_{it}^{child}, spends workless each year on the 1st generation workless experience using an OLS model, clustered at the individual level. A vector of age controls, $\boldsymbol{A}_i$, are included to remove any variation in age within the sample.

$$w_{it}^{son} = \alpha + \beta w_i^{father} + \theta a_{it}^{son} * w_i^{father} + \boldsymbol{A}_i \gamma + e_i \quad (9)$$

If the pattern is consistent across data sources, it may be plausible to assert an optimal age to estimate the intergenerational coefficient, as is the case in the mobility literature.

iii) ***Local labour market conditions***

---

[2] The unemployment rate for 16/17 year olds as of May-July 2011 was 36.9% compared to 18.7% for 18-24 year olds, 7.8% for 25-34 year olds and 5.3% for 35-42 year olds (ONS http://www.ons.gov.uk/ons/taxonomy/index.html?nscl=Unemployment+by+Age, 2011)

In addition to age effects across time, the intergenerational coefficient could be affected by changes in local labour market conditions that directly impact the individual's outside option in terms of the probability of finding a job. As noted, employment is highly pro-cyclical and more responsive to local labour market conditions than we might expect wages to be (Romer, 2006). Any potential bias could therefore be driven by both differences in where in the country the individual grows up and differences in their experience of the business cycle at different stages of their career.

Using LEA level information as a proxy for local labour market conditions on entry, a within-LEA model can be estimated with an LEA fixed effect, δ_r, to remove any local-area-specific effects from the estimated intergenerational coefficient.

$$w_{ir}^{son} = \alpha + \beta w_{ir}^{father} + \delta_r + \boldsymbol{A}_i\gamma + e_{ir} \tag{10}$$

By taking fixed effects of model (10), any unobserved heterogeneity driven by differences in employment experiences across regions can be removed, estimating the intergenerational coefficient for individuals within their own LEA. This model removes the fixed effect by using a centred observation, a deviation from the LEA average workless level, rather than a level. Therefore generations' that live in high unemployment areas, that both experience high proportions of time out of work appear less like an outlier as their LEA average proportion of time out of work is higher and hence their deviation from this is lower. However, generations that live in low unemployment areas that both still experience high proportions of time out of work will drive the data estimation as an intergenerational correlation will exist despite their local labour market experience. The reduction in the estimated intergenerational correlation from the baseline estimate indicates the extent to which the correlation is driven by differences across local areas.

To enhance this specification, more detailed information on county level unemployment rates across time can be used. Unemployment rates across time are only available for the slightly more aggregated county level rather than LEA level data. Equation (11) illustrates this specification, making use of the annual information on both the unemployment rate and the proportion of time spent workless. Now the focus is on person time observations by regressing the proportion of time the 2nd generation individual i in region r at time t, w_{irt}^{son}, spends workless each year on the 1st generation workless experience using an OLS model, clustered at the individual level controlling for the county level unemployment rate for each year, u_{rt}.

$$w_{irt}^{son} = \alpha + \beta w_{ir}^{father} + \tau u_{rt} + \boldsymbol{A}_i\gamma + e_{irt} \tag{11}$$

Any shift in the intergenerational correlation driven by model (10) or (11) tells us how much of the correlation is driven simply by fathers and sons living in the same local labour markets. However, it may instead be the case that the impact of the local labour market conditions on the 2nd generation varies by whether the 1st generation experience workless spells or not, much in the same way that the age-profiles of the 2nd generation may vary by the workless experiences of the 1st generation. In this case, the local labour market experience at the point of observation would directly bias any correlation across generations. To consider this, as in the case of life-cycle bias, an interaction between the county level unemployment rate and the 1st generation workless experience, $u_{irt} * w_i^p$, can be included to remove any differential local labour market effects.

$$w_{irt}^{son} = \alpha + \beta w_{ir}^{father} + \theta w_{ir}^{p} * u_{rt} + \tau u_{rt} + \boldsymbol{A}_i\gamma + e_{irt} \tag{12}$$

While the reduction in the correlation indicates any potential bias driven specifically by differences in the unemployment rate across different counties, the coefficient on the interaction term illustrates that specific impact of differential unemployment rates by counties at different time periods on those individuals with workless fathers. The benefit of using an interaction term in this setting is that from this we can derive information about the size of the intergenerational correlation at different levels of local unemployment, based on where the interaction is evaluated. In the baseline model (12), the intergenerational coefficient is evaluated at the average unemployment rate in the sample across all years, $\bar{u}$. To show this, a new parameter can be defined as the total effect of fathers workless spells on sons workless spells, $\varphi = \beta + \theta(\bar{u})$. This can be rearranged and substituted into (12)

$$w_{irt}^{son} = \alpha + \varphi_k w_{ir}^{father} + \theta w_{ir}^{p} * (u_{rt} - \bar{u}) + \tau u_{rt} + \boldsymbol{A}_i\gamma + e_{irt} \tag{13}$$

As within-county unemployment rates range from 1.6% to 23.3% across time within the sample of interest, (13) can be estimated across this range of values, substituting the average unemployment rate, $\bar{u}$, for different values of the unemployment rate $u = [2, 3, 4 \dots .23]$, resulting in a range of estimates of the intergenerational coefficients, $\hat{\beta}_k = \hat{\varphi}_k - \hat{\theta}(u_k)\ \ for\ k = 2 \dots 23$

$$w_{irt}^{son} = \alpha + \varphi_k w_{ir}^{father} + \theta w_{ir}^{p} * (u_{rt} - u_k) + \tau u_{rt} + \boldsymbol{A}_i \gamma + e_{irt} \tag{14}$$

The range of estimates can then be assessed to illustrate any differential impact on the intergenerational correlation of worklessness of differential unemployment rates across counties at various points in time.

Data

i) Data description

Three different data sources are used for this analysis, each with their own strengths and weaknesses that will be discussed in the next section. The two British birth cohort studies, the National Child Development Survey (NCDS) of all individuals born in one week in March, 1958, and the British Cohort Study (BCS) of all individuals born in one week in April 1970 are familiar datasets within the intergenerational mobility literature given their longitudinal nature. The British Household Panel Survey (BHPS) is beginning to be used in this context as the 2nd generation age into adulthood. The BHPS, unlike the cohort studies, is structured as a panel of households with all individuals within a survey household entering into the survey as they reach the age of 16. Individuals are then followed as they start their own households and new members entering into the new households also form part of the survey. The cohort studies, by contrast, follow the same individuals across their lifetimes at various ages with questions directed to the parents of cohort members throughout their childhood.

The employment measures available across the three surveys vary. In the BHPS, continuous work histories are available (Halpin, 1997) from 1990 until 2005. Within this file information from the employment status at the time of interview, throughout the last year and retrospective data, limited here to 1975 onwards, are combined to form episodes of different employment statuses throughout their adult lives. This data can be transformed into various different measures of workless spells. Generations can be linked within the data using the mother and father identification variables. From this, 1st and 2nd generation monthly work histories can be constructed from the series of episodes for groups of people where data is available in both generations. Individuals in both generations are defined as workless each month if they are not in employment or education with various continuous and more discrete measures of total worklessness then created including the proportion of time spent out of work and whether they have never worked. The major benefit of the BHPS over the cohort studies is the availability of continuous information on the 1st generation's workless experiences.

Various sample restrictions are placed on the BHPS data to minimise biases and maximise comparability across data sources. The first restriction is that both generations must be observed within the data for over two years. As table 2 shows, on average in the final sample the 1st generation are observed for 103 months and the 2nd generation for 90 months. The second restriction placed on the data is that the 1st generation must be observed before the 2nd generation turn 18. This is to ensure that the 1st generation work history occurs during childhood. The third is that the 2nd generation must be born before 1982. This is to ensure that the entire 2nd generation sample has a chance to reach age 23 in the latest wave of data so that those entering higher education can be observed for two years after leaving full time education. For this analysis the sample is also restricted to fathers and sons to remove participation issues for women.

In the cohort studies, fathers are observed at two discrete points in time, when the cohort member (son) is age 11 (1969) and 16 (1974) in the NCDS and 10 (1980) and 16 (1986) in the BCS. As in the BHPS the focus is on fathers and sons throughout. In the NCDS the question asks about the father's occupation at 11 and 16, requesting that if they are not currently working to put 'not working'. These are coded as workless. In the BCS the question asks the 'current (present) employment situation' of the father at 10 and 16. Both are coded as workless if the response is anything but 'regular paid job' or 'works occasionally'. For the 2nd generation, the cohort studies provide monthly work history data from 16-42 in the NCDS and 16-30 in the BCS (Galindo-Rueda, 2002). Sons are defined as workless if not in employment or education at each month observed with various continuous and more discrete measures of worklessness created including the proportion of time spent out of work and whether the son has spent a year or more in concurrent spells out of work or has never worked. The sample is restricted to sons and fathers with work history information available for the 2nd generation and at least one employment status observed for the 1st generation. The implications of this restriction on the 1st generation's employment status are discussed in section 4 iii).

In addition to the main employment status variables, data on the age of the 1st and 2nd generation across all three cohorts is used. For the local labour market analysis, we use local education authority (LEA) data from the BCS. Individuals within the sample live in 115 different LEAs in 1986 with an average of 32 final sample members per LEA. When the analysis is aggregated up to county level data, individuals live in 53 different counties with an average of 69 sample members per county. 21% of individuals in the final sample do not have LEA information available for this analysis. The implications of this are discussed in the results section 5 iii). Unemployment rates were matched into the BCS data using county level information on the ILO unemployment rate for the whole county from the Employment Gazette from 1986 to 1998. Each individual was assigned an annual unemployment rate for the county that the LEA was in based on the LEA they were observed in in 1986.

ii) Measurement issues

In order to minimise the impact of attenuation bias, we ideally want to observe the 1st generation for as long as possible in their adult lives. As noted, in the cohort studies the 1st generation are only observed at two points in time during the cohort members childhood, at 11(10) and 16 in the NCDS (BCS), and so are likely to be affected to some degree by attenuation bias. In the BHPS, longer spells of employment status are available for the 1st generation allowing us to move towards a lifetime measure of worklessness in the explanatory variable[3]. On average the 1st generation are observed for 103 months although the range is from 24 months to 276 months. In this case the average father is observed across the son's formative years from age 10-18. Robustness tests on restricting the window to longer periods are included in the analysis. By contrasting a continuous 'longer window' measure in the BHPS with a dichotomous 'small window' measure, constructed to replicate those available in the

[3] Of course, the 1st generation are also not observed across their entire working life in the BHPS but relative to the cohort studies, this is an improvement.

NCDS and BCS as closely as possible, any likely impact of measurement error in this context can be assessed.

All three data sources used are nationally representative although there may be some concern that they suffer from attrition, particularly in the cohort studies due to their longitudinal nature. Selection bias in samples can also lead to attenuation bias as discussed in Solon (1992). Table A1 compares the 1st generation unemployment rates in the cohort studies to the national unemployment rates for men aged 16 and over for corresponding years. Although this comparison is not perfect, the rates are very similar which is reassuring. Given the monthly nature of the 2nd generation measures there is no obvious national comparison group. At birth, the cohort members were a nationally representative group but the concern is that the individuals' for whom monthly work history data is available might vary systematically from those individuals' for whom this data is not available due to attrition. Table A2 compares parental education, fathers' class and 2nd generation IQ test score measures for the cohort members that have work history information and the cohort members that do not. The two samples are very similar in terms of characteristics although there is a suggestion that those who do not have work history information are from slightly less educated parents in lower social classes who do slightly worse on their IQ test at age 10 in both cohorts.

Problems with recall bias may occur if the 1st generation is asked to provide retrospective information on their work histories. In the cohort studies the father responds to questions about his current employment status and so this is unlikely to cause a problem. In the BHPS some of the information used in the work history data is retrospective, although limited to only as far back as 1975 to minimize the impact of recall bias. Despite this restriction, there may be some bias from those reporting work histories retrospectively. An important point is that recall bias will only affect the estimate of the intergenerational correlation if employed fathers recall things in a different way to workless fathers.

While the BHPS has the advantage of providing a continuous workless measure for fathers, sons are only observed at the early stages of their labour market experience given the sample design. This may lead to life-cycle bias as discussed in section 3 ii). The cohort studies by contrast have more complete monthly work history for sons. Using this monthly work history data, various measures of workless spells can be created in the cohort studies for the life-cycle bias analysis ranging from a yearly proportion of time spent workless, the proportion of time spent workless across the whole period, to more durational measures such as spending a year or more out of work or never working. The durational measures aim to directly consider any non-linearity across the workless distribution by looking at the more extreme cases of workless spells. As discussed in the methodology section, the yearly proportion of time spent workless allows us to analyse trends in age trajectories by considering the average proportion of time spent out of work each year for each cohort at each age.

The BCS data provides an opportunity to explore the impact of local labour market conditions as information is available on the LEA that the 2nd generation lived in at age 16. In addition, unemployment rates at the county level are matched into the BCS data using information from the Employment Gazette. LEA unemployment rates were not available so the use of county level involves aggregating the LEA data up slightly. Unemployment rates for every year from 1986 when the LEA is observed until 1998, the last full year of complete work history data in the BCS, were matched into the data. An implicit assumption when using this data is that individuals stayed in the same county they were observed in at 16 to experience this county level unemployment rate across time. Unfortunately the BCS does not provide any further regional information on the cohort member after age 16 and so this assumption is not testable, even at a more aggregated level.

iii) Descriptive statistics

Starting with the more continuous BHPS data, table 2 shows summary statistics for both generations in the BHPS. Individuals are observed for varying periods of time as the panel is not balanced. On average fathers spend 13% of the total time observed out of work compared to 8% of time for sons. Fathers’ spend an average of 10.8 months out of work compared to 6.8 months for sons. This is perhaps surprising given that the average age of fathers is 41 compared to the sons’ average age of 21. We might expect that sons would have higher levels of worklessness as we observe them earlier in their labour market experience. When considering more discrete measures of worklessness such as the percentage of the sample that spend any time out of work for the period observed, it becomes clear that there is more churning in the sons sample with 54% of sons having experienced a spell out of work compared to only 23% of fathers. The higher proportion of workless spells for fathers must therefore be driven by a smaller fraction of fathers with higher durations out of work. Compared to the

10% of fathers in the sample who have never worked for the period observed, only 1% of sons are never in work for the period observed supporting the evidence from the LFS statistics in the introduction that households with two or more generations that never work are very rare.

If we split the summary statistics for the son by the type of work experience of the father, a story begins to emerge about the scale of the intergenerational correlation of worklessness. Table 3 illustrates that for sons with fathers who are only ever employed, 50% still experience at least one month out of work but on average only 7% of their total time is spent workless in contrast to sons with fathers who never worked who spend around 18% of their time out of work themselves. In months, this equates on average to an extra 5.6 months out of work for sons with fathers' who are never observed to work compared to sons with fathers' that are always observed as employed. The middle group, sons with fathers who experience any worklessness sit in the middle of the other two categories in terms of percentage of time spent out of work with 13%. 16% more sons with fathers with any worklessness spent any time out of work themselves than sons with fathers who were always employed and 15% more sons with fathers who never worked spent a year or more out of work than sons with employed fathers.

Table 4 illustrates the 2nd generation data available in the NCDS and the BCS. Given that the cohort studies are observed across such a long window, two types of measures can be considered, a NEET sample of young adults from 16-23 for comparability with the BHPS data and a wider ranging sample using later information from the cohorts to consider life-cycle bias. As can be seen, individuals in the cohort studies spent less time workless than in the BHPS, with both cohorts experiencing a similar 4% average of time spent workless across the period observed, equivalent to 3.7 and 3.6 months in the NCDS and BCS respectively. Again, as seen in the BHPS, a very small percentage of the 2nd generation in the cohort studies never worked across all of the months observed, consistent with the LFS findings with less stringent restrictions on generations living in the same household. The NCDS looks much more similar to the BHPS in terms of individuals churning in and out of the labour force with around 50% of the sample experiencing at least one month out of work. By contrast for the BCS, only 20% of the sample are out of work at any point across the same age period[4]. The summary statistics for the wider age-range sample are similar to the NEET sample in terms of the average proportion of time spent out of work, however as this is observed over a longer period this is now equivalent to spending 8.1 months workless on average in the NCDS and 7.9 months out of work on average in the BCS. More individuals in both cohorts experience any workless spells, and both experience an increase in individuals experiencing a year or more out of work. As expected with the longer time window, the percentage of the sample never observed as working falls to practically zero.

Table 5 illustrates the sample composition of the two observed employment status variables for the 1st generation. By combining two observations of employment status the aim is to reduce the size of the error term, as discussed in section 3 i), for the two cohort studies. A comparable restricted measure can be created in the BHPS by using only a limited section of the longer window of work history available. As noted this increases the error component from equation (3) as it shortens the window that the 1st generation BHPS cohort are observed for. For the three data sources, there are three states that the father can be observed in at the time the employment status is measured; employed, workless or missing. Given the large amount of missing data in the cohort studies, particularly the BCS at 16, we construct a measure of father's worklessness to be 1 if the father is only ever observed as workless and 0 otherwise. The aim is to create a measure which proxies a lifetime measure of work experience well and limits the impact of measurement error. By restricting this measure to those only ever observed as workless we are making the assumption that for those observed workless in one period and missing in the other, the underlying propensity to experience workless spells is higher than those observed employed in one period and workless in the other and those observed employed in one period and missing in the other.

Given that a more continuous measure is available in the BHPS, this assumption can be tested by summarising the average proportion of time spent out of work in the longer window for three different categories of individuals; father's observed workless at 16 who are missing information at 12, father's observed

[4] There may be an issue with recall bias in the BCS cohort as suggested by this result. Work histories were constructed retrospectively from whenever individuals were interviewed. In the NCDS individuals were interviewed at 23 compared to 26 in the BCS and this could suggest that people in the NCDS remembered more spells out of work than those in the BCS. In a classical measurement error sense, this will not impact the estimation of the intergenerational correlation as attenuation bias only works through the right-hand side variables.

employed at 16 who are missing information at age 12 and father's observed employed in one period and workless in the other. The average proportion of time spent workless for fathers who are missing at 12 and observed employed at 16 is 0.007. For those fathers who are observed in both states, employed and workless at either age the proportion is 0.307. By contrast the proportion of time spent workless for fathers who are missing at 12 and observed workless at 16 is 0.883. This suggests that the underlying propensity to experience spells out of work is much higher for those observed workless in one period and missing in the other than those who are observed employed for one period and either missing or workless in the other.

Table 6 separates the comparable 2nd generation workless experiences from age 16-23 across the three cohorts for the discrete comparable measures of 1st generation worklessness discussed in table 5. For all three types of 1st generation worklessness, the BHPS has higher workless rates than the two cohort studies as seen in tables 2 and 4. Comparing the differences within cohorts across types of worklessness, sons of employed fathers in the NCDS spent on average 4.2% of their time out of work compared to 12.1% of time for sons with fathers defined as workless in our sample, equating to a difference of nearly 7 additional months spent out of work. In the BCS, sons with employed fathers spent a similar amount of time out of work, 3.7%, but sons with workless fathers spent 14.4% of time out of work, 9 months longer than their counterparts with employed fathers. In the BHPS, sons with workless fathers spent 14% more time out of work than sons with employed fathers, on average 12 months in total compared to 4.9 months for those with employed fathers. The middle grouping of sons with fathers in a transitory state of employment and worklessness look more similar to the sons of employed fathers across all three cohorts with an average 2-3% more time spent out of work. In the BCS they look most similar spending just over 1 month more out of work. This increases to 2.4 months more than employed fathers in the BHPS and 2.9 months more in the NCDS.

4. Results

The results in table 6 suggest that there is a sizeable intergenerational correlation in all three data sources observed. Table 7 presents results from univariate OLS regression results from equation (1) using three measures of sons workless experiences and the comparable two-point-in-time measures of 1st generation worklessness available in all three studies discussed in table 5. The intergenerational correlations of worklessness are large and significant across all cohorts when focusing on the proportion of time spent out of work in the 2nd generation. On average a son with a workless father spends 7.9% more time out of work in the NCDS and 10.6% more time out of work in the BCS and BHPS than sons with employed fathers. These are moderate effects with significant economic implications. The scarring literature suggests a wage penalty at 33 of between 15 percent compared to those with no youth unemployment and a future employment scar of a further 3 months by age 33 (Gregg and Tominey, 2005 Gregg, 2001).

Looking across all three measures of 2nd generation worklessness, the coefficients in the BCS, born only a few years before the average year of birth for the BHPS cohort are remarkably similar for the first two measures of sons' workless experiences. This is reassuring and suggests some degree of comparability in the data sources despite the obvious differences including smaller sample sizes and concern over recall bias in the BHPS. For the NCDS cohort, however, as has been seen in the intergenerational mobility literature, there is a lower correlation between the 1st and 2nd generations although the differences here are not statistically significant.

As we move down through the rows of table 7 we move towards a more duration based measure of worklessness in the 2nd generation. Sons with workless fathers are 15-18% more likely to spend a year or more out of work than sons with employed fathers. In the last row, the BCS cohort is the only cohort that has a significant correlation between sons who are never observed in work and workless fathers. For this cohort if the father is observed as workless the son is 3.3% more likely to never be observed in work than his counterpart with an employed father.

i) Measurement error

Utilising the longer window of work history data available in the BHPS to move towards a better measure of the intergenerational correlation, table 8 illustrates the best available estimate of $\hat{\beta}$ in the UK using all observed information for the 1st generation workless experiences in a measure of the proportion of time spent workless. Note that this measure may still suffer from attenuation bias as the 1st generation are not observed across their entire working lives but it will likely have less classical error than table 7. The correlation is 0.117, suggesting that for a standard deviation increase in the time spent out of work in the 1st generation; the son spends an extra 11.7% of time out of work in adulthood. The coefficient when using the more restricted measure of worklessness in the 1st generation, from two time periods of information, is smaller as would be expected if more measurement error was present. The reduction is small however; a little over 1pp reduction in the average time spent out of work if the father is workless rather than employed.

This longer window of work history can be used to estimate ‘w’ and ‘e’ and the distributions of x_i and y_i from section 3i). In the BHPS, 11.6% of fathers observed as employed in the shorter window measure do spend some time out of work when the longer window measure is used. 20.7% of fathers observed as workless in the shorter window spend some time in work in the longer window. Figure 1 plots the distribution of x_i , the proportion of time spent out of work for those observed as employed at a point in time and y_i, the proportion of time spent in work for those observed as workless at a point in time.

As can be seen there is a far greater skew to the right for those observed in work that experience some workless spells in the longer window indicating that many people in this category experience only a relatively small proportion of time out of work. By contrast the distribution of y_i is almost normal suggesting that employment spells for those who are observed as workless are more frequent. Combining these two effects using equation (8) suggests a total average error in the BHPS of 11%, seen in the reduction of $\hat{\beta}$ in table 8. The combination of the scale and distribution by type of employment status suggests that 8.9ppts of the total error is from those observed as workless at a point in time who actually work more than the measure suggests. The remaining 2ppts are from those observed as employed experiencing more workless spells. This suggests that there is more persistence in employment spells than workless spells.

As mentioned in the data section, the BHPS may also be affected by measurement error because the window of time that the 1st generation are observed for varies within the sample. Table A3 replicates table 8 restricting the sample to fathers’ only observed for a minimum of 5 years or more. With this more stringent restriction, the impact on the proportion of time that the son spends workless is very small for both intergenerational correlations. As noted there is a trade-off between observing the 1st generation for a longer period of time and losing valuable sample information but this evidence suggests that the intergenerational correlation is very similar for a sample of individuals observed for a longer periods of time and so the trade-off here is small.

Overall, there is a moderate correlation in intergenerational workless spells across the three cohorts with significant economic implications. Although measurement error may be causing problems in the various data sources it appears unlikely, from this evidence, to be leading to substantial biases in the estimated coefficients when using the more restricted 1st generation measures of worklessness rather than the more complete information available. Our attention now turns to a different form of errors-in-variable bias, life-cycle bias.

ii) Life-cycle bias

As noted in section 3 ii), the trajectories for age-workless trajectories may vary from the age-earnings trajectories observed in the mobility literature. Figure 2 uses quasi-cohorts of males of working age from the LFS for periods from 1992 until 2010 to plot workless rates by age across low educated and high educated groups. The aim of this is to give a sense of what age-workless trajectories we might expect to see and any likely direction of life-cycle bias. Low educated individuals are defined as level 2 (GCSEs) or below and high educated individuals are defined as level 3 (A-levels) and above. As can be seen from the four graphs, there is a fairly consistent stable pattern of a convex and increasing relationship between workless rates and age. The gap between low educated and high educated males is fairly stable across ages until individuals hit age 55 when high educated males catch up with the workless rates of low educated males through early retirement. This suggests that life-cycle bias may not be an issue for intergenerational workless correlations.

To explore this further the work histories in the 2nd generation from the two older cohorts can be used to calculate yearly proportions of time spent out of work. These annual measures can then by plotted by the age of the cohort member to show the age trajectories of the average proportion of time spent out of work by sons with workless fathers and sons with employed fathers. The BHPS is omitted at this stage as this places to many data constraints on the already small samples. Figure 3 illustrates these age trajectories by cohort rather than quasi-cohort. Strikingly, there is not a stable pattern across the two cohorts considered. The intergenerational life-cycle effect, or the changing gap between sons of employed fathers and sons of workless fathers, is relatively stable in the NCDS as seen in the LFS in figure 2, when taking into account any likely local labour market effects that will be discussed in the next section (the NCDS were 23 in 1981). By contrast, in the BCS, the trajectories for sons' with employed fathers remained relatively stable but sons' with workless fathers' consistently experienced higher spells out of work for every year they aged. This is in stark contrast to the mobility literature where the life-cycle bias in terms of the age-profiles in returns to education have been found to be very similar across a number of datasets both within and across countries.

Table 10 shows estimates from the interaction term between sons' age and fathers' workless experience from equation (9). As seen in figures 3 the trend in the NCDS is essentially flat with no significant difference by group. In the BCS, the life-cycle bias increases as son's age. The implications of this finding are that perhaps there is something other than a life-cycle bias driving the results we are seeing. Currently, for the NCDS cohort, it may not be problematic to measure the coefficient at any point in the 2nd generation life-cycle (local labour market conditions aside) as this is relatively stable. By contrast early measures of the intergenerational coefficient in the BCS will understate the lifetime intergenerational coefficient.

This point can be seen in table 11. By expanding out the window that the 2nd generation are observed for to take account of the later data available while still keeping the two cohorts comparable, the intergenerational coefficient can now be measured for individuals' workless experiences up to the age of 29. In the NCDS, for the more continuous measure of the proportion of time spent out of work, as observed in figure 3, the coefficient is very similar to that when only observing the cohort up to age 23. By contrast, in the BCS, increasing the window increases the intergenerational coefficient by 2.5ppts.

Interestingly, the life-cycle effect also seems to vary by the nature of the measure of 2nd generation worklessness used. For the higher duration measures of time spent out of work, the impact of viewing a longer window is much larger than the impact on the more continuous measure. This is because over a longer period individuals have more chance to experience a year or more out of work as seen in table 3. However an increase in the coefficient indicates that those with workless fathers are disproportionately more likely to be affected. In the NCDS, the likelihood of spending a year of more out of work increases by 6ppts when extending the window from 23 to 29 and in the BCS this effect is 8.5ppts larger. This also suggests that there is more churning for younger individuals and as people age there is more persistence in their workless experiences consistent with the scarring literature (Gregg, 2001). When considering the impact of 1st generation worklessness on those who never work in the 2nd generation, there is now a marginally significant effect in the NCDS, albeit of 0.7%, and the effect in the BCS decreases by .7ppts to just over 2%. This suggests that the estimates are becoming more precise with more information available and as the period considered increases by six extra years, the likelihood of never working throughout the whole period decreases.

To summarise although there seems to be little evidence of any life-cycle bias when looking at quasi-cohorts by education across ages in the LFS, there is a difference in the age-workless trajectories in the NCDS and BCS. This suggests that something other than age may be causing this divergence in trajectories across the two cohorts. The next section will look to examine whether local labour market conditions can account for some of these stark differences in life-cycle trajectories.

iii) Local labour market conditions

As discussed in section 3, a potential alternative explanation for the age profiles seen in the previous section could be the impact of external factors such as the local labour market conditions at the time of observation. Considering the contrast in the age-workless profiles in figure 2 by recession and non-recession periods, there is a bigger gap between high and low educated workless rates in the recession period compared to the non-recession periods. Low educated individuals aged 16-50 are out of work on average 22.3% of the time in a recession period compared to 18.5% of the time in a non-recession period. The corresponding figures for high educated individuals are 7.6% and 6.0% respectively. There is also a suggestion that low educated youths are particularly affected in recessions with the gap between low and high educated youths widening out in recession periods. This is consistent with Bound and Freeman (1992) who argue that young men are more sensitive to labour market occurrences than older men as older workers have more experience and seniority to buffer them somewhat from market developments.

Figure 4 considers this issue in the cohort studies by replicating figures 3 but rather than viewing the trends by the age of the son, instead viewing these trends by the year of observation. The annual average unemployment to population ratio is also included to show national trends across the period. In the NCDS, it is clear that the large spike seen in the proportion of time spent workless for sons with workless fathers early in their life-cycle is largely a product of the 1981 recession. While sons with employed fathers also experience an upturn in the proportion of time they spend out of work during this period the shock they experienced is far smaller and they return to a lower level of worklessness much faster. In the BCS, sons with workless fathers experience a shock from the 1991 recession. As in the NCDS, sons with employed fathers also experience a shock at this time but the effect is smaller. Unlike in the NCDS however, the BCS cohort sons with workless fathers do not appear to recover from the shock to employment in the early 1990s and continue to experience greater proportions of time out of work. Interestingly both groups from the older NCDS cohort do not seem to respond to the 1991 recession. Again, this may suggest that younger people are more susceptible to labour market shocks consistent with the evidence from the LFS.

Given this suggestive evidence that labour market conditions may affect the intergenerational correlation, the impact of local labour market conditions can be assessed directly. Table 11 reports the intergenerational coefficients from a range of models discussed in section 3 iii) controlling for local labour market conditions. The first row replicates the first result from column 4 of table 10 for a restricted sample of individuals for whom LEA information and county level unemployment data is available. It can be seen that individuals with this information are slightly more advantaged in terms of the impact of fathers workless spells on their own work experiences compared to the baseline sample used for the rest of the analysis, as the overall intergenerational coefficient falls by 1.2ppts. This is the result that the remainder of the analysis will be compared to for consistency. Running within LEA fixed effects from equation (10) surprisingly does very little to the point estimate of the intergenerational correlation. The correlation falls by .5ppt or only 4% of the total coefficient. This suggests that differences across local areas make little difference to the intergenerational relationship. Improving the information about local labour market conditions by using actual local unemployment rates rather than just fixed effects, as illustrated in equation (11) albeit at the county rather than LEA level, also does very little to the intergenerational correlation, reducing the estimate by only .4ppts. This suggests that very little of the intergenerational correlation in worklessness is due to fathers and sons experiencing the same local area labour market conditions.

The second panel of table 11 presents the results from the interaction model in equation (12). The intergenerational correlation decreases by a further 1ppt from the models with no interaction and the estimated interaction effect is striking. The impact of varying regional unemployment rates across time do not seem to affect the level of the intergenerational coefficient across regions, but instead the effect varies across the two groups; sons with workless fathers and sons with employed fathers. The coefficient is not driven by fathers and sons living in the same areas of high unemployment compared to fathers and sons living in the same areas of low unemployment. Instead, it is a combination of both the levels of unemployment experienced and the fathers' experience of workless spells. The vulnerable group, those with workless fathers, are hit harder by worse local labour market conditions.

Figure 5 illustrates this point, plotting the range of estimates from equation (13) across the various different unemployment rates observed in the data across time and county. The results are striking. Sons with workless fathers in weaker local labour markets with high unemployment spend over 25% more time workless than sons with employed fathers. By contrast, there is no significant difference in the time spent workless in tight local labour markets with low unemployment for sons with workless fathers compared to sons with employed fathers. The gap in the proportion of time spent out of work between those sons with employed fathers and those sons with workless fathers gets larger in weaker local labour markets. This trend is in line with what is observed for sons with workless fathers recovering more slowly from shocks than sons with employed fathers observed in both the NCDS and the BCS cohorts in figure 4.

Overall, there appears to be an important third factor to be considered when estimating the intergenerational correlation in worklessness. If individuals are only observed for a short window, the local labour market conditions at that snap shot of time need to be considered.

5. Conclusion

Despite the major developments in the measurement of intergenerational mobility over the past thirty five years, little work has focused on the intergenerational correlation of those out of work during this period. This group of individuals, of increasing interest in the public domain, are the most vulnerable group not only in terms of the poverty associated with periods out of work but also through later scarring penalties in terms of both wages and future employment and further behavioural related issues such as depression. This research uses the substantial progress made in measuring intergenerational mobility to measure the intergenerational correlation in worklessness for a number of cohorts in the UK.

There is a moderate significant correlation in spells out of work across generations with large economic implications. A son with a workless father is likely to experience between 8-11% more time out of work themselves between 16 and 23. In addition, they are 15-18% more likely to spend a year or more out of work in the same period. This increases to 20-25% when the period that the son is observed for is increased to 16 to 29. However, there is only a small significant effect for one data source when estimating the impact of fathers' worklessness on a son never observed to be working. This is due to the fact that in all data sources, only 1% of sons are never observed to be working. This is in contrast to some discussion that is currently taking place in the public domain. Sons with workless fathers are 3% more likely to never be in work from 16 to 23 than sons with employed fathers in the BCS cohort. There is no significant effect in the NCDS or BHPS. To place the magnitude of this correlation in context, further work is needed to measure this relationship in other countries.

When assessing the impact of measurement error, life-cycle bias and local labour market conditions on the estimated intergenerational correlation the story is mixed. Measurement error appears to have only a limited impact on measures of 1st generation workless spells in the BHPS. This is reassuring given that the available measures in the cohort studies are from only two points in time. Interestingly, unlike in the intergenerational mobility literature, the age trajectories of the workless correlation show no clear pattern across the NCDS and BCS. In the NCDS, the age-profile is flat compared to an increasing age-profile in the BCS. When looking at quasi-cohorts in the LFS there is little evidence of a life-cycle bias when considering differences in workless rates for high and low education groups. This suggests that something else may be driving difference in age-workless trajectories other than age. Controlling for regional variation at a disaggregated level also has surprisingly little impact on the intergenerational correlation. Rather, it seems that having a workless father is more harmful in worse labour market conditions. Evidence from the cohorts and the LFS combined suggests the existence of a local labour market conditions bias.

Careful consideration should therefore be given to local labour market conditions when measuring the intergenerational correlation in worklessness for snap-shots of time. More work is needed to assess the scale and direction of the bias in labour market conditions in other countries to see if this is possible to model.

These findings are the first attempt to quantify the intergenerational correlation in worklessness in the UK. Future work in this area should consider the issues raised in this research when attempting to estimate intergenerational correlations in worklessness. Further analysis is needed to examine the drivers of this intergenerational correlation and to attempt to identify causality in this relationship for policy prescription, something which the intergenerational mobility literature has struggled to achieve. Taking these results together, a picture begins to emerge for the UK where the intergenerational relationship is strong in weaker labour markets with high unemployment with no relationship in tight local labour markets with low unemployment. There are a number of competing hypotheses that may be driving this type of intergenerational relationship including a deprivation story or a welfare dependency story (Wilson, 1997). Further attempts to decipher between these hypotheses as to why such a relationship exists goes beyond the scope of this work. However, these findings alone are an important base to begin to understand the intergenerational correlations in workless spells.

Bibliography

Aigner, D (1973) ‘Regression with a binary independent variable subject to errors in observation’ in *Journal of Econometrics* Vol. 1

Becker, G. and Tomes, N. (1986) ‘Human Capital and the Rise and Fall of Families’ *The Journal of Labor Economics* Vol. 4 (3)

Black, S and Devereux, P. (2011) ‘Recent Developments in Intergenerational Mobility’ in Ashenfelter, O. and Card, D. (eds) *Handbook of Labor Economics, Volume 4B*, Amsterdam: North Holland.

Blanden, J., Gregg, P. and Macmillan, L. (2011) ‘Intergenerational Persistence in Income and Social Class: The Impact of Within-Group Inequality’ CMPO working paper 10/230

Corak, M., Gustafsson, B. and Osterberg, T. (2000) ‘Intergenerational influences on the receipt of Unemployment Insurance in Canada and Sweden’ IZA Discussion Paper Number 184

Dearden, L., Machin, S. and Reed, H. (1997) ‘Intergenerational mobility in Britain’ in *The Economic Journal* Vol. 107(440)

Ekhaugen, T. (2009) ‘Extracting the causal component from the intergenerational correlation in unemployment’ *Journal of Population Economics* Vol. 22

Galindo-Rueda, F. (2002) ‘Progression routes of cohort members: NCDS (1974-2000) and BCS70 (1986-2000). Description of the data work on British Cohort Surveys’. Centre for Economics of Education

Gottschalk, P. (1996) ‘Is the correlation in welfare participation across generations spurious?’ *Journal of Public Economics* Vol. 63

Grawe, N. (2006) ‘Lifecycle bias in estimate of intergenerational earnings persistence’ *Labour Economics* Vol.13

Gregg, P (2001) ‘The impact of youth unemployment on adult unemployment in the NCDS’ *The Economic Journal* Vol. 111 (Features)

Gregg, P and Tominey, E. (2005) ‘The wage scar from male youth unemployment’ *Labour Economics* Vol. 12

Gregg, P. and Wadsworth, J. (2011) ‘The Labour Market in Winter – the 2008-2009 Recession’ in *The Labour Market in Winter; the State of Working Britain.* Gregg, P and Wadsworth, J. (eds.) Oxford University Press: University of Oxford.

Haider, S. and Solon, G. (2006) ‘Life-cycle Variation in the Association between Current and Lifetime Earnings’, *American Economic Review* Vol.96.

Halpin, B. (1997) .Unified BHPS work-life histories: combining multiple sources into a user-friendly format., Technical Papers of the ESRC Research Centre on Micro-Social Change, 13, Colchester: University of Essex

Johnson, P. & Reed, H. (1996) 'Intergenerational Mobility Among the Rich and Poor: Results from the National Child Development Survey' *Oxford Review of Economic Policy*, Vol. 12 (1) p.127-143

Lee, C. and Solon, G. (2009) 'Trends in Intergenerational Income Mobility' in *The Review of Economic and Statistics* Vol.91(4)

Levine, P. and Zimmerman, D. (1996), "The Intergenerational Correlation in AFDC Participation: Welfare Trap or Poverty Trap?" IRP Discussion Paper No. 1100- 96.

Mazumder, B (2005) "Fortunate sons: New estimates of intergenerational mobility in the U.S. using social security earnings data' *Review of Economics and Statistics* Vol. 87

O'Neill, D. and Sweetman, O. (1998) 'Intergenerational mobility in Britain: Evidence from Unemployment Patterns' *Oxford Bulletin of Economics and Statistics* Vol. 60 (4)

Page, M. (2004). 'New Evidence on Intergenerational Correlations in Welfare Participation' in Generational Income Mobility in North America and Europe, ed. by M. Corak, Cambridge University Press

Romer, D. (2006) 'Chapter 4 - Real Business Cycles' in *Advanced Macroeconomics, 3rd edition* London: McGraw-Hill

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

Solon, G. (1999), 'Intergenerational mobility in the labour market' in Ashenfelter, O. and Card, D. (eds) *Handbook of Labor Economics*, *Volume 3A*, Amsterdam: North Holland.

Stadler, C. (1994) 'Real Business Cycles' *Journal of Economic Literature* Vol 33

Wilson, WJ. (1997) '*When Work Disappears: The World of the New Urban Poor'* (Random House: New York)

Zimmerman David (1992), 'Regression Toward Mediocrity in Economic Stature.' *American Economic Review*, 82, 409-42

Table 1: Population estimates from the April-June 2010 Labour Force Survey of the number of workless households in the UK

	ONS figures (students workless)	Students not workless
Total number of households of working age	20,818,429	20,818,429
Workless households of working age	3,876,892	3,659,907
Total number of households with 2 or more generations	4,199,974	4,199,974
Households with 2 generations where both are workless	358,769	178,742
Households with 2 generations where both are workless >1yr	295,085	141,147
Households with 2 generations where both are workless >2yr	243,419	109,304
Households with 2 generations where both are workless >5yr	184,252	80,084
Households with 2 generations where both never worked	38,481	15,350
Panel B: The length of time 2nd generation have been out of full time education for	**ONS figures (students workless)**	**Students not workless**
Total number of households with 2 generations where both never worked	38,481	15,350
2nd generation out of education <1 year	25,717	5,387
2nd generation out of education 1-2 years	1,485	992
2nd generation out of education 2-5 years	1,361	1,361
2nd generation out of education 5 or more years	5,625	5,625
2nd generation no information on leaving full time education	4,293	1,985

Not including student only households. Whilst the ONS count full time education as workless it is preferred here to not include full time students in workless numbers

Table 2: Descriptive statistics from the BHPS intergenerational work history data

	1st generation (father)	2nd generation (son)
Average proportion of time workless	0.134	0.087
Average number of months workless	10.78	6.80
Average total months observed	103.43	90.33
Range of total months observed	24-276	24-244
Percentage any month worklessness	22.63	53.63
Percentage a year or more workless	16.70	19.12
Percentage never worked	10.11	1.10
Average Age	41.24	21.14
Range of average age observed	28-56	17-32
N	455	455

Table 3: Descriptive statistics from the BHPS 2nd generation (son) work history data by the workless experiences of the 1st generation (father)

	1st generation (father) work experiences		
2nd generation (son) descriptives	Never workless	Any workless	Never worked
Average proportion of time workless	0.071	0.139	0.183
Average number of months workless	6.1	9.1	11.7
Average total months observed	91.61	85.93	92.53
Range of total months observed	24-204	24-244	24-244
Percentage any month workless	50.00	66.02	63.04
Percentage a year or more workless	17.61	24.27	32.61
Percentage never worked	0.85	1.92	2.17
Average Age	21.30	20.60	21.16
Range of average age observed	17-30	18-32	18-32
N	352	103	46

Table 4: Descriptive statistics from the NCDS and BCS 2nd generation (son) work history data

	NCDS	BCS
16-23 (NEET)		
Average proportion of time workless	0.044	0.042
Average number of months workless	3.74	3.57
Total months observed	84	84
Percentage any month workless	49.06	19.72
Percentage a year or more workless	5.93	8.27
Percentage never worked	1.60	1.74
16 - 29		
Average proportion of time workless	0.052	0.050
Average number of months workless	8.13	7.85
Total months observed	156	156
Percentage any month workless	56.61	27.72
Percentage a year or more workless	13.59	13.88
Percentage never worked	0.17	0.56
N	4635	4646

Table 5: Creating comparable 1st generation (father) workless measures from the NCDS, BCS and BHPS work history data

NCDS (1958)

Age of son 16 / 11	Father employed	Father workless	Missing	**TOTAL**
Father employed	2931 (63.24)	123 (2.65)	1031 (22.24)	**4085 (88.13)**
Father workless	38 (0.82)	41 (0.88)	35 (0.76)	**114 (2.46)**
Missing	410 (8.85)	26 (0.56)	0 (0.00)	**436 (9.41)**
TOTAL	**3379 (72.90)**	**190 (4.10)**	**1066 (23.00)**	**4635**

Dark shaded region represents those counted as workless, light shaded region corresponds to group in table 6 who are counted as not workless for the remainder of the analysis

BCS (1970)

Age of son 16 / 10	Father employed	Father workless	Missing	**TOTAL**
Father employed	1855 (39.93)	213 (4.58)	2100 (45.20)	**4168 (89.71)**
Father workless	27 (0.58)	39 (0.84)	213 (4.58)	**189 (4.07)**
Missing	245 (5.27)	44 (0.95)	0 (0.00)	**289 (6.22)**
TOTAL	**2127 (45.78)**	**296 (6.37)**	**2223 (47.85)**	**4646**

Dark shaded region represents those counted as workless, light shaded region corresponds to group in table 6 who are counted as not workless for the remainder of the analysis

BHPS (1977)

Age of son 16 / 10	Father employed	Father workless	Missing	**TOTAL**
Father employed	268 (58.90)	17 (3.74)	1 (0.22)	**286 (62.86)**
Father workless	5 (1.10)	32 (7.03)	0 (0.00)	**37 (8.13)**
Missing	105 (23.08)	26 (5.71)	1 (0.22)	**132 (29.01)**
TOTAL	**378 (83.08)**	**75 (16.48)**	**2 (0.44)**	**455**

Dark shaded region represents those counted as workless, light shaded region corresponds to group in table 6 who are counted as not workless for the remainder of the analysis

(light shading)	Defined not workless
(dark shading)	Defined workless

Table 6: Average proportion of time the 2nd generation (son) spent workless by the comparable measures of workless experiences of the 1st generation (father)

	NCDS	BCS	BHPS
1st generation only employed	0.042 (3.5 mth)	0.037 (3.1 mth)	0.097 (4.9 mth)
1st generation observed employed and workless	0.071 (5.9 mth)	0.050 (4.2 mth)	0.121 (7.1 mth)
1st generation only workless	0.121 (10.2 mth)	0.144 (12.0 mth)	0.237 (12.0 mth)

Shading corresponds to groupings in table 5

Table 7: Intergenerational worklessness correlations for comparable 1st generation measures and varying measures of worklessness in the 2nd generation for the cohort studies and the BHPS

1st generation measure	**NCDS** Observed workless at 11/16	**BCS** Observed workless at 10/16	**BHPS** Observed workless at 12/16
2nd generation measure (NEET) 16-23			
Proportion of time out of work	0.0786 (.010)***	0.1060 (.010)***	0.1057 (.022)***
A year or more workless	0.1498 (.024)***	0.1726 (.019)***	0.1756 (.054)***
Never working	0.0037 (.013)	0.0325 (.009)***	0.0071 (.015)
N	4635	4646	454

Standard errors in parenthesis. * 90% confidence, ** 95% confidence, *** 99% confidence.

Table 8: Assessing the impact of measurement error in the BHPS by comparing measures of worklessness for fathers' using a longer and shorter time window

1st generation measure	Proportion of time out of work	Only ever observed workless at 12/16
2nd generation measure (NEET) 16-23		
Proportion of time out of work	0.1173 (.023)***	0.1057 (.022)***
N	455	454
'w' - Proportion observed employed with some workless spells in longer window		0.116
'e' – Proportion observed workless with some employment in the longer window		0.207
$\bar{x}_i$ – Average proportion of time spent workless for those observed employed in short window		0.181
$\bar{y}_i$ – Average proportion of time spent employed for those observed as workless in short window		0.431

Standard errors in parenthesis. * 90% confidence, ** 95% confidence, *** 99% confidence. Sample smaller in only observed workless as one observation missing information at both 12 and 16. See table 5.

Table 9: Exploring the life-cycle bias in the cohort studies interacting the age of the son with fathers' workless experience

1st generation measure	Observed workless at 11/16	
2nd generation measure	Intergenerational correlation ($\hat{\beta}$)	Interaction ($\hat{\theta}$)
NCDS 16 to 42	0.0612 (.019)***	-0.0010 (.013)
BCS 16 to 29	0.0992 (.018)***	0.0084 (.002)***

Standard errors in parenthesis. * 90% confidence, ** 95% confidence, *** 99% confidence. N=4635 in NCDS, 4646 in BCS,

Table 10: Intergenerational worklessness correlations for the cohort studies for a longer time window in the 2nd generation

1st generation measure	**NCDS** Observed workless at 11/16		**BCS** Observed workless at 10/16	
2nd generation age	16 - 23	16 - 29	16 - 23	16 - 29
Proportion of time out of work	0.0734 (.010)***	0.0699 (.012)***	0.1006 (.010)***	0.1244 (.010)***
A year or more workless	0.1398 (.024)***	0.1985 (.034)***	0.1662 (.020)***	0.2511 (.025)***
Never working	0.0010 (.013)	0.0072 (.004)*	0.0298 (.009)***	0.0226 (.005)***
N	4635	4635	4646	4646

Standard errors in parenthesis. * 90% confidence, ** 95% confidence, *** 99% confidence..
Note the 16-23 numbers vary from table 7 as parental age controls are now included in the regression

Table 11: Intergenerational worklessness correlations for the proportion of time spent workless in the 2nd generation on a discrete measure of worklessness for the 1st generation, controlling for various local labour market conditions in the BCS

1st generation measure	Observed workless at 10/16	
Model type	Intergenerational correlation ($\hat{\beta}$)	Interaction ($\hat{\theta}$)
Full model	0.1244 (.010)***	
Restricted BCS sample (observed LEA and county level data)	0.1123 (.020)***	
Within LEA (δ_r)	0.1074 (.020)***	
Controlling for county level unemployment (u_{irt})	0.1087 (.020)***	
Interaction model (Controlling for county level unemployment (u_{irt}))	0.0974 (.018)***	0.0131 (.005)***
N	3672	

Standard errors in parenthesis. * 90% confidence, ** 95% confidence, *** 99% confidence..

Figure 1: Distribution of the errors from a short window point in time observation compared to a long window continuous proportion of time spent workless for fathers in the BHPS

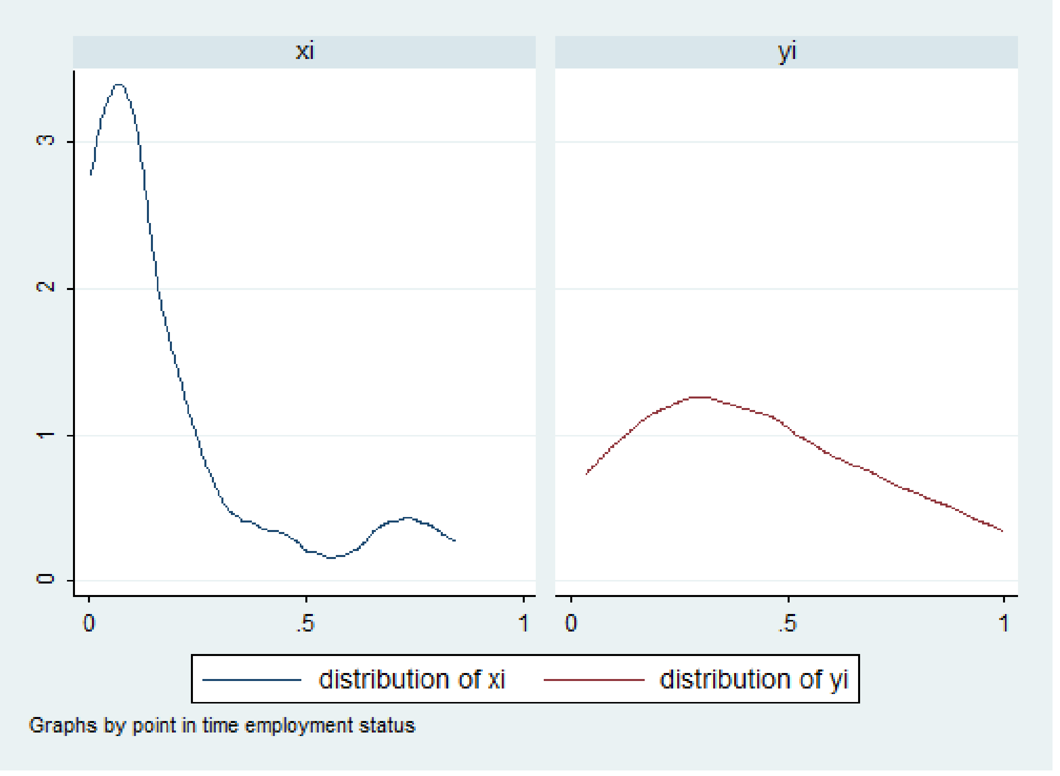


Figure 2: Age profiles of the average proportion of time individuals spend out of work and education by education level

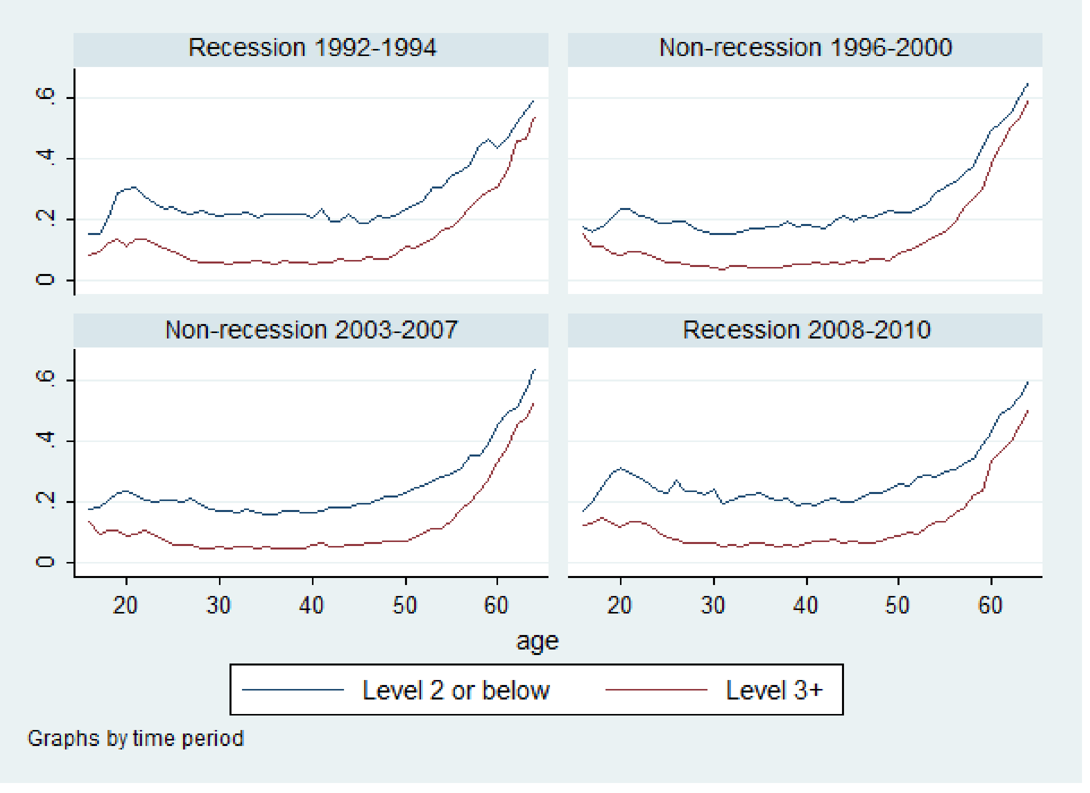


Figure 3: Age profiles of the average proportion of time the 2nd generation spent out of work for sons with workless fathers and sons with employed fathers

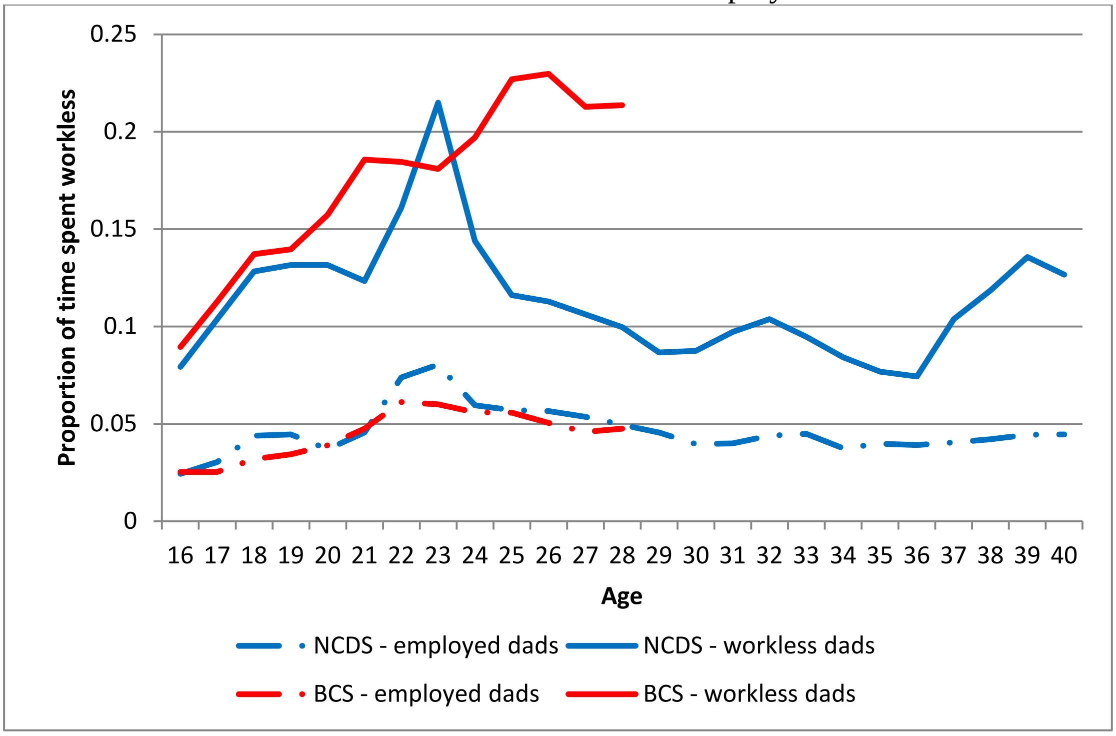


Figure 4: The average proportion of time the 2nd generation spent out of work for sons of workless fathers and sons of employed fathers by the year of observation

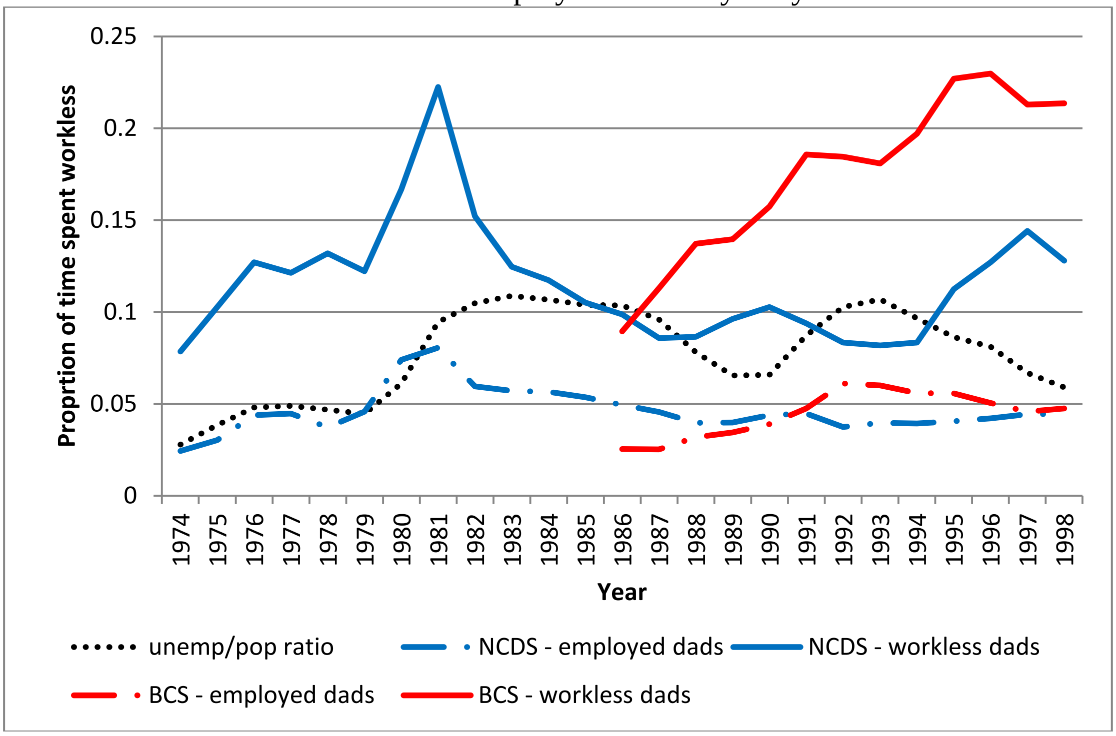
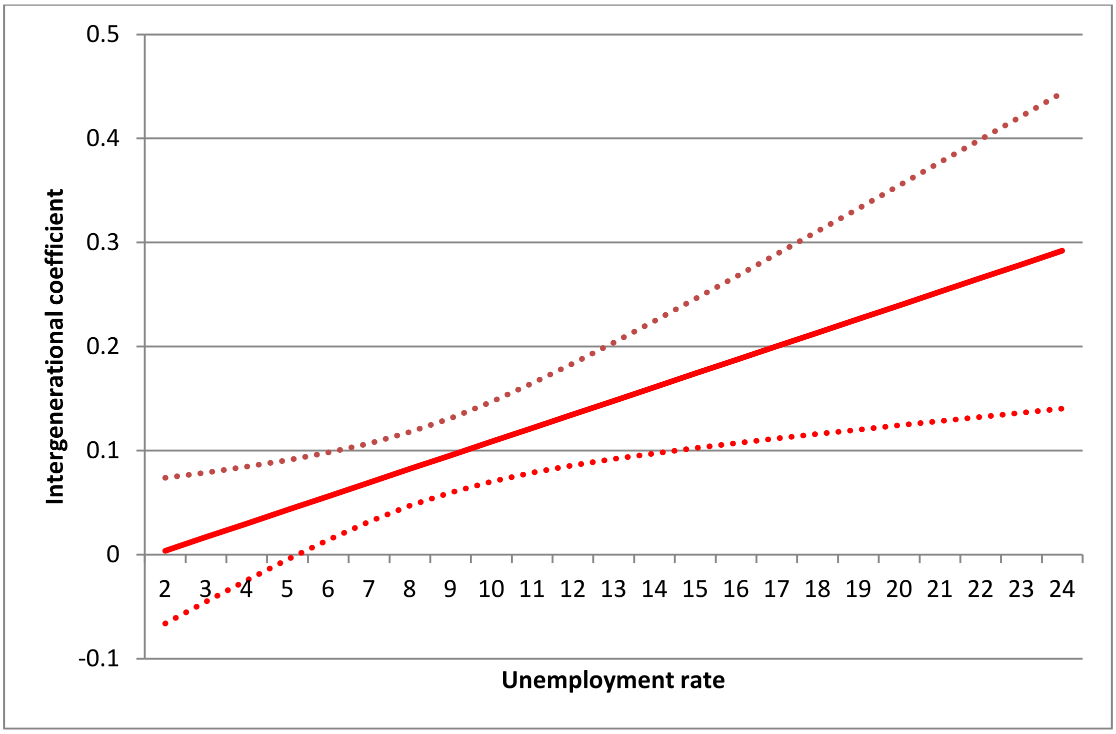


Figure 5: Variation in the intergenerational correlation of worklessness in the BCS by the county level unemployment rate based on the families' county of residence in 1986



Appendix

Table A1 Unemployment rates for the 1st generation cohort data

	Cohorts – all	**Cohorts – sample**	**National unemp. Rate**
	Father-son pairs		**Males 16+**
NCDS			
1969	3.94	2.74	2.8*
1974	6.48	5.35	3.0
BCS			
1980	5.85	4.34	6.6
1986	14.21	12.22	11.7

Table A2: Background characteristics for sons with and without available work history data in the cohort studies

	NCDS		**BCS**	
	Work history	**No work history**	**Work history**	**No work history**
Father's education				
Before school leaving age	2.01	3.91	0.55	0.87
School leaving age	57.25	57.48	64.67	69.36
O-levels	17.48	19.99	14.14	13.20
A-levels	1458	11.72	11.30	9.30
Higher education	8.67	6.90	9.34	7.27
Mother's education				
Before school leaving age	1.66	3.71	0.87	1.16
School leaving age	46.79	46.55	63.74	68.06
O-levels	29.29	30.55	16.74	16.03
A-levels	16.54	14.46	12.31	9.93
Higher education	5.72	4.72	6.33	4.82
Father's social class				
I	4.99	4.04	5.82	4.35
II	14.19	11.43	12.49	10.64
III nm (BCS only)			13.58	10.20
III (m in BCS)	61.25	60.28	47.71	48.91
IV	11.80	11.46	14.51	17.38
V	7.77	12.79	5.89	8.51
Sons characteristics				
IQ test score	100.58	97.32	101.01	98.49

Father's education observed for 63% of total sample in NCDS and 94% in BCS, Mother's education observed for 65% of total sample in NCDS and 98% in BC. Father's class observed for 93% of total sample in NCDS and 91% in BCS, IQ observed for 80% of total sample in NCDS and 67% in BCS http://www.cls.ioe.ac.uk/studies.asp?section=000100020003

Table A3: A robustness test on the sample restriction of the minimum length of the window that the 1st generation are observed for in the BHPS (replicating table 8)

1st generation measure **Observed for at least 5 years**	Proportion of time out of work (continuous)	Only ever observed workless at 12/16 (discrete)
2nd generation measure (NEET) 16-23		
Proportion of time out of work (continuous)	0.1189 (.030)***	0.1053 (.030)***
N	333	332

Standard errors in parenthesis. * 90% confidence, ** 95% confidence, *** 99% confidence. Sample smaller in only observed workless measure as one observation missing information at both 12 and 16.