

Education: Risk Enhancing or Insurance Mechanism?

Judith Maria Delaney

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
of
University College London.

Department of Economics
University College London

May 31, 2017

I, Judith Maria Delaney, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Abstract

In the first chapter, I examine the returns to education for both males and females with a particular focus on the effect of wage risk and periods of non-employment. I also account for selection in to the labour market using a Heckman selection equation and decompose earnings in to permanent and transitory components in an effort to understand the components of wage risk. My results suggest that failure to account for periods of non-employment, wage risk and selection in to the labour market when calculating returns to education leads to biased estimates.

In the second chapter, along with my co-author, Paul Devereux, we look at the causal effect of education on earnings uncertainty and volatility and the effect of education on sheltering workers from the adverse effects of recessions. We use the 1973 change in compulsory schooling law to provide exogenous variation in education. Our regression discontinuity estimates suggest that men whose education was increased by the law subsequently had lower earnings volatility, less pro-cyclical earnings, and were less likely to experience real pay cuts.

In the third chapter, I analyse the role of risk, family background, cognitive and non-cognitive skills in determining college attendance. I use a structural life cycle model explicitly capturing the decision to go to college and incorporating important features which impact the returns to college such as savings, labour supply, human capital accumulation and depreciation, wage risk and employment risk. It is estimated that grants, parental background, non-cognitive skills and risk significantly impact the decision to go to college. However, the biggest factor in determining college attendance is cognitive skills. This is driven both by differences in returns to college conditional on cognitive skills and by the larger psychic costs faced by those with low cognitive skills.

Acknowledgements

I am deeply indebted to my supervisors Sir Richard Blundell and Eric French. They have been very supportive, kind and encouraging throughout the entire PhD. I am very grateful and feel extremely privileged to have been able to discuss my research with such wonderful minds. It was a pleasure to have had such kind supervisors and while my thesis has benefited greatly from such meetings, I know that their influence will last long after the PhD.

I would also like to thank Fabien Postel-Vinay for some insightful discussions and also other UCL faculty including Pedro Caneiro, MariaChristina De Nardi, Christian Dustmann, Suphanit Piyapromdee, Imran Rasul and Michela Tincani for their help. I am also grateful to the administrative staff at UCL for their help along the way.

A special thanks to Paul Devereux who inspired me to do a PhD. Ever since I was an undergrad, Paul has been supportive, friendly and very helpful. Chapter 3 in this thesis is joint work with Paul and I am so lucky to have had the opportunity to write a chapter with such a great person and scholar.

Finally, I would like to thank my family and friends. I don't think I could have done this PhD without your support. Thank you all so much for your encouragement and for all the fun we had which definitely made the PhD a lot easier, I am very grateful.

Thanks so much to my mam, Frances, whose endless support, encouragement and sense of humour have been so important. Thanks for listening to me going on about the PhD and I hope that knowing the difference between structural and reduced form economics comes in useful at some stage! This PhD is as much yours as it is mine. Thanks for everything!

Contents

| | | |
|----------|---------------------------------------------------|-----------|
| 1 | Introduction | 10 |
| 2 | Risk-Adjusted Returns | 13 |
| 2.1 | Introduction | 13 |
| 2.2 | Methodology | 16 |
| 2.3 | Earnings Process | 18 |
| 2.4 | Selection Bias | 20 |
| 2.5 | Endogenous Labour Market Participation | 21 |
| 2.6 | Data | 24 |
| 2.7 | Risk Differentials | 25 |
| 2.8 | Results | 27 |
| 2.8.1 | <i>Males</i> | 27 |
| 2.8.2 | <i>Females</i> | 30 |
| 2.9 | Advanced Information and Insurance | 31 |
| 2.10 | Conclusion | 33 |
| 2.11 | References | 35 |
| 2.12 | Appendix | 39 |
| 3 | Education, Uncertainty and Business Cycles | 51 |
| 3.1 | Introduction | 51 |
| 3.2 | Data | 53 |
| 3.3 | Earnings Volatility Measures | 56 |
| 3.4 | Estimation Strategy | 60 |
| 3.5 | Results | 62 |
| 3.5.1 | Log Weekly Pay | 62 |

| | | |
|----------|---------------------------------------------------------------------------------|-----------|
| 3.5.2 | Standard Deviation of Earnings | 64 |
| 3.5.3 | Earnings Cyclicalty | 68 |
| 3.5.4 | Pay cuts | 69 |
| 3.6 | Robustness Checks | 71 |
| 3.6.1 | Bandwidth and Model Specification | 71 |
| 3.6.2 | Excluding Scotland | 72 |
| 3.6.3 | Standard Errors Clustered at the Individual Level | 72 |
| 3.7 | Conclusion | 73 |
| 3.8 | References | 74 |
| 3.9 | Appendix | 79 |
| 4 | To Go to College or Not? The Role of Ability, Family Background and Risk | 91 |
| 4.1 | Data | 96 |
| 4.2 | Reduced Form Analysis | 98 |
| 4.3 | Model | 102 |
| 4.3.1 | Set-up | 102 |
| 4.3.2 | Utility | 103 |
| 4.3.3 | Budget Constraint | 103 |
| 4.3.4 | Earnings Process | 104 |
| 4.3.5 | Value Functions | 105 |
| 4.3.6 | Education Decision | 105 |
| 4.4 | Estimation | 106 |
| 4.4.1 | Calibration | 107 |
| 4.4.2 | Parameter Estimates | 108 |
| 4.4.3 | Compensating Variation | 110 |
| 4.4.4 | Heterogeneity in College Attendance | 111 |
| 4.5 | Policy Experiments | 111 |
| 4.6 | Conclusion | 114 |

List of Figures

| | | |
|------|---------------------------------------------------------------------------------------------------------|-----|
| 3.1 | Regional Unemployment Rates | 59 |
| 3.2 | First stage effect using BHPS | 79 |
| 3.3 | Young | 80 |
| 3.4 | Older | 80 |
| 3.5 | All | 80 |
| 3.6 | Young | 81 |
| 3.7 | Older | 81 |
| 3.8 | All | 81 |
| 3.9 | Young | 82 |
| 3.10 | Older | 82 |
| 3.11 | All | 82 |
| 3.12 | Young | 83 |
| 3.13 | Older | 83 |
| 3.14 | All | 83 |
| 3.15 | Young | 84 |
| 3.16 | Older | 84 |
| 3.17 | All | 84 |
| 4.1 | Cognitive ability, non-cognitive ability and parental interest by parental education | 99 |
| 4.2 | Non-cognitive ability and parental interest by cognitive ability | 99 |
| 4.3 | Parental interest by non-cognitive ability | 100 |
| 4.4 | College by cognitive ability, non-cognitive ability, parental education and parental interest | 100 |

List of Tables

| | | |
|------|---------------------------------------------------------------------------------------------------------------------------------------------|----|
| 2.1 | Males: P-values for test of null hypotheses of no correlation between consumption growth and future income growth for all years | 32 |
| 2.2 | Females: P-values for test of null hypotheses of no correlation between consumption growth and future income growth for all years | 32 |
| 2.3 | Males: OLS Regression of Log Annual Net Earnings | 39 |
| 2.4 | Males: First Stage Estimates Probit of Employment using Non Labour Income as an instrument | 40 |
| 2.5 | Males: Including Inverse Mills Ratio in OLS Regression of Log Annual Net Earnings | 41 |
| 2.6 | Males: Probit Regression of Employment | 42 |
| 2.7 | Females: OLS Regression of Log Annual Net Earnings | 43 |
| 2.8 | Females: First Stage Estimates Probit of Employment using Non Labour Income as an instrument | 44 |
| 2.9 | Females: Including Inverse Mills Ratio in OLS Regression of Log Annual Net Earnings | 45 |
| 2.10 | Females: Probit regression of Employment | 46 |
| 2.11 | IRR | 47 |
| 2.12 | IRR with CRRA equal 1.5 | 47 |
| 2.13 | IRR with CRRA equal 3 | 47 |
| 2.14 | IRR Corrected for Endogenous Employment | 48 |
| 2.15 | IRR Corrected for Endogenous Employment CRRA equal 1.5 | 48 |
| 2.16 | IRR Corrected for Endogenous Employment CRRA equal 3 | 48 |
| 2.17 | IRR Females | 49 |
| 2.18 | IRR with CRRA equal 1.5 | 49 |
| 2.19 | IRR with CRRA equal 3 | 49 |

| | | |
|------|-------------------------------------------------------------------------------------------------|-----|
| 2.20 | IRR Females Corrected for Endogenous Employment | 50 |
| 2.21 | IRR Females Corrected for Endogenous Employment CRRA equal 1.5 | 50 |
| 2.22 | IRR Corrected for Endogenous Employment CRRA equal 3 | 50 |
| 3.1 | Descriptive Statistics for Males NESPD 1983-2013 | 55 |
| 3.2 | Effect of the Law on Log Weekly Pay | 64 |
| 3.3 | Effect of the Law on the Standard Deviation of Log Earnings | 66 |
| 3.4 | Effect of the Law on the Standard Deviation of Earnings | 67 |
| 3.5 | Effect of the Law on Earnings Cyclicalilty | 68 |
| 3.6 | Effect of the Law on Regional Earnings Cyclicalilty | 69 |
| 3.7 | Effect of the Law on the Probability of a Real Weekly Pay Cut (over 5-Year Period) | 71 |
| 3.8 | Effect of the Law on Log Weekly Pay | 85 |
| 3.9 | Effect of the Law on the Standard Deviation of Weekly Pay | 85 |
| 3.10 | Effect of the Law on Weekly Pay Cyclicalilty | 86 |
| 3.11 | Effect of the Law on Regional Weekly Pay Cyclicalilty | 86 |
| 3.12 | Effect of the Law on Probability of Real Weekly Pay Cut over 5-Year Period | 87 |
| 3.13 | Effect of the Law after Omitting Scottish Residents | 88 |
| 3.14 | Effect of the Law on Log Weekly Pay | 89 |
| 3.15 | Effect of the Law on the Standard Deviation of Log Weekly Pay | 89 |
| 3.16 | Effect of the Law on the Probability of a Real Weekly Pay Cut (over 5-Year Period) | 90 |
| 4.1 | Regression of Log Earnings at Ages 33 and 42 | 102 |
| 4.2 | Model Estimates with standard errors in parenthesis | 108 |
| 4.3 | Psychic Cost Equation | 108 |
| 4.4 | NCDS 1958 Cohort Compensating Variation | 110 |
| 4.5 | NCDS 1958 Cohort College Attendance | 111 |
| 4.6 | NCDS 1958 Cohort Counterfactuals | 112 |
| 4.7 | NCDS 1958 Cohort Counterfactuals | 113 |
| A1 | Comparing Simulated and Data Moments - NCDS 1958 | 120 |

Chapter 1

Introduction

There is much evidence to suggest that having more education leads to positive labour market returns in the form of higher earnings. However, there is little evidence about whether education reduces earnings risk or how risk impacts the decision to go to college. This thesis takes a rigorous account of the role of risk in education using both rich UK household data and large administrative panel data.

To start with I look at how standard returns to education change once we allow for risk to impact returns. The literature on returns to education has been focused on obtaining the causal return to education absent of any endogeneity. While obtaining the correct causal estimate is important, this preoccupation has led the literature to omit other equally important facets in returns to education. Investing in education is risky as there is no guarantee of any future income stream and moreover no insurance market to insure against low returns. Therefore, the return to education should be augmented to allow for the fact that individuals do not like risk and so would prefer a more stable earnings stream even if the average return is smaller. It might be difficult to directly say how wage risk may impact returns but it is plausible to imagine that accounting for employment risk will tend to increase the returns to education since those with higher education have higher employment rates and so are in receipt of income for a larger fraction of their lives. Ignoring employment differentials across education levels implicitly assumes that unemployment rates are education invariant. In chapter two, I examine the returns to education for both males and females focusing on the effect of both wage risk and employment risk. I also account for selection in to the labour market using a Heckman selection equation to account for the fact that those who are in work may be different than those who are not in work which may be more important for females. Understanding the sources of wage risk is important from a policy perspective,

the welfare effects of transitory shocks may not be so large if such shocks can be easily smoothed from savings or the tax and transfer system. On the other hand permanent shocks which are long lasting will have large effects on welfare. Therefore I decompose wages in to both transitory and permanent shocks and only allow permanent shocks to impact returns. I find unadjusted returns to education similar to previous estimates in the range of 6 to 7 percent. Accounting for employment risk leads to quite substantial changes in returns particularly for high school graduates while the wage risk adjusted returns are much smaller. Correcting for non random selection in to the labour market results in small changes in returns for males but quite significant changes for females. My results suggest that failure to account for employment risk, wage risk and selection in to the labour market when calculating returns to education leads to biased estimates.

In chapter 3, along with my co-author, Paul Devereux, we look at the direct causal effect of education on labour market risk. More specifically we examine the effect of an extra year of schooling on wage volatility, wage cyclicality and the probability of getting a wage cut. We allow the effects to vary over the life cycle. There is little causal evidence about whether education reduces earnings uncertainty and volatility or helps to shelter workers from the adverse effects of recessions. To address endogeneity of schooling we use the raising of the school leaving age from 15 to 16 in the UK in 1973. Combined with the New Earnings Survey Panel Dataset (NESPD) which is a large administrative dataset that follows 1 percent UK workers this allows us to use a regression discontinuity design to examine the causal impact of an extra year of education on each of our risk outcomes. Our estimates suggest that men whose education was increased by the law subsequently had lower earnings volatility, less pro-cyclical earnings, and were less likely to experience real pay cuts. In general, these effects were larger for men aged less than 40. Our results provide another avenue through which education may lead to an increase in welfare.

To take the thesis full circle, I look at why people choose to go to college in the first place and the importance of risk, ability and family background. I use a structural life cycle model to address these questions. While returns to college are estimated to be somewhere between 30 and 60 percent, still less than half of people in the UK go to college. If individuals are fully optimising then it must be that they take non-pecuniary aspects of college in to consideration. To account for this, I allow psychic costs to impact the decision to go to college to capture the fact that those with lower cognitive skills may lack the ability to do

the coursework while those with lower non-cognitive skills may not have the determination, self application or perseverance to finish the course. In addition, family background will play a vital role by providing children with an environment more conducive to skill acquisition and a culture which makes college going a given. I estimate a rich structural model which explicitly captures the decision to go to college and incorporates important features which impact the returns to college such as savings, labour supply, human capital accumulation and depreciation, wage risk and employment risk. The model is estimated using rich UK cohort data containing early measures of both cognitive and non-cognitive ability, in addition to rich family background information. It is estimated that grants, parental background, non-cognitive ability and risk significantly impact the decision to go to college. However, the biggest factor in determining college attendance is cognitive ability; if high school graduates had the same cognitive ability distribution as college graduates then college attendance would increase by roughly 20%. This is driven both by differences in returns to college conditional on ability and by the larger psychic costs faced by those with low ability. This suggests that policies aimed at increasing the cognitive ability of individuals at early ages are key to increasing college attendance.

Chapter 2

Risk-Adjusted Returns

2.1 Introduction

The returns to education is one of the most widely researched topics in economics. Hundreds of papers have been written using a variety of different techniques and using many different countries but there is still no general consensus on what the returns to education actually are. Early studies followed the Mincer set up and regressed log earnings on years of schooling and a quadratic in experience and took the coefficient on years of schooling to be the return (Mincer, (1974)). Criticism of this approach for failure to control for the endogeneity of schooling led to a plethora of papers making use of instrumental variables techniques which attempt to exogenously isolate the causal effect of schooling on earnings (see Card 1995 for a review). However, despite the huge number of papers examining the returns to education, there is a dearth of papers which incorporate risk. Since almost all investment decisions, be they property investment decisions, investment in stocks or setting up a new business involve some consideration of the amount of risk involved and the risk return trade off, it is surprising that such an important factor is almost completely omitted in the literature. Human capital like any other investment is subject to the perils of risk and so it is imperative to include this when analysing the return. An individual deciding whether to invest in education faces a huge amount of uncertainty concerning future labour market conditions, completion of the schooling level, length of life, future earnings and what fraction of time will be spent in employment. In this paper I focus on the effect of unemployment risk and wage risk. Given that there does not exist any market that insures against low returns to education this is a very important topic and may explain part of the reason why many students do not progress to further education despite the perceived benefits. Investing in education does not guarantee that one will obtain a high paying job; indeed due

to the skewed nature of the earnings distribution many people will earn significantly less than mean earnings. Income levels among observationally similar people may differ due to luck, social connectedness, illness, promotions, ability, different training opportunities or motivation and therefore there is a wide range of potential earnings outcomes that may be realised. Thus it is likely that individuals care not only about the mean of the earnings distribution but also the variance.¹ The variance is a measure of risk and if an individual is mildly risk averse they will prefer the earnings distribution with the lower variance conditional on the mean being the same. The returns to education will differ if one attempts to adjust for the riskiness of different education levels although the effects of employment risk and wage risk are not clear cut. If higher educated individuals have lower risk of unemployment then their lifetime expected earnings will be higher and this will increase the returns to education. On the other hand if increased education comes at the cost of higher wage risk then individuals will place less value on higher education and the returns will be lower. Higher educated males have higher employment rates due to higher levels of human capital, they are less likely to be fired due to higher training costs and in addition they can downgrade in times of economic downturn, all of which contributes to a lower risk of unemployment. For females it is not so clear how the returns to schooling may be biased. Higher educated females have a higher probability of finding a job but for many reasons may decide not to participate in the labour force. The opportunity cost of being out of the labour market is higher for more educated than for less educated females and therefore it is more costly for higher educated females to not participate in the labour market; however if there is assortative mating, higher educated females may choose to opt out of the labor market due to higher household earnings. Lower educated females on the other hand have lower income and therefore due to financial reasons would like to participate in the labour market but as it is harder for lower educated women to find a job they may be discouraged from looking. Despite the fact that the low educated have higher unemployment probabilities, unemployment benefits represent a large chunk of their pre-displacement income and this may be enough to negate the adverse effects of employment. Added to this, the existence of minimum wage laws provide a lower bound to the wage that one can receive and so this may make the risk adjusted returns smaller than one may have previously thought.

As alluded to at the beginning, the literature is scant with research on this topic. Weiss

¹If returns are normally distributed this is enough to summarise the entire distribution.

(1972) looks at the risk in occupation and education using data on scientists in the US. In his model an individual seeks to maximize the expected discounted sum of future returns with utility dependent on current income only; the subjective discount rate that equates two earnings streams is used to estimate the return to education. His measure of risk is the coefficient of variation. He finds that risk is decreasing in education and suggests that this may be because “more educated scientists are not only more productive but also more uniform in their skills, abilities and attitudes”. He finds that the inclusion of risk has negligible effects on rates of returns but that increasing the degree of risk aversion substantially reduces the return to education (despite the smaller coefficient of variation) due to decreasing marginal utility. Olson, Shefrin and White (1979) estimate a model similar to Weiss but allow borrowing while in school with repayments being made in fixed instalments once schooling has been completed. They use the residual from a regression of log earnings on a set of fixed and time varying covariates such as marital status and region and find that risk adjusted returns to college are small but positive. Nickell (1979) using UK data finds that after correcting for unemployment, the return to education using pre-tax weekly income rises by 0.6 percentage points. However including unemployment benefits and using post-tax weekly income leads the adjusted rate of return to rise by only 0.2 percentage points. He concludes that “the monetary impact of the extra schooling, insofar as it reduces the chances of unemployment, is of no great consequence”. Pistaferri and Padula using both US and Italian data find the returns to education are significantly higher when accounting for both wage and employment risk.

In this paper I analyse the returns to different education levels: college versus high school and high school versus less than high school. In keeping with earlier studies I use the IRR framework to back out the rate of return but undertake a more rigorous analysis of the amount of risk involved in each education level by separating the risk in to permanent and transitory components. In addition I look at the returns for both males and females, correct for endogenous labour market participation and use UK data.

2.2 Methodology

Individuals have CRRA preferences and choose schooling to maximize the expected utility of lifetime income

$$E_s \sum_{j=s}^T (1 + \rho)^{j-s} \frac{y_{ij}(s)^\gamma}{\gamma}$$

where $1 - \gamma$ is the coefficient of relative risk aversion, ρ is the discount rate and E_s is the expectation operator conditional on information at time s and T is the retirement age which is assumed known with certainty at the beginning of life. In this paper I abstract from consumption and thus implicitly assume that individuals care only about lifetime income. This would be the case if there were incomplete markets with no borrowing or saving. Although this may seem like a strong limitation of the paper, the literature on returns to education focus almost exclusively on the effect of education on earnings rather than utility and therefore my set up will be useful for comparisons with previous estimates.

Following Becker (1964) and Hanoch (1967) the internal rate of return (IRR) is defined as the discount rate that equates the present value of the discounted net lifetime earnings for two different schooling levels.

$$\sum_{j=t}^T (1 + \rho^*)^{j-t} \frac{E_t[(y_{ij}(s)^\gamma)]}{\gamma} = \sum_{j=t}^T (1 + \rho^*)^{j-s} \frac{E_t(y_{ij}(s')^\gamma)}{\gamma}$$

In this equation t denotes the school leaving age at which the lowest of the two schooling levels being evaluated is, while s and s' denote the two education levels being compared. I assume individuals with academic qualifications lower than a high school degree leave school at 16 and enter the labour market at age 17, those with a high school degree finish at 18 and enter the labour market at 19 while and those with a college degree are assumed to enter the labour market at 22 allowing them to stay in college until age twenty one. It is assumed that the retirement age is independent of education level and is set at age sixty five.² When comparing high school graduates and those with less than high school it is assumed that the earnings of those with a high school degree is zero while in school

²Heckman, Lochner and Todd (2006) find that allowing the retirement age to differ by education level does not lead to large changes since the earnings at the end of the life are so heavily discounted.

but they face no living costs. It is assumed that those in college have zero earnings while attending college but pay 5,000 pounds a year throughout their 3 years in college.³ This set up is advantageous over the standard Mincer regression in obtaining the IRR since the coefficient on schooling does not give an estimate of the IRR except under certain conditions which Heckman, Lochner and Todd (2006) find to be rejected in recent data. The Mincer approach gives an estimate of the growth rate of earnings with schooling. An individual deciding whether to quit schooling and enter the labour market or to continue with schooling will be most interested in the internal rate of return which requires explicitly accounting for all costs and benefits associated with each schooling level. The income process is estimated assuming income is log normally distributed $\ln y_{ij} \sim N$ and the moments are converted back using standard log normal formulae:

$$E_s[y_{ij}(s)] = \exp(E_s[\ln y_{ij}(s)] + 0.5 \text{var}_s[\ln y_{ij}(s)])$$

$$\text{var}_s[y_{ij}(s)] = \exp(2E_s[\ln y_{ij}(s)] + \text{var}_s[\ln y_{ij}(s)])(\exp(\text{var}_s[\ln y_{ij}(s)]) - 1)$$

In order to estimate the expected utility I use a second order Taylor approximation around mean earnings.

$$EU_{wr} = \frac{E_s[y_{ij}^\gamma]}{\gamma} \approx \frac{[E_s(y_{ij})]^\gamma}{\gamma} + \frac{\gamma-1}{2} \text{var}_s[y_{ij}] [E_s(y_{ij})]^{\gamma-2}$$

In terms of risk neutrality, that is when γ is equal to 1, the second part of the equation drops out and the utility is solely dependent on the mean of the distribution. When γ is equal to zero then it is a log utility function and the Taylor approximation is slightly modified. For each year the certainty equivalent value of the expected utility is calculated and used in the IRR calculation.

Two approaches are used in order to understand the effect of unemployment on the IRR. In the first case expected earnings are used whereby expected earnings is got by $E[y] = \pi_{it} y_{it} + (1 - \pi)b$ where π is the probability of employment and b denotes unem-

³In 2009 the UK government increased the maximum amount of tuition fees to 3,290 pounds per year but by 2012 the cap had been raised to 9,000 pounds. However this is just a maximum and it is possible that some universities will charge less. In future work I will look at the effect of using different costs and allow for borrowing while in college and fixed loan repayments throughout the working life in conjunction with the new maintenance loans that are available to allow for a more complete analysis of the IRR.

ployment benefit. I make no distinction between those who are unemployed and those who are out of the labour force similar to Low, Meghir and Pistaferri (2010). The unemployment benefit used is the standard weekly benefit for a single adult in 2012. This benefit differs depending on whether the individual is under 25 years of age or older and this is incorporated in to the analysis. Although the unemployment benefit is only given for a maximum of 26 weeks (after which it is means tested against household income including housing costs and household composition) for simplicity I assume that it is available for one year. Since one period corresponds to one year this also assumes that individuals are unemployed for the full year which may bias the results if the duration of unemployment varies by education. The second approach is to use the utility framework when analysing the effect of unemployment. Although the probability of employment is increasing in education, the replacement rate is significantly lower for the higher educated and this may be enough to offset any gains since risk averse individuals dislike large fluctuations in their income. I look at both approaches because only looking at unemployment risk in a utility set up may lead to changes in returns that are due solely to the parametrization of the utility function rather than the unemployment risk. The probability of employment is calculated from a probit regression of employment on a set of covariates including a quadratic in age, year dummies, dummies for parental education and a dummy denoting whether the individual is white or not. This is performed separately for each education level and for each gender. For robustness results I also examined the effects of using the mean value of employment and there was very little change in the results.

2.3 Earnings Process

In the following earnings specification log net earnings are regressed on a quadratic in age, year dummies, dummies for mother's highest education level, father's highest education level and a dummy denoting whether the individual is white or not. I do not directly control for cohort but by using year dummies and a linear age variable, I am in effect allowing for a linear cohort trend. For robustest checks I include a quartic polynomial in cohort but the results are very similar and so to increase the precision of my estimates I neglect directly controlling for cohort. Unlike many studies which use the standard Mincer regression, I use a quadratic in age rather than experience. There are two reasons for doing this; firstly experience is endogenous and secondly by controlling for experience the benefit to leaving schooling early via the effect on increased labour market experience is eliminated.

Conditioning on family background will pick up unobserved differences such as tastes and possibly ability. The regression is performed separately by gender and by education level. By estimating the regression separately by education level I allow all covariates to vary with education. This is necessary since earnings growth rates are not parallel across schooling levels.⁴ One reason for non-separability would be differential on the job investment, for example, Mincer (1991) finds that the higher educated are more likely to receive training.⁵ Log earnings are defined as usual net monthly earnings multiplied by twelve. I decided to use monthly earnings rather than annual earnings as the latter would be affected by periods of unemployment throughout the year and thus when I am computing the unemployment adjusted return, I would in effect be overestimating the effects of unemployment. Most papers use some measure of the gross wage as the dependent variable but with a progressive tax system this is not suitable since the individual will only receive a certain proportion of their income - that proportion decreasing with education - and therefore I use net earnings. I do not include variables such as marital status, region or industry since these variables can be considered as intermediate variables in that they may be outcomes themselves of the education decision and thus by controlling for these variables, I would eliminate some of the pathways through which higher returns to education are realised.

The error term consists of an individual fixed effect, a random walk and a purely transitory iid component. The transitory and permanent shocks are mean zero and serially uncorrelated.

$$Y_{i,a,t} = X_{i,a,t}\beta + f_i + u_{i,a,t} + v_{i,a,t}$$

The permanent component has a unit root such that

$$v_{i,a,t} = v_{i,a-1,t-1} + \zeta_{i,a,t}$$

⁴Migali and Walker (2011) and Heckman, Lochner and Todd (2006) find evidence against separability

⁵Additionally, this set up is advantageous over earlier studies which included years of schooling as a dummy variable in the regression and thus impose linearity in returns to schooling. The existence of "sheep skin effects" whereby the returns to schooling vary with each qualification completed and therefore the returns to an extra year of schooling which does not warrant a qualification will be less than a year of schooling which does are captured by performing the regression separately by education level.

Residual income is given by $y_{i,a,t} = Y_{i,a,t} - X_{i,a,t}\beta = f_i + u_{i,a,t} + v_{i,a,t}$

Residual income growth is therefore $g_{i,a,t} = y_{i,a,t} - y_{i,a-1,t-a} = \zeta_{i,a} + \Delta u_{i,a,t}$

Thus the variance of the permanent and transitory components can be identified by

$$Cov(g_{i,a,t}, g_{i,a,t-1} + g_{i,a,t} + g_{i,a,t+1}) = var(\zeta_{i,a,t})$$

$$Cov(g_{i,a,t}, g_{i,a,t+1}) = -var(u_{i,a,t})$$

This earnings specification assumes that measurement error in earnings is negligible or subsumed by the transitory component. However, since I am only interested in comparing the returns at two different education levels, this should not be a problem if there is no systematic difference in measurement error across education levels. Bound and Krueger (1994) find that measurement error is uncorrelated with education.⁶ I allow the variance to differ across ages as it is likely that the variance of wages will change across the life cycle due to an array of different reasons such as workers and firms gradually learning about the individual's productivity (Faber and Gobbons, 1997), differential investment in human capital and the increased occurrence of health shocks at the end of the life cycle.

2.4 Selection Bias

There are two main channels by which the IRR estimates may be biased – endogenous employment and endogenous schooling. Individuals are not randomly assigned to education or employment and therefore if there is non random selection in to either the estimates will be biased. It is likely that those individuals who acquire additional education or who choose to become employed may have higher ability or better non cognitive skills and thus it is erroneous to compare the earnings of individuals at two different education levels. In an ideal set up, earnings would be obtained on the same individual for the two different education

⁶Additionally, if measurement error is constant over time then by using the growth of the residual, any measurement error will be eliminated.

levels and the subsequent earnings would be compared, however since the individual can only choose one education stream to follow, it is necessary to somehow impute the earnings she would have received had she chosen the alternate education path. This is analogous to a Roy model (1958) where there are two states of the world but the individual can only choose one. The unobserved earnings stream is what is known as the counterfactual. In this paper I abstract from the issue of endogenous schooling but deal with endogenous employment. In the following analysis identification of the correct counterfactual is achieved by using a Heckman two step correction model which subject to some distributional assumptions allows one to overcome the issue of selection bias. Although a control function approach is less robust than other experimental estimators such as IV due to the strong distributional assumptions that are required, the advantage is that by specifying the distribution of unobservables one can identify the Average Treatment Effect. Instrumental variables on the other hand while more robust can only identify a local average treatment effect, that is, the affect of treatment on the sub population induced in to treatment due to the instrument which may not have any meaningful economic interpretation⁷.

2.5 Endogenous Labour Market Participation

Almost all studies examining the returns to schooling use current earnings as a proxy for lifetime earnings. However, the use of current earnings will lead to inconsistent estimates because earnings vary systematically over the life cycle. Workers with high lifetime earnings tend to have higher earnings growth rates than workers with low lifetime earnings and thus a comparison of earnings at the early stage of the life cycle will lead to a downward bias while comparing individuals late in life will lead to an upward bias in the estimates. Any attempt to overcome this problem by controlling for age or experience will not eliminate this bias because the result is due to heterogeneous variation around the central tendency of earnings growth (Haider and Solon, 2006). Bhuller, Mogstad and Salvanes (2011) using Norwegian data on males with almost career long earnings histories find substantial evidence of a life-cycle bias in the returns to schooling. They find a strong positive relationship between the mean age in the sample and the returns to schooling and suggest that in order to minimize bias, the sample should be restricted to individuals aged 32 to 33. However, if there is differential selection in to employment at these ages there may still be

⁷The raising of the school leaving age (Harmon and Walker (1995)) is one instance where the identified LATE may have particular policy relevance since it gives the effect of extra schooling for those who would have left without the law but now acquire extra schooling as a result of the law.

bias in the returns. They also only estimate the life-cycle bias for males so the bias may be very different for females. Given that female earnings growth rates are not as large as males one might expect any bias to be substantially smaller. However given the complexity of the interaction between female labour supply, childbearing years and education it is hard to pin point the direction of bias. If higher educated females have children at an older age then comparing females at different stages of the life cycle would induce a child bearing bias in addition to any life cycle bias. Therefore only by looking over the whole life cycle and correcting for any endogeneity can one be certain to obtain the correct earnings differential.

When looking at the returns over the whole life cycle accounting for endogenous employment decisions is crucial. Most studies looking at the return to education focus on prime labour market participation ages when individuals are most likely to be in the labour market to avoid dealing with this issue. There are substantial differences in labour market participation over the life cycle, across education groups and in the interaction of education and the life cycle. At younger ages it is likely that there is positive selection in to the labour market for the low educated if those with the least labour market value are hit with unemployment shocks. There may be negative selection for the higher educated if those with high earnings capabilities undertake MBAs or PhDs. At the end of the life cycle there is less attachment to the labour market for all education groups due to a variety of reasons including early retirement, unemployment and the fact that at the end of the life cycle the returns from investing in one's human capital are diminished due to the small amount of time left in the labour market to recoup the returns to experience (Shaw, 1989). Towards the end of the life cycle it is possible that there is positive selection in the labour market for the low educated if those who are hit with negative unemployment shocks or are discouraged from their low growth rates leave the labour market while there may be positive selection for the higher educated if those who have high earnings stay in the labour market as the opportunity cost of leaving is too great. Conversely, there may be negative selection for all education levels if better individuals who have amassed enough wealth retire early. Obviously many different hypotheses can be put forward regarding the way selection into the labour market works for each education level and for each gender; if there is positive selection of low educated workers then comparing returns at two different levels (assuming random selection in to employment of the higher educated) will lead to a downward bias in returns to education; on the other hand it could be that the average high educated worker is

of higher quality than a potential high educated worker and in this case (assuming random selection in to the lower education level) returns will be overestimated. Correcting for non random selection in to the labour market is very important if one wants to get an unbiased estimate of the returns to education. In order to address this issue I use a Heckman selection equation. To avoid identification coming exclusively from the non linearity of the inverse mills ratio, I use annual non labour income net of annual means tested cash benefits as an instrument for labour market participation. This is valid as long as non labour income significantly affects the decision to work while having no affect on subsequent earnings.⁸

Denote the latent variable for labour market participation:

$$P_{it}^* = r'_{it}\theta + \pi_{it}$$

where

$$P_{it} = 1 \quad \text{if} \quad P_{it}^* > 0$$

$$P_{it} = 0 \quad \text{otherwise}$$

Assume the errors are joint normally distributed such that

$$\begin{pmatrix} \varepsilon_{it} \\ \pi_{it} \end{pmatrix} \sim N \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{\varepsilon}^2 & \rho_{\varepsilon\pi} \\ \rho_{\varepsilon\pi} & 1 \end{pmatrix}$$

where ε_{it} is the unobserved component of the log wage equation and σ_{π}^2 has been normalised to 1.

Then

$$E[y_{it}|P_{it} = 1, x, z] = X_{i,t}\beta + \rho_{\varepsilon\pi}\lambda(r'_{it}\theta)$$

$$Var[y_{it}|P_{it} = 1, x, z] = \sigma_{y_{i,t}}^2 - \rho_{\varepsilon\pi}^2\lambda(r'_{it}\theta)(r'_{it}\theta + \lambda(r'_{it}\theta))$$

⁸ Low, Pistaferri and Meghir (2004) have used this as an instrument for participation.

where

$$\lambda(r'_{i,t}\theta) = \frac{\phi(r'_{i,t}\theta)}{\Phi(r'_{i,t}\theta)}$$

Correcting the variance of the permanent component is slightly more involved since the permanent component is identified from the growth of residual earnings. Therefore the permanent component is only identified if the individual is in the labour market for two consecutive time periods. However, assuming that the permanent error component is independent and serially uncorrelated across time periods then the variance of the permanent component can be identified from the following equations:

$$E[g_{it}|P_{it} = 1, P_{i,t-1}, x, z] = \rho_{\zeta\pi}\lambda(r'_{it}\theta)$$

$$\text{Var}[g_{it}|P_{it} = 1, P_{i,t-1} = 1, x, z] = \sigma_{\zeta,t}^2 + \sigma_{u_{i,t}}^2 + \sigma_{u_{i,t-1}}^2 + \rho_{\zeta\pi}^2\lambda(r'_{it}\theta)(r'_{it}\theta + \lambda(r'_{it}\theta))$$

$$\sigma_{u_{i,t}}^2 = -\text{Cov}(g_{i,a,t}, g_{i,a,t+1})$$

2.6 Data

The data I use is the first 18 waves of the British Household Panel Survey (BHPS).⁹ The BHPS started in 1991 and collected information on approximately 5,500 households and 10,300 individuals in England. Supplementary samples covering Scotland and Wales each containing 1,500 households were added in 1999 while in 2001 some 2,000 households representing Northern Ireland were added. The BHPS contains rich information on education, income, family background, employment and consumption. In the analysis I drop self employed individuals, those who are still in full time education and those who are older than 65 or younger than seventeen. Those with missing information on usual net monthly pay, highest education qualification, race, age or parental education are also dropped. The education variable I use is a derived variable in the BHPS denoting highest academic qualification. I group those with higher degree, first degree, HNC, HDC or teaching to the highest

⁹The BHPS in 2009 became part of the new Understanding Society Survey. This the largest longitudinal survey of it's kind in the UK sampling some 40,000 households.

schooling level which I refer to as “College”, those with A-levels comprise the schooling level which I refer to as “High School” and those with anything less than A-levels make up the lowest schooling level. Each parental education variable is categorised into no qualifications/some qualifications/further education/university qualification and finally a variable denoting whether an individual is white or not is derived. The bottom and top 1 percent of earnings in each year are trimmed to eliminate any measurement error or outliers. All imputed earnings are set to missing. Earnings and consumption are deflated to 1991 prices using the UK retail price index (RPI).

2.7 Risk Differentials

There are quite substantial differences in employment levels across education groups and over the life cycle for both males and females. For males, over the whole working life those with less than high school have average employment levels of 72.1%, for those with a high school degree it's 79.3% while those with college are employed 83% of the time. For females it's 60%, 76% and 77% respectively. Thus employment levels are increasing in education but with a larger difference between high school and less than high school degree than between college and high school. The lack of full employment highlights the importance to correct for non-random selection into the labour market. Turning to the variance of earnings where variance is got by taking the square of the residual from a regression of log earnings on a set of explanatory variables, there appears to be a u-shaped pattern in education with the variance for less than high school being 0.1405, the variance for those with a high school degree being 0.1404 while the variance of college graduates is 0.1439. Although the difference is not large across education groups, it appears that college education has the most risk.¹⁰ However once I correct for endogenous labour market participation the results are very different. The lowest education group sees the largest rise with the variance rising to 0.1807 which is not surprising given the low labour market attachment for this group. The increase for the other two groups is a lot smaller; high school variance increases to 0.159 and the variance of college graduates increases to 0.1568. For females the uncor-

¹⁰There is concern in the literature that the risk may be known to the individual in advance as they may know their own ability, motivation and other skills. To get around this I look at the variance of the growth in the residual which will eliminate the effect of any time invariant person specific effects on the variance. This leads to a variance which is decreasing in education level: 0.0568, 0.0452, 0.044. Therefore it seems that a large part of the risk in education is attributable to person heterogeneity which is fixed over time. However whether the risk is known or unknown is still an empirical question that has not been convincingly answered in the literature to date.

rected variance is largest for the low educated and smallest for the higher educated while the selection adjusted variance leads to a substantial increase in the variance of the college graduates, a small increase for the low educated and virtually no change for the high school graduates. This suggests that there is large positive or negative selection of college educated females resulting in a distribution that is truncated at the lower tail or upper tail while the unchanged variance for the high school graduates suggests that despite the less than full employment, selection in to employment is random.

Next I decompose the variance in to permanent and transitory shocks. It is important to distinguish between these two aspects since they have very different welfare effects. Transitory shocks which are temporary and short lasting include a short illness, a bonus, overtime labour supply and any mean reverting shock. It is argued that transitory shocks average out over one's lifetime or can be smoothed through savings and so do not effect welfare, Cochrane (1991) and Blundell, Pistaferri and Preston (2008) find full insurance against transitory shocks.¹¹ Conversely, permanent shocks such as low productivity, disability, promotion, and skill biased technological change are less likely to be insured against and have persistent effects on a person's earnings and welfare. For males the variance of the transitory shock is decreasing in education; 0.0149 for those with less than high school, 0.0106 for those with a high school degree and 0.0069 for those with college. Thus it appears that going to college will decrease the effects of transitory shocks which is what one would expect given that unemployment is decreasing in schooling level. The variance of transitory shocks for females is u-shaped in education with the college educated having a variance of transitory shocks equal to .023, high school graduates 0.012 and low educated 0.018. The quite substantial difference between the two highest education levels is not due to higher levels of unemployment given that they are very similar and so may be due to higher levels of mobility between jobs or differences in variation on the intensive margin if higher educated females are more likely to move in and out of part time employment.¹²

The variance of the permanent shocks for males is u-shaped in schooling level. For those with less than high school it is 0.013, those with high school degree it is 0.0102 and for those with college it is 0.0118. Correcting the variance of permanent shocks for partic-

¹¹One way transitory shocks could affect welfare is if individuals faced liquidity constraints.

¹²It could also be due to the fact that higher educated individuals have children at older ages than their lower educated counterparts and given that the distribution of earnings is more dispersed at these ages due to differential growth rates the variance of transitory shocks will be larger. This is assuming that maternity leave represents a temporary change in income and so that females do not stay out of the market for too long once they give birth.

ipation leads to increases for all education levels while maintaining the u-shaped pattern. Therefore it appears that males who do not obtain a high school degree have higher risk in terms of unemployment, higher transitory shocks and higher permanent shocks than those with either a high school degree or college education. This suggests that accounting for risk will substantially increase the returns to a high school degree. The variance of permanent shocks for females is decreasing in education while correcting for employment leads to increases at all levels with the largest increase being for the low educated which is not surprising given their low employment rates. Therefore although college educated females face larger short term risk, they face less volatility in uninsurable permanent shocks which may lead to an increase in returns to education. It should be noted that these results which use after tax earnings will differ from those using gross earnings leading to lower variance of the permanent component but increasing the volatility of the transitory shocks because now tax changes become an additional source of uncertainty.

2.8 Results

2.8.1 Males

First I look at the IRR without unemployment risk or wage risk. In this set up individuals receive expected earnings over the whole lifetime. This leads to an IRR of 6.3% to college and 7.32% to high school.¹³ Adjusting for unemployment risk but still assuming agents are risk neutral increases the return to college to 7.32% and to high school graduates to 10.72%. Allowing for risk aversion and assuming a coefficient of relative risk aversion equal to 1.5 (Attanasio and Weber (1995)) the unemployment adjusted return to college is 8.5% while it's 13.16% for high school graduates. Thus as expected the lower probability of unemployment for those with higher education leads to a higher return to schooling. Focusing on wage risk now and still assuming a coefficient of risk aversion equal to 1.5 and using the variance of the residual as a measure of wage risk, the adjusted return to college is 6.17% while the adjusted IRR for high school graduates is 8.25%. Thus the inclusion of wage risk decreases the returns to college due to the higher risk relative to high school graduates while the return to high school is increased due to the higher risk of those with

¹³This result is consistent with returns found in the literature, for example, Grenet (2009) uses the raising of the school leaving age in Britain and finds an extra year of schooling increases hourly wages by 6 to 7% and Ashenfelter et al (2000) find that studies investigating the return to an extra year of schooling, on average, find a return of 6.6%. The similarity of the results suggests that neither endogeneity nor sheepskin effects may be an issue in estimating returns.

less than high school degree although the change is not very substantial: 0.15% and 0.25% respectively. Using the variance of the permanent shocks instead of the residual to denote risk the wage risk adjusted return to college is 6.26% while it is 8% for high school. Thus using the permanent shock dampens any wage risk adjustment in returns. This is due to the fact that transitory shocks are decreasing in education and it is the transitory shocks that are driving the results with a more modest differential in permanent shocks leading to returns that are not so much affected. Although the returns to college are lower when adjustment is made for wage risk due to the higher risk involved, Weisbrod (1962) stresses that the extra risk may not be considered as decreasing utility if going to college increases the potential occupations available to the individual. Looking at the combined effect of both wage risk and unemployment risk results in returns that are similar to the unemployment adjusted return. This is due to the fact that the effect of unemployment risk is quite large while the effect of wage risk is negligible.

Next, I examine the effect of increasing the degree of relative risk aversion. Assuming a coefficient of risk aversion of 3 leads to a drop of approximately 1.4% and 1% (in absolute value) to the returns to both college and high school respectively. As the probability of employment is increasing in education this is surprising as one would expect those with higher levels of risk aversion to favour education levels with lower unemployment risk. The increase in the RRA leads to a slight reduction in the wage risk adjusted return to college and a slight increase in the wage risk adjusted return to high school. The increase in the coefficient of relative risk aversion places more weight on the reduction of risk and less weight on higher income due to the concavity of the utility function. Since having a high school degree has less wage risk the IRR to college falls while the IRR to high school rises but interestingly the increase in degree of risk aversion leads to a decrease in the unemployment adjusted IRR to both schooling levels with the biggest decrease occurring for those with college. This is due to the set up of the UK benefit system. The UK has one of the lowest replacement rates in the OECD.¹⁴ Since the unemployment benefit is independent of predisplacement earnings, the higher educated lose a higher fraction of income when unemployed¹⁵. This flat benefit system results in large swings in earnings. Thus although college graduates face a lower probability of unemployment, the large decline in earnings

¹⁴The UK ranks low amongst OECD countries with a replacement rate of under 20% compared with replacement rates of over 60% for countries such as Sweden and Portugal.

¹⁵Although not modelled here, higher educated lose most when unemployed not only due to lower replacement rates but they also have higher human capital depreciation rates and forego large gains to work experience.

on becoming unemployed makes it less valuable when an individual is averse to large deviations in earnings. In contrast, those with low education face smaller relative declines in earnings when they become unemployed. However this model abstracts from the value of leisure which may play a small part in alleviating the negative impacts of unemployment. If consumption and leisure are substitutes then the negative income effects of unemployment may be offset by the utility from increased leisure.

Turning to the results from the control function approach which correct for endogenous employment. The first stage results show that the non labour income net of benefits has a significant negative effect on employment at all education levels. While the inclusion of the inverse mills ratio in the regression suggests that workers are negatively selected in to employment at all education levels. This result could be due to negative selection in older ages outweighing positive selection at other stages of the life cycle.¹⁶ Correcting for endogenous labour market participation leads to a fall in the IRR to college and high school with the IRR to college now being 5.81% and the IRR to high school being 5.79%. The unemployment adjusted return is similar to the baseline case without the employment correction with the effect of unemployment risk increasing the IRR at all levels but with a small decrease when the degree of risk aversion is increased. The wage risk adjusted IRR is now different than the baseline case increasing returns to both college and high school. The wage risk adjusted return increases the return to high school by almost 2% while slightly increasing the returns to college. This difference in the wage risk adjusted return is because the degree of employment truncation is higher for those with lower education. This highlights the importance of accounting for selection in employment as without this correction it would appear that wage risk lowers the IRR to college while having little effect on high school returns. Increasing the coefficient of relative risk aversion to 3 leads the wage risk adjusted return to increase for both college and high school although the increase is not very large. Similar to the baseline case, using permanent shocks dampens the impact of wage risk for all levels of education. Finally, accounting for both wage risk and unemployment risk results in returns which are similar to the unemployment adjusted return.

¹⁶Interacting the inverse mills ratio with age and adding a quadratic in this interaction term shows the higher educated select positively in to the labour market but the estimate is not significant and adjusting the mean and variance for the additional selection terms leads to unreasonable results and so I decide to use the standard approach and only include the inverse mills ratio in the regression.

2.8.2 Females

The uncorrected IRR for females is similar to that for males with the return to high school being slightly larger than the return to college: 7.06% for college and 9.85% for high school. Given that males are more likely to be unemployed due to job destruction or failure to find a job, the same cannot be said for females. Many females choose not to work for a variety of reasons and so it would be incorrect to refer to non-employment as "unemployment risk" since it may be a choice. Therefore when talking about females, rather than calling it the unemployment adjusted return I will denote it as the return after correcting for periods of non-employment. This adjustment leads to a drop in the IRR to college of approximately 2.5% while the return to high school almost doubles to 16.6%. This is due to the low employment levels of low educated females. Looking at wage risk lowers the IRR to college by approximately 0.5% while increasing the IRR to high school by almost 3% signifying the u-shaped pattern in the variance across education levels. When only permanent shocks are used as a source of wage risk the adjusted return to college is now slightly larger than the baseline case while the return to high school is slightly lower and almost 3% lower than the wage risk adjusted return that uses the variance. This is in contrast to the findings for males whereby using both the variance and the variance of the permanent shocks resulted in changes in the same direction. Looking at both the unemployment and wage risk adjusted returns leads to estimates that are similar in magnitude to the unemployment adjusted return.

Increasing the degree of risk aversion to 3 leads to a negative IRR to college of -0.4% while the IRR to high school is still larger than the baseline case but significantly less than the return correcting for periods of non-employment when the degree of risk aversion is 1.5. This highlights that returns to education are heterogeneous in the degree of risk aversion. Those with extremely large levels of risk aversion may be put off from college by the large swings in potential earnings due to the low replacement rates. The wage risk adjusted return now falls slightly for college and rises slightly for high school. Increasing the degree of risk aversion has large effects on the unemployment adjusted return but the effect on wage risk is fairly innocuous. The low return to college when unemployment risk is included suggests that given the likelihood that females may drop out of the labour market once they start a family and given that labour supply is the channel by which returns to human capital investment are reaped, it may seem that investing in college for females may not be such a worthwhile pursuit. However, if by going to college females marry college educated men

this will tend to increase the overall returns.

Correcting for participation leads the return to college to more than double to 14.5% while the return to high school falls to 6.2%. The large increase in the return to college is consistent with the hypothesis that the lower educated select positively in to employment although the coefficient on the inverse mills ratio is negative suggesting negative selection in to employment at all education levels.¹⁷ If those college educated women who would earn the most if they participate in the labour market are married to high earning men and due to high levels of wealth decide not to participate in the labour market then selection adjusted returns will be larger. Increasing the degree of risk aversion leads to similar results to the baseline case. Overall it appears that correcting for selection in to employment is particularly large for females leading to substantial changes in the rate of return. Unemployment risk decreases the IRR to college while significantly increasing the IRR to high school; wage risk decreases the returns to college while increasing the returns to high school but using only uninsurable risk as represented by the permanent shocks counteracts these effects resulting in a wage risk adjusted return that is similar to the unadjusted return.

2.9 Advanced Information and Insurance

There is a huge debate about whether the variance observed by the econometrician represents risk from the agent's point of view. Recent studies by Heckman, Cunha and Navarro (2005) find that individuals know at least 40 percent of the risk due to heterogeneity and therefore neglecting this insight will lead to an overestimate of risk. They estimate the correlation between observed outcomes and the agent's schooling decision to infer the amount of risk that is known in advance. This could be due to individual's knowing their own ability, motivation etc. Since in my analysis I used the variance of the permanent shocks as a measure of risk which uses the residual growth rate, this will eliminate any fixed unobserved factors such as ability that may be known to the individual. However it is possible that there are time varying factors such as promotion or demotion that may be known in advance to the individual. It is possible to test for advanced information if one has data on consumption. In the BHPS there is data for each year on usual weekly food expenditure.¹⁸ While it would be better to have consumption items such as expenditure on clothing, travel, and other non durables, due to a lack of data covering these items, previous studies have also used food

¹⁷The first stage results show that the instrument has negative effects on employment at all education levels.

¹⁸This includes takeaways but excludes meals eaten in restaurants.

expenditure, for example, Zeldes (1989). The permanent income hypothesis predicts that consumption should only react to unanticipated permanent income shocks as it is assumed that agents can perfectly insure themselves against transitory shocks via savings.¹⁹ It is likely that if one expects a promotion at the end of the term then this will be factored in to consumption decisions from today and although from the econometricians point of view it registers as a shock when the agent gets promoted, this will not be a shock to the agent. Following Blundell, Pistaferri and Preston (2008) I test for evidence of superior information. The basic idea is that if income is anticipated by the agent then future income growth should be correlated with current consumption growth. If there is a significant correlation between current consumption growth and future income growth then this implies that the agent has more information and such shocks that appear as risk may in fact be already known to the individual. I regress the real values of log consumption and log earnings on a wide set of covariates including dummies for year of birth, year, household size, job status, number of children, region, marital status and race separately for each gender and education level and use the residuals from these regressions in the test. Tables 1 and 2 shows that for those with college the test of no correlation between current consumption growth and future income growth is rejected for both males and females.

Table 2.1: Males: P-values for test of null hypotheses of no correlation between consumption growth and future income growth for all years

| | Less than HS | High School | College |
|------------------------------------------------------------|--------------|-------------|---------|
| Test $Cov(\Delta y_{i,t+1}, \Delta c_{i,t}) = 0$ for all t | 0.197 | 0.286 | 0.112 |
| Test $Cov(\Delta y_{i,t+2}, \Delta c_{i,t}) = 0$ for all t | 0.202 | 0.745 | 0.921 |
| Test $Cov(\Delta y_{i,t+3}, \Delta c_{i,t}) = 0$ for all t | 0.593 | 0.502 | 0.604 |
| Test $Cov(\Delta y_{i,t+4}, \Delta c_{i,t}) = 0$ for all t | 0.694 | 0.747 | 0.669 |

Table 2.2: Females: P-values for test of null hypotheses of no correlation between consumption growth and future income growth for all years

| | Less than HS | High School | College |
|------------------------------------------------------------|--------------|-------------|---------|
| Test $Cov(\Delta y_{i,t+1}, \Delta c_{i,t}) = 0$ for all t | 0.456 | 0.996 | 0.268 |
| Test $Cov(\Delta y_{i,t+2}, \Delta c_{i,t}) = 0$ for all t | 0.091 | 0.853 | 0.219 |
| Test $Cov(\Delta y_{i,t+3}, \Delta c_{i,t}) = 0$ for all t | 0.607 | 0.070 | 0.413 |
| Test $Cov(\Delta y_{i,t+4}, \Delta c_{i,t}) = 0$ for all t | 0.111 | 0.642 | 0.878 |

¹⁹However observing a muted response of consumption to permanent shocks to income could be due to advanced information or insurance. Van Rens and Primiceri (2009) attribute any observed change in permanent shocks which do not translate in to consumption changes as information that must have been known in advance. However this assumes that there is no insurance available for permanent shocks.

On the other hand if there are mechanisms available that provide insurance against income shocks this will help to alleviate the negative impact of the shocks. The availability of insurance may differ across education levels and this will affect the internal rate of return. Possible insurance mechanisms include savings, borrowing, spousal labour supply, social networks and government transfers. It is likely that higher educated individuals would benefit most from the savings and borrowing channel since their high earnings mean they can afford to build up a buffer stock of precautionary savings to insure against adverse shocks while also making them more attractive from a lender's view point and furthermore it is likely they have better credit history compared to the low educated who are more likely to default on loans. There is large evidence that spousal labour supply can act as an insurance mechanism, for example, Shelly Lundberg (1985) finds evidence that wife's labour supply increases in response to husband's negative income shocks, Devereux (2003) finds that a 10 percent fall in husband's wage leads to a 4% increase in wife's hours of work while Blundell, Pistaferri and Saporta-Ecksten (2012) find a 10% decrease in male wages leads to only a 4.4% decrease in household consumption due to an increase in spouse's labour supply. While this may differ across education levels, assortative mating (Neal, 2004) would imply that a higher educated spouse would command a higher wage and be more likely to get a job than a lower educated spouse. Similarly, if those who attend college have a social network which includes individuals who have also attended college then it is more likely that this channel would provide some benefit. The role of government transfers is the most important and widely available avenue for providing insurance to individuals through unemployment insurance, disability benefits, etc. To the extent that government transfers are means tested the lower educated will benefit the most from this insurance channel. Blundell, Graber and Mogstad (2012) using Norwegian data find that taxes and transfers play a substantial role in sheltering individuals from the adverse consequences of income shocks with particular benefit for the low educated.

2.10 Conclusion

The majority of studies investigating the return to education do not adjust returns to account for employment or wage risk which is equivalent to assuming that risk is constant across education levels. This paper has provided evidence that there are significant differences in employment probabilities and wage risk across education levels and that failure to account for these differences in estimating the return to education will lead to substantially biased

estimates. The main finding is that for males the returns to both high school and college increase substantially once unemployment risk is taken in to account. Correcting for endogenous employment decisions results in significant changes in the results highlighting the need for studies investigating the returns to education to correct for non random selection in to employment. Overall, the results suggest that accounting for risk leads to large changes in returns via employment risk while the wage risk adjusted returns are not as substantial. It is imperative that future studies investigating the returns to education take into consideration the effects of risk and also endogenous employment decisions.

The results in this paper only focus on monetary returns but it is conceivable that the non pecuniary returns are quite large. There is strong evidence that increased education lowers crime, reduces the incidence of teenage birth, increases health and happiness and leads to better decision making (see Oreopoulos and Salvanes (2011) and the references therein). Thus incorporating these non-monetary returns may lead to quite different results. Of course this paper has not accounted for the endogeneity of education. Finding a credible instrument for schooling is quite difficult but a must if one wants to be sure that the results represent the casual effect of schooling and therefore future work will attempt to correct for this endogeneity. Furthermore, the selection corrected results rely on the assumption that the error terms determining the earnings and employment are jointly normal. Further work relaxing this assumption is warranted. Das, Newey and Vella (2003) have used non parametric selection models to investigate the returns to schooling. Lastly, the use of a structural dynamic model may lead to insight in to the dynamics of unemployment risk and wage risk over the life cycle and how the effect of differential experience accumulation and human capital depreciation which vary by education level may affect the returns.

2.11 References

Ashenfelter, Orley., Harmon, Colm., Oosterbeek, Hessel. (1999). "A review of estimates of the schooling/earnings relationship, with tests for publication bias," *Labour Economics*, Elsevier, vol. 6(4), pages 453-470, November.

Attanasio, Orazio P., Weber, Guglielmo. (1995). "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey" *Journal of Political Economy*, Vol. 103, No. 6, pp. 1121-1157 Decemeber.

Becker, G.S. (1962). "Investment in Human Capital: A Theoretical Analysis", *Journal of Political Economy*". Vol. 70, No. 5. October.

Blundell, R., Graber, Michael., Mogstad, Magne. (2012). *Labor Income Dynamics and the Insurance from Taxes, Transfers and the Family*. Working Paper.

Blundell, Richard, Pistaferri, Luigi, Preston, Ian, (2008). "Consumption inequality and partial insurance". *American Economic Review*, 98 (5), 1887-1921.

Blundell, Richard, Reed, Howard, Stoker, Thomas M. (2003). "Interpreting Aggregate Wage Growth: The Role of Labor Market Participation". *The American Economic Review*, Vol. 93, No. 4. , pp. 1114-1131. November

Blundell, Richard., Pistaferri, Luigi., Saporta-Eksten, Itay. (2012). "Consumption Inequality and Family Labor Supply," NBER Working Papers 18445, National Bureau of Economic Research, Inc.

Bhuller, Manudeep, Mogstad, Magne, Salvanes, Kjell G., (2011). "Life-Cycle Bias and the Returns to Schooling in Current and Lifetime Earnings," IZA Discussion Papers 5788, Institute for the Study of Labor (IZA).

Cameron. Stephen V., Taber, Christopher (2004). "Estimation of Educational Borrowing Constraints Using Returns to Schooling" *Journal of Political Economy*, Vol. 112, No.

1, pp. 132-182, February.

Card, D. (1995). Earnings, Schooling, and Ability Revisited. *Research in Labor Economics*, edited by Solomon Polachek. Greenwich, CT: JAI.

Card, D. (2001). "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems". *Econometrica*, vol. 69(5), pages 1127-60. September.

Chen, Stacey H. (2008) "Estimating the variance of wages in the presence of selection and unobserved heterogeneity". *The Review of Economics and Statistics*, 90(2) 275-289. May

Cochrane, John H. (1991) "A Simple Test of Consumption Insurance" , *Journal of Political Economy*, Vol. 99, No. 5, pp. 957-976. October.

Cunha, F., Heckman, J.J., Navarro, S. (2005). Separating uncertainty from heterogeneity in life cycle earnings, The 2004 Hicks lecture. *Oxford Economic Papers* 57 (2), 191-261. April.

Devereux, Paul J., (2004). "Changes in Relative Wages and Family Labor Supply". *The Journal of Human Resources* , Vol. 39, No. 3, pp. 696-722. Summer.

Gronau, Reuben (1974). "Wage Comparisons—A Selectivity Bias *Journal of Political Economy*", Vol. 82, No. 6, pp. 1119-1143, November-December.

Hanoch, Giora. (1967). "An economic analysis of earnings and schooling". *The Journal of Human Resources*

Harmon, C. and Walker, I. (1995). 'Estimates of the economic return to schooling for the UK', *American Economic Review*, vol. 85, pp. 1278-86.

Heckman, James. (1979). "Sample Selection Bias as a Specification Error." *Econometrica* 47(1): 153-62.

Heckman, J.J, Lochner, L.J., Todd, P.E. (2006). "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond". Handbook of the Economics of Education, Volume 1 Chapter 7

Heckman, J.J, Lochner, L.J., Todd, P.E. (2008). "Earnings Functions and Rates of Return". Journal of Human Capital , Vol. 2, No. 1 , pp. 1-31. Spring

Heckman, J.J, Vytlačil E.J. (2001). "Identifying the role of cognitive ability in explaining the level of and change in the return to schooling". The Review of Economics and Statistics, Vol. LXXXIII. February

Lillard, L. A., Weiss, Y. (1977). "Components of Variation in Panel Earnings Data: American Scientists 1960-70." Nat. Bur. Econ. Res. Working Paper no. 121, rev. Stanford, Calif.: Center Econ. Analysis Human Behavior and Soc. Inst.

Lundberg, Shelly. (1988). "Labor Supply of Husbands and Wives: A Simultaneous Equations Approach". The Review of Economics and Statistics, Vol. 70, No. 2, pp. 224-235, May.

Mincer, J. (1974). "Schooling, Experience and Earnings". Columbia University Press: New York.

Mincer, J. (1991). Education and unemployment (No. w3838). National Bureau of Economic Research.

Meghir, Costas., Pistaferri, Luigi. (2004). "Income Variance Dynamics and Heterogeneity, , Econometrica, Vol 72 No 1, January.

Meghir, Costas., Pistaferri, Luigi. (2011). "Earnings, Consumption and Life Cycle Choices," Handbook of Labor Economics, Elsevier.

Nickell, S. (1979). "Education and Lifetime Patterns of Unemployment". *Journal of Political Economy*, Vol. 87, No. 5, pp. S117-S131. October.

Olson, Lawrence., Shefrin, H.M., White, Halbert. (1979). "Optimal Investment in Schooling When Incomes Are Risky". *Journal of Political Economy* , Vol. 87, No. 3, pp. 522-539. June.

Oreopoulos, Philip., Salvanes, Kjell G. (2011) "Priceless: The Nonpecuniary Benefits of Schooling" , *Journal of Economic Perspectives*, American Economic Association, vol. 25(1), pages 159-84, Winter.

Primiceria, Giorgio E., Van Rens,Thijs. (2009). "Heterogeneous life-cycle profiles, income risk and consumption inequality" *Journal of Monetary Economics*, Volume 56, Issue 1, January 2009, Pages 2039

Shaw, Kathryn L. (1989) "Life-Cycle Labor Supply with Human Capital Accumulation". *International Economic Review*, Vol. 30, No. , pp. 431-456. May

Weisbrod, Burton A. (1962). "Education and Investment in Human Capital," *Journal of Political Economy*, University of Chicago Press, vol. 70, pages 106.

Weiss, Yoram. (1972) "The Risk Element in Occupational and Educational Choices" *Journal of Political Economy*, Vol. 80, No. 6 , pp. 1203-1213, November-December.

Zeldes, Stephen. P. (1989). "Consumption and Liquidity Constraints: An Empirical Investigation," *Journal of Political Economy*, University of Chicago Press, vol. 97(2), pages 305-46, April.

2.12 Appendix

Table 2.3: Males: OLS Regression of Log Annual Net Earnings

| | Low Education (1) | High School (2) | College (3) |
|--------------------------|------------------------|------------------------|------------------------|
| Age | 0.0816*** (0.0031) | 0.0877*** (0.0068) | 0.0942*** (0.0080) |
| Age Squared | -0.0947*** (0.0039) | -0.1002*** (0.0087) | -0.1070*** (0.0097) |
| Mother_Some Quals | 0.0669*** (0.0214) | 0.0784** (0.0330) | -0.0197 (0.0272) |
| Mother_Further Education | 0.0633*** (0.0236) | 0.0564 (0.0390) | -0.0217 (0.0393) |
| Mother_University | 0.1386*** (0.0487) | -0.0196 (0.0593) | -0.0103 (0.0517) |
| Father_Some Quals | 0.0054 (0.0223) | 0.0607* (0.0364) | 0.0673** (0.0317) |
| Father_Further Education | 0.0174 (0.0205) | 0.0480 (0.0331) | -0.0126 (0.0290) |
| Father_University | 0.0567 (0.0408) | 0.1574*** (0.0517) | 0.0593 (0.0453) |
| White | 0.0453 (0.0470) | 0.0444 (0.0995) | 0.0732 (0.0624) |
| Observations | 18270 | 7406 | 7983 |
| N_clust | 2346 | 850 | 895 |

Standard Errors reported in parentheses are clustered at the individual level.

Significance levels: * 10%, ** 5%, *** 1%. Regressions include year dummies.

(d) denotes dummy variable. The base category for White is Non-White.

The base category for parental education is No Qualifications.

Table 2.4: Males: First Stage Estimates Probit of Employment using Non Labour Income as an instrument

| | Low Education (1) | High School (2) | College (3) |
|------------------------------|------------------------|------------------------|------------------------|
| Age | 0.0488*** (0.0030) | 0.0310*** (0.0034) | 0.0231*** (0.0039) |
| Age Squared | -0.0649*** (0.0035) | -0.0411*** (0.0042) | -0.0306*** (0.0045) |
| Mother_Some Quals (d) | 0.0231 (0.0229) | 0.0398* (0.0207) | 0.0188 (0.0116) |
| Mother_Further Education (d) | 0.0024 (0.0332) | 0.0366** (0.0176) | -0.0057 (0.0191) |
| Mother_University (d) | -0.0758 (0.0574) | 0.0355 (0.0257) | 0.0031 (0.0306) |
| Father_Some Quals (d) | 0.0166 (0.0227) | -0.0058 (0.0270) | 0.0176 (0.0132) |
| Father_Further Education (d) | -0.0027 (0.0248) | -0.0348 (0.0214) | 0.0191 (0.0122) |
| Father_University (d) | 0.0670** (0.0284) | -0.0743 (0.0605) | 0.0128 (0.0200) |
| White (d) | 0.0649 (0.0551) | -0.0679*** (0.0114) | -0.0444*** (0.0099) |
| Non-Labour Income | -0.0508*** (0.0043) | -0.0302*** (0.0040) | -0.0290*** (0.0031) |
| Observations | 9748 | 4772 | 6001 |
| N_clust | 1984 | 739 | 897 |

Standard Errors reported in parentheses are clustered at the individual level.

Significance levels: * 10%, ** 5%, *** 1%. Regressions include year dummies.

(d) denotes dummy variable. The base category for White is Non-White.

The base category for parental education is No Qualifications.

Non labour income refers to annual non labour income excluding benefits.

Table 2.5: Males: Including Inverse Mills Ratio in OLS Regression of Log Annual Net Earnings

| | Low Education (1) | High School (2) | College (3) |
|-----------------------------|------------------------|------------------------|------------------------|
| Age | 0.0736*** (0.0069) | 0.0724*** (0.0125) | 0.0684*** (0.0112) |
| Age Squared | -0.0807*** (0.0091) | -0.0812*** (0.0161) | -0.0763*** (0.0140) |
| Mother_Some Quals | 0.0525 (0.0329) | 0.0334 (0.0495) | -0.0180 (0.0304) |
| Mother_Further Education | 0.0512 (0.0355) | 0.0147 (0.0507) | -0.0610 (0.0461) |
| Mother_University | 0.1550* (0.0796) | -0.1270 (0.0846) | -0.0566 (0.0566) |
| Father_Some Quals | 0.0128 (0.0329) | 0.0843* (0.0497) | 0.0353 (0.0338) |
| Father_Further Education | 0.0376 (0.0267) | 0.0415 (0.0472) | -0.0605* (0.0344) |
| Father_University | 0.0508 (0.0618) | 0.2302*** (0.0770) | 0.0590 (0.0518) |
| White | 0.0407 (0.0685) | 0.0235 (0.1304) | 0.0356 (0.0684) |
| Inverse Mills Ratio_Low | -0.3464*** (0.0727) | | |
| Inverse Mills Ratio_HS | | -0.2354* (0.1275) | |
| Inverse Mills Ratio_College | | | -0.2532** (0.1106) |
| Observations | 7157 | 3815 | 4833 |
| N_clust | 1537 | 625 | 771 |

SEs in parentheses are computed using block bootstrap method.

Based on 500 replications and used to *a/c* for the pre-estimated IVM ratio.

Significance levels: * 10%, ** 5%, *** 1%. Regressions include year dummies.

(d) denotes dummy variable. The base category for White is Non-White.

The base category for parental education is No Qualifications.

Non labour income refers to annual non labour income excluding benefits.

Table 2.6: Males: Probit Regression of Employment

| | Low Education (1) | High School (2) | College (3) |
|------------------------------|------------------------|------------------------|------------------------|
| Age | 0.0410*** (0.0025) | 0.0364*** (0.0034) | 0.0365*** (0.0036) |
| Age Squared | -0.0582*** (0.0030) | -0.0502*** (0.0041) | -0.0491*** (0.0041) |
| Mother_Some Quals (d) | 0.0605*** (0.0179) | 0.0692*** (0.0233) | 0.0135 (0.0145) |
| Mother_Further Education (d) | 0.0456** (0.0223) | 0.0334 (0.0229) | -0.0043 (0.0184) |
| Mother_University (d) | 0.0426 (0.0393) | 0.0653*** (0.0247) | -0.0214 (0.0343) |
| Father_Some Quals (d) | 0.0266 (0.0206) | -0.0351 (0.0340) | 0.0074 (0.0170) |
| Father_Further Education (d) | 0.0199 (0.0185) | -0.0421* (0.0242) | 0.0197 (0.0130) |
| Father_University (d) | 0.0595* (0.0360) | -0.0551 (0.0441) | 0.0135 (0.0209) |
| White (d) | 0.0673 (0.0418) | -0.0455 (0.0302) | 0.0451 (0.0478) |
| Observations | 26728 | 9702 | 10178 |
| N_clust | 3035 | 994 | 1040 |

Standard Errors reported in parentheses are clustered at the individual level.

Significance levels: * 10%, ** 5%, *** 1%. Regressions include year dummies.

(d) denotes dummy variable. The base category for White is Non-White.

The base category for parental education is No Qualifications.

Table 2.7: Females: OLS Regression of Log Annual Net Earnings

| | Low Education (1) | High School (2) | College (3) |
|--------------------------|------------------------|------------------------|------------------------|
| Age | 0.0311*** (0.0042) | 0.0318*** (0.0102) | 0.0487*** (0.0111) |
| Age Squared | -0.0429*** (0.0054) | -0.0447*** (0.0135) | -0.0587*** (0.0141) |
| Mother_Some Quals | 0.0781*** (0.0264) | 0.0580 (0.0455) | 0.0219 (0.0437) |
| Mother_Further Education | 0.1111*** (0.0309) | 0.1856*** (0.0495) | 0.0132 (0.0478) |
| Mother_University | 0.3232*** (0.0540) | 0.1976*** (0.0750) | -0.0690 (0.0686) |
| Father_Some Quals | 0.0405 (0.0292) | -0.0024 (0.0556) | -0.0683 (0.0482) |
| Father_Further Education | 0.0589** (0.0256) | -0.0590 (0.0439) | -0.0164 (0.0446) |
| Father_University | 0.0440 (0.0542) | 0.0436 (0.0698) | -0.1045 (0.0648) |
| White | -0.1650*** (0.0615) | -0.0705 (0.1284) | 0.0117 (0.0899) |
| Observations | 24201 | 6577 | 7505 |
| N_clust | 3092 | 823 | 923 |

Standard Errors reported in parentheses are clustered at the individual level.

Significance levels: * 10%, ** 5%, *** 1%. Regressions include year dummies.

(d) denotes dummy variable. The base category for White is Non-White.

The base category for parental education is No Qualifications.

Table 2.8: Females: First Stage Estimates Probit of Employment using Non Labour Income as an instrument

| | Low Education (1) | High School (2) | College (3) |
|------------------------------|------------------------|------------------------|------------------------|
| Age | 0.0537*** (0.0036) | 0.0277*** (0.0062) | 0.0445*** (0.0066) |
| Age Squared | -0.0755*** (0.0042) | -0.0417*** (0.0073) | -0.0590*** (0.0075) |
| Mother_Some Quals (d) | 0.0387 (0.0258) | 0.0093 (0.0327) | 0.0635** (0.0298) |
| Mother_Further Education (d) | 0.0558** (0.0283) | 0.0314 (0.0326) | 0.0396 (0.0307) |
| Mother_University (d) | 0.0957* (0.0529) | 0.0163 (0.0488) | 0.0932** (0.0376) |
| Father_Some Quals (d) | -0.0018 (0.0312) | -0.0344 (0.0389) | -0.0241 (0.0406) |
| Father_Further Education (d) | -0.0179 (0.0235) | -0.0133 (0.0325) | 0.0036 (0.0319) |
| Father_University (d) | -0.0274 (0.0502) | -0.0038 (0.0421) | -0.0382 (0.0479) |
| White (d) | 0.1311** (0.0612) | 0.0535 (0.0913) | 0.0990 (0.1186) |
| Non-Labour Income | -0.0417*** (0.0044) | -0.0373*** (0.0060) | -0.0406*** (0.0058) |
| Observations | 15382 | 4155 | 5723 |
| N_clust | 2822 | 705 | 886 |

Standard Errors reported in parentheses are clustered at the individual level.

Significance levels: * 10%, ** 5%, *** 1%. Regressions include year dummies.

(d) denotes dummy variable. The base category for White is Non-White.

The base category for parental education is No Qualifications.

Non labour income refers to annual non labour income excluding benefits.

Table 2.9: Females: Including Inverse Mills Ratio in OLS Regression of Log Annual Net Earnings

| | Low Education (1) | High School (2) | College (3) |
|-----------------------------|------------------------|------------------------|------------------------|
| Age | 0.0736*** (0.0072) | 0.0724*** (0.0119) | 0.0684*** (0.0112) |
| Age Squared | -0.0807*** (0.0096) | -0.0812*** (0.0152) | -0.0763*** (0.0139) |
| Mother_Some Quals | 0.0525* (0.0303) | 0.0334 (0.0463) | -0.0180 (0.0294) |
| Mother_Further Education | 0.0512 (0.0340) | 0.0147 (0.0547) | -0.0610 (0.0463) |
| Mother_University | 0.1550** (0.0750) | -0.1270 (0.0867) | -0.0566 (0.0586) |
| Father_Some Quals | 0.0128 (0.0344) | 0.0843* (0.0497) | 0.0353 (0.0343) |
| Father_Further Education | 0.0376 (0.0270) | 0.0415 (0.0455) | -0.0605* (0.0356) |
| Father_University | 0.0508 (0.0564) | 0.2302*** (0.0774) | 0.0590 (0.0529) |
| White | 0.0407 (0.0677) | 0.0235 (0.1385) | 0.0356 (0.0665) |
| Inverse Mills Ratio_Low | -0.3464*** (0.0768) | | |
| Inverse Mills Ratio_HS | | -0.2354* (0.1278) | |
| Inverse Mills Ratio_College | | | -0.2532** (0.1131) |
| Observations | 7157 | 3815 | 4833 |
| N_clust | 1537 | 625 | 771 |

SEs in parentheses are computed using block bootstrap method.

Based on 500 replications and used to *a/c* for the pre-estimated IVM ratio.

Significance levels: * 10%, ** 5%, *** 1%. Regressions include year dummies.

(d) denotes dummy variable. The base category for White is Non-White.

The base category for parental education is No Qualifications.

Non labour income refers to annual non labour income excluding benefits.

Table 2.10: Females: Probit regression of Employment

| | Low Education (1) | High School (2) | College (3) |
|------------------------------|------------------------|------------------------|------------------------|
| Age | 0.0456*** (0.0027) | 0.0235*** (0.0052) | 0.0512*** (0.0054) |
| Age Squared | -0.0631*** (0.0031) | -0.0365*** (0.0063) | -0.0679*** (0.0062) |
| Mother_Some Quals (d) | 0.0768*** (0.0177) | 0.0243 (0.0277) | 0.0484* (0.0264) |
| Mother_Further Education (d) | 0.0955*** (0.0204) | 0.0335 (0.0289) | 0.0252 (0.0273) |
| Mother_University (d) | 0.1208*** (0.0413) | 0.0376 (0.0388) | -0.0118 (0.0410) |
| Father_Some Quals (d) | 0.0276 (0.0208) | 0.0270 (0.0307) | -0.0159 (0.0335) |
| Father_Further Education (d) | 0.0134 (0.0170) | 0.0348 (0.0262) | 0.0107 (0.0275) |
| Father_University (d) | 0.0151 (0.0375) | 0.0350 (0.0357) | -0.0086 (0.0371) |
| White (d) | 0.1678*** (0.0432) | 0.0842 (0.0816) | 0.0611 (0.0794) |
| Observations | 42208 | 9210 | 10664 |
| N_clust | 4187 | 946 | 1081 |

Standard Errors reported in parentheses are clustered at the individual level.

Significance levels: * 10%, ** 5%, *** 1%. Regressions include year dummies.

(d) denotes dummy variable. The base category for White is Non-White.

The base category for parental education is No Qualifications.

Table 2.11: IRR

| | No Risk | Expected Pay |
|-------------|---------|--------------|
| College | 0.063 | 0.0732 |
| High School | 0.0789 | 0.1072 |

Table 2.12: IRR with CRRA equal 1.5

| | Unemployment Risk | WR (res) | Both (res) | WR (perm) | Both (perm) |
|-------------|-------------------|----------|------------|-----------|-------------|
| College | 0.0850 | 0.0617 | 0.08164 | 0.0626 | 0.0844 |
| High School | 0.1316 | 0.0825 | 0.1306 | 0.08 | 0.1320 |

Table 2.13: IRR with CRRA equal 3

| | Unemployment Risk | WR (res) | Both (res) | WR (perm) | Both (perm) |
|-------------|-------------------|----------|------------|-----------|-------------|
| College | 0.0711 | 0.0609 | 0.0669 | 0.0621 | 0.0705 |
| High School | 0.1208 | 0.0832 | 0.1173 | 0.0811 | 0.1208 |

Note: The term in parentheses denotes which measure of wage risk was used in calculating the estimates. res denotes the variance of the residual and perm denotes the variance of permanent shocks.

Table 2.14: IRR Corrected for Endogenous Employment

| | No Risk | Expected Pay |
|-------------|---------|--------------|
| College | 0.0581 | 0.0677 |
| High School | 0.0579 | 0.092 |

Table 2.15: IRR Corrected for Endogenous Employment CRRA equal 1.5

| | Unemployment Risk | WR (res) | Both (res) | WR (perm) | Both (perm) |
|-------------|-------------------|----------|------------|-----------|-------------|
| College | 0.0804 | 0.0598 | 0.0792 | 0.0586 | 0.0805 |
| High School | 0.1232 | 0.0762 | 0.1306 | 0.0658 | 0.1284 |

Table 2.16: IRR Corrected for Endogenous Employment CRRA equal 3

| | Unemployment Risk | WR (res) | Both (res) | WR (perm) | Both (perm) |
|-------------|-------------------|----------|------------|-----------|-------------|
| College | 0.0661 | 0.0603 | 0.0579 | 0.0588 | 0.0657 |
| High School | 0.1169 | 0.0815 | 0.1163 | 0.0718 | 0.1183 |

Table 2.17: IRR Females

| | No Risk | Expected Pay |
|-------------|---------|--------------|
| College | 0.0706 | 0.0656 |
| High School | 0.0985 | 0.1651 |

Table 2.18: IRR with CRRA equal 1.5

| | Periods of Non-Employment | WR (res) | Both (res) | WR (perm) | Both (perm) |
|-------------|---------------------------|----------|------------|-----------|-------------|
| College | 0.0449 | 0.0669 | 0.0425 | 0.0718 | 0.0457 |
| High School | 0.1873 | 0.1258 | 0.1903 | 0.0982 | 0.1853 |

Table 2.19: IRR with CRRA equal 3

| | Periods of Non-Employment | WR (res) | Both (res) | WR (perm) | Both (perm) |
|-------------|---------------------------|----------|------------|-----------|-------------|
| College | -0.004 | 0.0653 | 0.0024 | .0729 | -0.0029 |
| High School | 0.1285 | 0.1308 | 0.1293 | 0.0979 | 0.1275 |

Table 2.20: IRR Females Corrected for Endogenous Employment

| | No Risk | Expected Pay |
|-------------|---------|--------------|
| College | 0.1448 | 0.1310 |
| High School | 0.0617 | 0.1214 |

Table 2.21: IRR Females Corrected for Endogenous Employment CRRA equal 1.5

| | Periods of Non-Employment | WR (res) | Both (res) | WR (perm) | Both (perm) |
|-------------|---------------------------|----------|------------|-----------|-------------|
| College | 0.0821 | 0.1016 | 0.0579 | 0.1510 | 0.0867 |
| High School | 0.1629 | 0.0973 | 0.17 | 0.1766 | 0.1668 |

Table 2.22: IRR Corrected for Endogenous Employment CRRA equal 3

| | Periods of Non-Employment | WR (res) | Both (res) | WR (perm) | Both (perm) |
|-------------|---------------------------|----------|------------|-----------|-------------|
| College | 0.0003 | 0.095 | 0.0038 | 0.1559 | 0.11 |
| High School | 0.1209 | 0.098 | 0.1191 | 0.088 | 0.1252 |

Chapter 3

Education, Uncertainty and Business Cycles

**This is joint work with Paul J Devereux (University College Dublin)*

3.1 Introduction

Earnings volatility is a feature of modern labour markets as individual workers are subject to wage and employment variation. In the absence of full insurance against labour market adversities, volatility can have large effects on the welfare of individuals (Low et al (2010), Heathcote et al. (2014), French (2005), Banks et al (2001), Blundell et al (2008)). These issues are especially pertinent in times of recession when employers faced with rising costs and falling demand seek to cut costs by either firing workers and/or implementing wage cuts. Can investments in education provide shelter against these economic uncertainties? And do policies that increase educational attainment, such as compulsory schooling laws, reduce the earnings volatility faced by workers? These questions are important given the recent great recession which had large negative effects on many individuals.

While some research has studied the relationship between education and earnings volatility, there is no general consensus.¹ Moreover, these estimates represent correlations rather than the causal effect of education on earnings volatility. Persons with more education may have higher ability and differ in other unobservable ways that lead them to have less or more earnings volatility independent of their

¹Meghir and Pistaferri (2004) using the PSID find a u-shaped pattern with high school graduates having lower earnings volatility than high school dropouts but higher than college graduates, while Blundell et al (2015) using Norwegian data find that the interaction of earnings volatility and the life cycle differs by education.

educational attainment.

There is very little evidence about the causal effects of education on earnings volatility. While it is not the central focus of her paper, Chen (2008) uses U.S. data to estimate the transitory component of earnings using a parametric selection model. Contemporaneous to our research, Liu et al. (2015) use compulsory schooling changes in Norway to estimate the transitory component of earnings for individuals with low levels of education. Our focus is different in that we study direct measures of earnings volatility experienced by individuals including the standard deviation of individual earnings, the cyclical nature of earnings, and the probability of a pay cut. In fact, to our knowledge, this is the first paper that studies the direct **causal effect** of education on these particular measures. Additionally, we contribute to the literature by studying a highly cyclical labour market up to and including the great recession.

Our analysis uses a regression discontinuity design based on a change in compulsory schooling combined with a large panel dataset. We find that an additional year of education leads to a reduced standard deviation of log earnings of about 0.01 (around 10% of the mean standard deviation), decreased probability of receiving a pay cut of about 3.5 percentage points and to a lower level of earnings cyclical nature. These results are robust to many specifications. A recent survey paper highlights that education leads to large benefits in terms of wages, health, employment, voting behaviour, crime, teenage pregnancy, decision making and in many other dimensions (Oreopoulos and Salvanes, 2011); we show an additional channel through which education may lead to welfare gains for individuals – lower earnings volatility.

There are many reasons to expect that earnings volatility may be influenced by education level. Education may affect the industry or occupation chosen by the individual and, thus, may affect the labour market shocks to which they are exposed. For example, more educated individuals may work in less volatile industries such as services rather than manufacturing or construction. When searching for jobs, more educated workers may be more effective (have greater search capital) and may achieve better job matches with lower subsequent wage volatility and a

lower likelihood of layoff (Mincer 1991). More educated workers are likely to be more mobile geographically and hence able to move region in order to reduce the effects of local shocks (Machin et al 2012). Firms may be less likely to lay off more educated workers if hiring and vacancy costs are greater for them. Finally, more educated workers may be quicker to adapt to technological advances and/or have skills which are complementary to technology, and so may retain their jobs and have increased wage growth in times of structural change within the economy. While these factors tend to imply lower earnings volatility for more educated workers, there are other factors that suggest the opposite. More education typically comes with the likelihood of greater specialisation that may make the worker more exposed to specific shocks. Moreover, since the more educated typically have higher wages, they can more easily absorb wage cuts as opposed to low wage workers. The presence of the minimum wage may also lead to less volatility for those with lower education since it provides a lower bound on wages. As such, whether or not greater education lowers earnings volatility is an empirical question which we attempt to answer in this paper.

The paper proceeds as follows: section 2 describes the data, section 3 describes the measures of earnings volatility that we study, section 4 describes the estimation strategy, section 5 contains the results, and section 6 does some robustness checks. Finally, section 7 concludes.

3.2 Data

The New Earnings Survey Panel Dataset (NESPD) is a large administrative dataset covering the period 1975 to 2013. It follows a random sample of 1% of the British population whose national insurance number ends in a certain pair of digits. The survey refers to a specific week in April each year and excludes the self employed. Due to the fact that the survey uses national insurance numbers, the attrition rate is very low since if an individual temporarily drops out of the labour market, becomes unemployed or changes job, they will be picked up again once they begin working. The main advantage of the dataset, apart from the large sample sizes, is that the data are very accurate. Employers are obliged by law to fill out the employee information and thus there is less measurement error or non-response than is typically the case

in household surveys.² Also, the long period of time covered by the dataset makes it ideal for estimating cyclical effects. The data span the recessions of the early 80s, early 90s and the recent ‘Great Recession’.

Our measure of earnings is the log of weekly pay including overtime. Because of the difficult issues involved with dealing with non-participation of women in the labour market, we focus our analysis on males. We exclude those individuals whose pay was affected by absence and limit the sample to those aged between 20 and 60 to reduce selection effects that typically occur at the beginning and end of the life cycle.³ Because the compulsory schooling law changed for the 1957 cohort, we only include those born between 1947 and 1967 and, in our primary analysis, we study cohorts born between 1952 and 1962.

For survey years prior to 2004, the age variable in the survey refers to age as at January 1st. Therefore we calculate year of birth as year - age - 1. From 2004 onwards the age variable refers to age as at the time of survey which implies that assigning year of birth to be equal to year - age - 1 will only be correct roughly two thirds of the time. We contacted the UK’s Office for National Statistics (ONS) who kindly provided us with the actual year of birth of individuals from 2004 onwards. Since the first persons affected by the law change were born in September 1957, this implies that our 1957 cohort includes both treated and untreated individuals. Therefore, we drop this cohort in all our analyses. We do know, however, that all persons in our 1956 cohort were untreated and all persons in our 1958 cohort were treated.⁴

Persons born in 1962 are recorded as aged 20 in the 1983 survey year. 1983 is, therefore, the first survey year in which all persons born between 1952 and 1962 are potentially in our sample. Likewise, persons born in 1952 are aged 60 in 2013 so this is the last year in which all persons born between 1952 and 1962 are potentially in our sample. For this reason, we restrict our analysis to the 1983 to 2013 years

²One concern is that since the survey is based on payroll records it only samples those who earn enough to be above the PAYE threshold; however Devereux and Hart (2010) have shown that the exclusion of those who earn very low earnings is not important.

³Since we are interested in the effect of an extra year of school at age 15, by age 20 we expect that the complier group will mostly be in the labour market.

⁴We have information on school cohort from 2004 but not for earlier survey years. Because most of our data are pre-2004, we use calendar year cohort in our analysis. This affects the efficiency of our estimates but not consistency.

of the NESPD. If we were to use earlier survey years, some cohorts would not be present in all years and our estimates could be affected by year effects arising from periods of very high or low earnings volatility.

We deflate the weekly pay measure to 2013 prices using the Retail Price Index for April and trim the bottom and top 0.5% of earnings each year to eliminate any outliers or measurement error.⁵ In addition, we drop observations with sex or age discrepancies. Finally we drop those with hours of work less than 1 hour per week. The resulting sample is 1,140,024 observations. The unemployment rate we use refers to the claimant count rate for the April corresponding to the survey year. When conducting regional analysis we use the corresponding regional unemployment rate.

Table 3.1 displays the descriptive statistics for our broad sample that includes cohorts from 1947 to 1967 and our primary sample that includes just the 1952-62 birth cohorts. The "Law Affected" variable is 1 if the person was subject to the higher compulsory schooling age of 16 and zero otherwise.

Table 3.1: Descriptive Statistics for Males NESPD 1983-2013

| Variable | Observations | Mean | Standard Deviation |
|----------------------------|--------------|---------|--------------------|
| Cohorts = 1947-1967 | | | |
| Year | 1,104,104 | 1997.23 | 8.48 |
| Cohort | 1,104,104 | 1957.37 | 6.08 |
| Age | 1,104,104 | 38.95 | 9.57 |
| Log Weekly Pay | 1,104,104 | 6.33 | 0.514 |
| Law Affected | 1,104,104 | .536 | 0.499 |
| Cohorts = 1952-1962 | | | |
| Year | 574,498 | 1997.15 | 8.76 |
| Cohort | 574,498 | 1957.25 | 3.33 |
| Age | 574,498 | 38.99 | 9.35 |
| Log Weekly Pay | 574,498 | 6.34 | 0.512 |
| Law Affected | 574,498 | 0.537 | 0.499 |

Note: Observations refers to number of person-year observations

⁵The RPI is the only price index in the UK which covers the sample period. However, there is some concern that the method overestimates price inflation and, since 1998, there has been a new adjusted RPI called the RPIJ which uses a similar calculation method to the CPI but unlike the CPI includes housing costs and mortgages in the basket of goods.

3.3 Earnings Volatility Measures

We focus on three different measures of earnings volatility – the individual-level standard deviation of earnings, the degree of earnings cyclical, and the probability of experiencing a pay cut. Note that our focus is on variation over time in earnings for individual workers and we do not study how the cross-sectional variance of earnings differs by education level. In this section, we discuss the rationale for each measure and how exactly we implement it in our data.

Standard Deviation of Earnings: We use the standard deviation of log(weekly pay) as our primary measure of earnings uncertainty faced by men. Since education may have differing effects on earnings variability at different ages, we construct the standard deviation for each person at each age using the variation in log(weekly pay) over the five year period centred on that age. To help describe our basic approach, we start with the simple statistical model:

$$y_{it} = X_{it}\beta_t + \varepsilon_{it} \quad (3.1)$$

Here y_{it} is log(weekly pay) and X is a full set of cohort and year indicators. As these subsume age indicators, X controls for predictable life-cycle effects on earnings as well as aggregate shocks that differ across years. The error term, ε , then reflects that part of earnings that is not systematically related to cohort, age, or year. We take a particular year and then keep all observations in the 5 year window centred on it. So, for example, for 1985, we keep observations from 1983 to 1987. We then estimate the regression above on these data and estimate the standard deviation of the residual for each individual separately. That is, for each individual, we calculate the standard deviation of their earnings residual over this 5-year period. This gives us a measure of earnings volatility for that person in that year.⁶ This procedure gives us one observation for each individual for the middle year of each 5-year period and allows us to estimate the effect of education on earnings volatility at each age. We use the terms uncertainty and volatility interchangeably, however,

⁶An alternative to this approach would be to estimate the parameters of an earnings process and use it to estimate the variances of transitory and permanent components. Our approach has the advantage of not relying on the specification of some arbitrary parametric form for the income process.

we acknowledge that part of our measure of earnings volatility may not actually represent true uncertainty if individuals know in advance the variation in earnings they face.⁷

In practice, for various reasons, it is not the case that every person is in our sample in every year of each rolling 5-year panel. Therefore, we face a trade-off in that, if we require people to have valid wage observations in all 5 years, we will have a much reduced sample size and a quite selected sample. However, if we estimate the standard deviation in all feasible cases (ie where there are at least 2 observations on the individual), some standard deviations will be much more precisely estimated than others. In practice, we have taken a compromise position of requiring a valid wage observation for at least 4 years out of the 5 although we show later that the results are robust to restricting the sample to those with at least 2, 3, or 5 observations.⁸

Earnings Cyclicity: The extent to which earnings move in line with the business cycle is another important measure of earnings volatility. There is a large literature that studies the relationship between education and the degree of wage cyclicality but none of these papers attempt to estimate the causal effect of education.⁹ The papers usually estimate regressions with a wage variable as the dependent variable and some business cycle proxy such as the unemployment rate as a right hand side regressor and either look separately by education group or interact the unemployment rate with education. However, this may represent only correlations and obscure the fact that those with different levels of education may differ in other unobservable ways that affect wage cyclicality.

⁷ There has been a series of papers in the literature attempting to address this issue – separating what is known in advance from what is actual uncertainty using, for example, information on education choices (Cunha et al (2005)) and consumption (Blundell et al (2008)). We do not have data on either of these variables and therefore we do not attempt to identify how much of the variability is known in advance. As a result, we interpret our volatility measure as representing an upper bound on the amount of uncertainty faced by an individual.

⁸We have examined whether the law affects the probability of having a wage observation in our sample for at least 4 of the 5 years and found no evidence of any relationship. So, it is unlikely that selection bias is a problem when we require presence in 4 out of 5 years.

⁹The literature looking at the effects of wage cyclicality across education groups finds mixed results. On the one hand, Keane, Solon, Barsky (1991), Stockman (1983), Bils (1985) and Keane and Prasad (1991) find no difference in wage cyclicality across education groups. However, Freeman (1991), Bartik (1991), Hines et al (2001) and Hoynes et al (2012) find strong evidence that the low educated are most sensitive to business cycles while Ammermueller et al. (2009) find the opposite.

With the recent Great Recession and the associated ‘UK productivity puzzle’ there has been a renewed interest in understanding the effects of recessions on wage volatility and unemployment in Britain (Gregg et al. (2014), Blundell et al. (2014), Elsby et al (2010)).¹⁰ However, very few of these papers look at how the effects differ across education levels. One paper which does provide some evidence of differences across education is Blundell et al (2014) who show that the large and unprecedented wage cuts experienced in the Great Recession affected all education groups uniformly.¹¹

Our first measure of the business cycle is the unemployment rate in April of the survey year (the NESPD survey takes place in April). The basic idea is to first estimate the earnings cyclicity coefficient for each cohort and then to treat the earnings cyclicity of the cohort as the dependent variable in the second cohort-level step. We obtain the cyclical coefficients for each cohort by estimating the following regression for each cohort separately.

$$\Delta y_{it} = \alpha_0 + \alpha_{1c}\Delta u_t + \alpha_2 year + \varepsilon_{it}$$

Here Δy denotes the change in log earnings and Δu denotes the change in the aggregate unemployment rate while controlling for year allows for a linear trend in earnings growth.¹² We use the estimated coefficient from each cohort-specific regression α_{1c} , as our measure of the earnings cyclicity experienced by cohort c .

In order to allow for more variation in unemployment rates we add to our aggregate analysis by also exploiting variation in regional unemployment rates. We use the 12 standard statistical regions used for the UK. The empirical analysis follows exactly as before except we add region dummies in the first step and use the region unemployment rate rather than the national unemployment rate. Here r de-

¹⁰The UK productivity puzzle refers to the fact that since the onset of the recession output fell by 6% while unemployment only decreased by 2% resulting in falling output per worker.

¹¹ They find that between 2009 and 2012 real wages decreased by about 10 percent for all education groups.

¹²By taking the difference in log earnings we increase the robustness of our estimates by getting rid of unobserved heterogeneity that is fixed over time. This specification is fairly standard in the wage cyclicity literature and is used, for example, by Solon et al. (1994)

notes region.

$$\Delta y_{itr} = \alpha_0 + \alpha_{1c}\Delta u_{tr} + \alpha_2 year + region_r + \varepsilon_{itr}$$

Figure 3.1 shows the unemployment rates in each region over the sample period. While the unemployment rates typically tend to move in unison, there is some divergence in unemployment rates between regions.¹³ Once again, we use the estimated coefficient from each cohort-specific regression α_{1c} , as our measure of the earnings cyclicity experienced by cohort c .

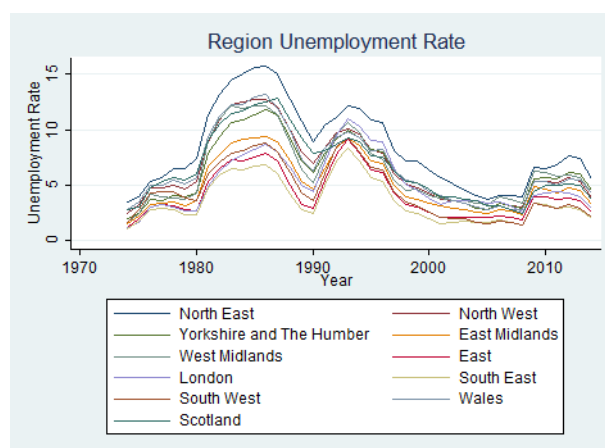


Figure 3.1: Regional Unemployment Rates

Pay Cuts: We use the prevalence of pay cuts as an additional measure of earnings volatility. Earnings generally rise over the life-cycle so pay cuts are likely to be unexpected and unpleasant for workers. Nickell et al. (2002) use occupation coding to assign workers to skill groups and find that low skilled workers are more likely to experience nominal pay cuts. But, again, we are unaware of analysis in the literature of the causal effects of education on the probability of a pay cut. These issues have become particularly relevant in the recent great recession where there has been a large surge in both real and nominal pay cuts. In keeping with our analysis of standard deviations, we measure pay cuts as occurring if the weekly pay of a

¹³ We have also looked at estimates whereby we add year dummies to the first step to allow us to control for year effects so that we are only left with the variation in unemployment rates that is derived uniquely from variation across regions and abstracts from any national business cycles. However, due to the fact that the regional rates are so highly correlated, the standard errors become very large and the estimates are not informative.

worker is lower than he received 5 years previously. We define a pay cut to be equal to 1 if the current weekly pay is less than weekly pay at period $t-4$, i.e. if at the end of the 5-year period, weekly pay is less than it was at the beginning of the period. By studying cuts over a 5-year period, we reduce the number of cuts that occur due to extremely short-term changes and so should have less noise in our measure. We focus our analysis only on real pay cuts.

3.4 Estimation Strategy

In 1973 the UK government raised the minimum school leaving age from 15 to 16. This law affected all students in England and Wales born on or after September 1st 1957 and was a follow up to the first raising of the school leaving age (RoSLA) in 1933.¹⁴ These laws have been much utilised in the literature in order to estimate the returns to additional years of education (Harmon and Walker (1995), Oreopoulos (2006), Grenet (2013), Devereux and Hart (2010), Clark and Royer (2013), Buscha and Dickson (2015)). We use a regression discontinuity design which provides very robust estimates since we only focus on changes in outcomes of those born on either side of the cut-off. The resulting estimates will produce a Local Average Treatment Effect that relates to those induced to increase their schooling due to the law i.e. the compliers (Angrist and Imbens (1994)).

Our primary approach is to estimate the model non-parametrically using a local linear regression with rectangular kernel weights (Hahn et al. (2001)).¹⁵ To focus on cohorts born close to the law change, we restrict the sample to cohorts born between 1952 and 1962 corresponding to 5 years on either side. However, we also show estimates where we use bigger bandwidths (cohorts born up to 10 years each side of the discontinuity) and other specifications such as a global polynomial approach, similar to Devereux and Hart (2010) and Oreopoulos (2006). One important issue with regression discontinuity designs is determining the best way to conduct inference. With cohort-based designs like ours, researchers often choose to cluster by cohort. However, this has been shown to be very unreliable if the

¹⁴In 2013 the school leaving age in England and Wales was raised from 16 to 17 and in 2015 it was raised to 18

¹⁵Imbens and Lemieux (2008) suggest using a simple rectangular kernel since using different weights only changes those estimates which are already sensitive to the bandwidth and thus are already invalid. Moreover the asymptotic bias is independent of the choice of weights.

number of clusters is small such as in our case where we have 10 cohorts in the local linear regressions. To avoid this issue we conduct all our analysis at the cohort level and weight by the number of observations in each cohort to control for any heteroskedasticity.¹⁶

The NESPD dataset contains very accurate earnings data, has large sample sizes, and has repeated observations on individuals that allow us to construct individual measures of earnings volatility. However, it does not contain data on month of birth. Given persons born after September 1st 1957 were subject to the new compulsory schooling law, we use a ‘donut’ style approach whereby we omit the year 1957 from the estimation.

We focus on the reduced form relationship between the law and our outcome variables:

$$Y_{it} = \theta_0 + \theta_1 Law_i + f(YOB_i) + u_{it} \quad (3.2)$$

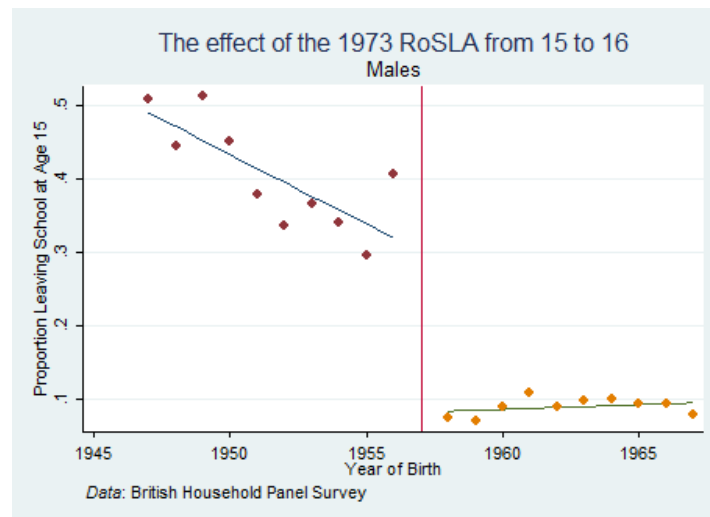
Here Y refers to the dependent variable of interest and Law is an indicator variable denoting whether the individual was born before or after 1957. The function $f(\cdot)$ represents a smooth function of year of birth. When we use local linear regression this is a linear function of year of birth that is allowed to have different slopes on each side of the discontinuity. In other specifications $f(\cdot)$ is proxied by a low order polynomial.

Our dataset does not have information on years of education. To deal with the absence of an education variable, we use a Two Sample Two Stage Least Squares approach similar to Angrist and Krueger (1992) and Devereux and Hart (2010). The basic idea is to use a separate dataset to obtain the first stage estimate of the effect of the law on education. Since we have only one instrument for education (the law change), the TS2SLS estimator is equal to the reduced form effect of the law divided by the first stage effect of the law on education. If the first stage effect is π_1 , then

$$\hat{\gamma}_1 = \frac{\hat{\theta}_1}{\hat{\pi}_1}$$

¹⁶This approach is conservative as it implies that our main regressions have only 10 observations.

Given that the UK compulsory schooling laws have been so widely studied, we follow Buscha and Dickson (2015) and use 0.30 as our estimate of the first stage. This is due to the general consensus in the literature that the law increased years of schooling by about 0.3 for males. Therefore the results that we find using the NESPD can be multiplied by 3.33 to provide the effect of an extra year of education on the outcome variable. As a robustness check we also calculated the first stage using the British Household Panel Survey (BHPS) and found a first stage effect of 0.332 with a standard error equal to 0.061 (see figure 3.2 in the Appendix for a visual representation). This is very similar to previous studies. The figure below shows the first stage effect of the compulsory schooling change on the proportion leaving school at age 15 using the BHPS.



3.5 Results

3.5.1 Log Weekly Pay

While our interest is in earnings volatility rather than in the level of earnings, we provide some context by first showing the effects of the reform on log weekly pay. Recent research (Bhuller et al. (2013), Buscha and Dickson (2015)) has emphasised the importance of variation in the return to education over the lifecycle. Of particular relevance is Buscha and Dickson (2015) who estimate the returns to the compulsory schooling law by year using the NESPD. Because there are many differences between our analysis and theirs, we show the return at each age using our

sample and methods. Table 3.2 reports estimates by year and thus tracks out returns over the life-cycle. For example, the first estimate is for 1983 when the 1957 cohort is aged 25 (and the 52-62 cohorts used in estimation are aged between 20 and 30). The final estimate is for 2013 at which point the 1957 cohort is aged 55 and the cohorts used in estimation are aged between 50 and 60. Importantly, we find that the results vary over the life cycle with the biggest return when individuals are in their early 30s. Indeed between 1991 and 1994, when the men are aged 33-36, the coefficient on the law is about 4% which scales up to a very large return to education of about 13%. The returns are lower when men are in their 20s and early 30s. Also, after the late 30s, the returns decline and the coefficient is even negative in a few of the later years.¹⁷

We also show results where we pool across ages in Table 3.2. First, we take average log(weekly pay) across all years for each cohort and use it as the dependent variable, weighting the cohort-specific means by the number of observations for that cohort so as not to induce heteroskedasticity. The resultant coefficient of 0.018 is not statistically significant but implies a return to an extra year of education of about 6%. This is similar to what other researchers have found for this law change using other data sets (Grenet (2013), Dickson (2012))¹⁸. We also split the sample in two into a younger group (when the 1957 cohort is aged 25 to 39) and an older group (when the 1957 cohort is aged 40 to 55) and report separate estimates for these two groups. Consistent with the life-cycle patterns, we find a larger effect for the younger group of 0.025. This is statistically significant at the 10% level and implies a return to education of about 8%. Figures 3.3 - 3.5 in the Appendix provide a visual impression of these estimates by plotting out average log(weekly pay) by cohort. Note that there is a clear pattern of earnings falling with cohort for younger men. This presumably arises because they are on the upward sloping part of their

¹⁷While many studies have shown life-cycle relationships between education and earnings, we are only aware of two studies that show causal estimates by detailed age. Using a very different estimation method Buscha and Dickson (2015) show a broadly similar pattern of increasing then decreasing returns using the NESPD. Our estimates are, however, in contrast to results from Norway which show that the returns are typically lower at the beginning of the life cycle and higher at the end (Bhuller et al. 2013). However the authors note in their paper that these types of estimates are likely to vary across national and institutional settings.

¹⁸Grenet (2013) uses the Quarterly Labour Force Survey from 1993 to 2004 and finds a return of 6-7% while Dickson (2012) using the British Household Panel Survey 1991 to 2006 finds a 10% return. While our estimate is lower than these, we show later that our estimates are slightly higher when we use a bigger bandwidth

life-cycle earnings profile.

It is important to note that while our sample gets older as we move to later survey years, we cannot be sure that differences in estimates by age are true effects of ageing. This is because the return to education could change across time even if the average age of sample members was time invariant. One possible reason is secular changes in the return to education that might arise because of technological change. Another is cyclical effects that could arise if education influences how earnings respond to the business cycle. We examine this directly later in the paper.

Table 3.2: Effect of the Law on Log Weekly Pay

| Year (age of 1957 cohort) | Effect | Year (age of 1957 cohort) | Effect |
|---------------------------|--------------------|---------------------------|------------------|
| 1983 (25) | 0.0179 (0.0169) | 1999 (41) | 0.007 (0.0174) |
| 1984 (26) | 0.0266 (0.0177) | 2000 (42) | 0.0251 (0.0261) |
| 1985 (27) | 0.0213 (0.0206) | 2001 (43) | 0.0165 (0.0206) |
| 1986 (28) | 0.0285 (0.0169) | 2002 (44) | 0.001 (0.0199) |
| 1987 (29) | 0.0261 (0.0211) | 2003 (45) | 0.0172 (0.0206) |
| 1988 (30) | 0.0106 (0.0229) | 2004 (46) | 0.0054 (0.0233) |
| 1989 (31) | 0.0292 (0.0166) | 2005 (47) | 0.0226 (0.0256) |
| 1990 (32) | 0.0301 (0.0182) | 2006 (48) | 0.0239 (0.0239) |
| 1991 (33) | 0.0354* (0.0181) | 2007 (49) | 0.0375 (0.0219) |
| 1992 (34) | 0.0468*** (0.0113) | 2008 (50) | 0.0255 (0.0282) |
| 1993 (35) | 0.0457* (0.0192) | 2009 (51) | 0.0317 (0.0191) |
| 1994 (36) | 0.0426** (0.0144) | 2010 (52) | -0.0143 (0.0126) |
| 1995 (37) | 0.0275* (0.0133) | 2011 (53) | 0.0104 (0.0167) |
| 1996 (38) | 0.0213 (0.0171) | 2012 (54) | 0.0052 (0.0122) |
| 1997 (39) | 0.0176 (0.01) | 2013 (55) | -0.0089 (0.0206) |
| 1998 (40) | -0.0029 (0.0183) | | |
| Young (aged less than 40) | 0.0252* (0.0123) | n=302,527 | |
| Older (aged 40+) | 0.0143 (0.0159) | n=271,971 | |
| All | 0.018 (0.0121) | n=574,498 | |

All regressions done at the cohort level weighted by the number of observations per cohort. Standard errors in parenthesis. Significance level: *** at .01, ** at .05 and * at .10

3.5.2 Standard Deviation of Earnings

Table 3.3 shows the effects of the law on the standard deviation of earnings over the life cycle. Because, we can only estimate this for the middle year in each rolling 5-year period, we have estimates by year from 1985 to 2011. What is immediately clear from the table is that most of the effect of education on earnings volatility happens at young ages with more education leading to less volatility. There are large effects of education around age 27-32 in the range of about 0.01. The average standard deviation for the 1957 cohort in this age range is 0.15. Therefore a coefficient

of -0.01 implies about a 7% reduction. When translated into our 2SLS estimate, this implies that an extra year of education decreases earnings volatility by around 0.03 or by about 20% of the mean which is quite a large effect. However, the size of the effects gets smaller as men age and many of the point estimates even become positive (albeit statistically insignificant) as men approach their 50s. One interpretation of this pattern is that more education helps people to find better and more stable job matches in their early career but the effect becomes unimportant at older ages when most individuals have found suitable job matches. Of course, as mentioned earlier, an alternative possibility is that these are time rather than age effects and that education particularly sheltered workers during the 1985-1990 period. While we can't rule out this possibility, the general pattern seems more consistent with an age rather than a cyclical effect. As in Table 3.2, we also report estimates where we increase precision by pooling across age groups. We find a statistically significant effect of -0.0065 for men aged less than 40 but no evidence of any effect for older men. Over the entire life cycle, the average effect is -0.0042 indicating that, while allowing that the effects differ over the life-cycle, on average, men experience lower earning volatility if they were subject to the law change. Scaling up by the first stage to extrapolate to the effect of an extra year of education, we find that over the whole life cycle an extra year of education leads to a lower standard deviation of earnings of about 0.012 or about 10% of the mean. Figures 3.6 - 3.8 in the Appendix provide a visual impression of these estimates. While the discontinuity is obvious for younger men, there is no obvious jump for the 1957 cohort in the older sample. Note that there is a clear pattern of earnings uncertainty rising with cohort for younger men. This arises because earnings uncertainty falls with age as workers become more settled into the labour market and the variability of earnings changes declines.

Table 3.3: Effect of the Law on the Standard Deviation of Log Earnings

| Year (age of 1957 cohort) | Effect | Year (age of 1957 cohort) | Effect |
|---------------------------|---------------------|---------------------------|--------------------|
| 1985 (27) | -0.0111** (0.0029) | 1999 (41) | -0.0055 (0.0051) |
| 1986 (28) | -0.0118 (0.0067) | 2000 (42) | -0.0068** (0.0019) |
| 1987 (29) | -0.0081 (0.0056) | 2001 (43) | -0.0042 (0.0056) |
| 1988 (30) | -0.0111*** (0.0024) | 2002 (44) | -0.0043 (0.0049) |
| 1989 (31) | -0.0102*** (0.0025) | 2003 (45) | 0.0001 (0.0056) |
| 1990 (32) | -0.0068 (0.0056) | 2004 (46) | 0.0034 (0.0058) |
| 1991 (33) | -0.005 (0.0034) | 2005 (47) | 0.001 (0.0068) |
| 1992 (34) | -0.0015 (0.0058) | 2006 (48) | -0.0036 (0.0053) |
| 1993 (35) | -0.0042 (0.0063) | 2007 (49) | -0.0088 (0.0061) |
| 1994 (36) | -0.005 (0.0041) | 2008 (50) | -0.0016 (0.0038) |
| 1995 (37) | -0.0047 (0.0047) | 2009 (51) | -0.0009 (0.0066) |
| 1996 (38) | -0.0059 (0.0034) | 2010 (52) | 0.0032 (0.0071) |
| 1997 (39) | -0.0044 (0.0034) | 2011 (53) | 0.006 (0.0049) |
| 1998 (40) | -0.006 (0.0044) | | |
| Young (aged less than 40) | -0.0065** (0.0027) | n = 197,741 | |
| Older (aged 40+) | -0.0021 (0.0028) | n = 175,943 | |
| All | -0.0042 (0.0026) | n = 373,684 | |

Standard deviations measured in the 5-year period centred around the listed year with at least 4 observations per window. Local linear regression estimates using cohorts born between 1952 and 1962 with the 1957 cohort omitted. All regressions done at the cohort level with weights equal to the number of observations per cohort. Standard errors in parenthesis. Significance level: *** at .01, ** at .05 and * at .10

In the estimates so far, we require at least 4 wage observations per person in each 5-year window. In Table 3.4, we examine whether the estimates are sensitive to using different requirements. First, looking at the sample sizes, we see that as we move from the least restrictive (only requiring 2 observations) to the most restrictive (requiring all 5), we lose over half the person-year observations. However, the estimates are quite robust – the effect of the law on the younger group varies between -0.0063 and -0.0072 and is always statistically significant. Likewise, there is never any evidence of an effect on the older group.

Table 3.4: Effect of the Law on the Standard Deviation of Earnings

| | Effect | | Sample Size |
|-------------------------------------------|-----------|----------|-------------|
| SD calculated using at least 2 obs | | | |
| Young | -0.0064** | (0.0026) | 255,679 |
| Old | -0.0017 | (0.0028) | 235,281 |
| All | -0.0041 | (0.0025) | 490,960 |
| SD calculated using at least 3 obs | | | |
| Young | -0.0063* | (0.0026) | 238,787 |
| Old | -0.002 | (0.0025) | 218,468 |
| All | -0.004 | (0.0024) | 457,255 |
| SD calculated using at least 4 obs | | | |
| Young | -0.0065** | (0.0027) | 197,741 |
| Old | -0.0021 | (0.0028) | 175,943 |
| All | -0.0042 | (0.0026) | 373,684 |
| SD calculated using at least 5 obs | | | |
| Young | -0.0072* | (0.0031) | 126,388 |
| Old | 0.0011 | (0.0025) | 111,500 |
| All | -0.0031 | (0.0026) | 237,888 |

All regressions done at the cohort level with weights equal to the number of observations per cohort. Standard deviations measured in the 5-year period centred around the listed year. Local linear regression estimates using cohorts born between 1952 and 1962 with the 1957 cohort omitted. Standard errors in parenthesis. Significance level: *** at .01, ** at .05 and * at .10

3.5.3 Earnings Cyclicalities

As mentioned earlier, we calculate an earnings cyclicalities coefficient for each cohort. To estimate the effect of the law on earnings cyclicalities, we do a cohort-level regression where we regress the cyclicalities coefficient for each cohort on the law using the local linear specification. As before, in this second step, the regression is weighted by the number of observations in each cohort. The estimates are reported in Table 3.5.

Because we need a large time dimension to reliably estimate the earnings cyclicalities coefficients, we cannot study how the effect of the law on earnings cyclicalities varies at each individual age. We report estimates for the full set of years (1983-2013) and also split the sample by age group (younger versus older) as before. However, the standard errors become very high for the older group so this estimate is not very informative. As is the case for our other outcomes, the effect for the younger group is larger than that for workers as a whole and suggests that more education reduces the earnings cyclicalities experienced by men. In Table 3.5 we see that the effect of the law on earnings cyclicalities is about .0039 overall and 0.0044 for the younger group and the estimate is very significant.

Table 3.5: Effect of the Law on Earnings Cyclicalities

| Sample | Effect | Sample Size |
|---------------------------|--------------------|-------------|
| Young (aged less than 40) | 0.0044*** (0.0008) | 223,974 |
| Older (40+) | 0.0002 (0.0056) | 202,412 |
| All | 0.0039*** (0.0008) | 441,890 |

Local linear regression estimates using cohorts born between 1952 and 1962 with the 1957 cohort omitted. All regressions done at the cohort level with weights equal to the number of observations per cohort. Standard errors in parenthesis. Significance level: *** at .01, ** at .05 and * at .10

To get a sense of the magnitudes, it is helpful to look at figures 3.9 - 3.11. The pictures show the estimated earnings cyclicalities coefficient by cohort. For young men, this is always negative, reflecting the fact that earnings are mildly procyclical with estimates ranging from close to zero to -0.008. These suggest that a one point increase in the unemployment rate reduces earnings by between 0 and 1%. There is clear jump for the 1957 cohort with the levels of cyclicalities being lower (in absolute terms) after the law change than before. This is despite the fact that earnings

cyclicality is decreasing with age for our sample.

In practice, we estimate earnings cyclicality over the 1983-1997 period when studying the younger groups and the 1998-2013 period for the older groups. Therefore, the greater effect of the law on earnings cyclicality in the younger group could reflect the fact that education had less protective effect during the Great Recession than it had in the recession of the 1990s. This is plausible as the Great Recession had disproportionate impacts on financial industries that employ a lot of people with higher education. It is difficult to distinguish this effect from the possibility that education has less impact on the cyclical swings experienced by older people. However, our previous estimates of decreased standard deviation of earnings at younger ages would tend to support the age rather than time based explanation.¹⁹

3.5.3.1 Regional shocks

Table 3.6: Effect of the Law on Regional Earnings Cyclicality

| Sample | Effect | Sample Size |
|---------------------------|--------------------|-------------|
| Young (aged less than 40) | 0.0038*** (0.0007) | 223,974 |
| Older (40+) | 0.0007 (0.0053) | 202,412 |
| All | 0.0034** (0.001) | 441,890 |

The cyclical indicator is the contemporaneous regional unemployment rate for April. All regressions done at the cohort level with weights equal to the number of observations per cohort. Standard errors in parenthesis. Significance level: *** at .01, ** at .05 and * at .10

Table 3.6 shows the regional cyclicality estimates. The results are similar to the estimates using the national unemployment rate both in terms of the coefficient estimates and standard errors. This is not surprising since Figure 3.1 highlights that unemployment rates tend to move in parallel across regions. However, it is reassuring that, once we look within region and so control for any regional effects, the negative effect of the law change on earnings cyclicality remains.

3.5.4 Pay cuts

Table 3.7 shows the effect of the law on the probability of a real pay cut in each 5-year period centred on the reported year. We define the pay cut variable to be equal to 1 if the real weekly pay at the end of the 5-year period is less than what it was

¹⁹Much research has found that younger people are more sensitive to the business cycle, for example, Hoynes et al (2012) and Elsby et al (2010).

at the beginning of the period. So, for example, for the year 1987 in Table 3.7, the estimate shows the effect of the law change on the probability that the real weekly pay in 1989 is lower than that in 1985. There is no clear pattern in the results with education leading to a lower probability of receiving a cut in some years, for example, in 1989, 1992 and 2004 but at other times leading to an increased likelihood of a pay cut. The effect of education on lowering the probability of receiving a pay cut in 1992 is consistent with the fact that the 1990's recession had large adverse effects on those with the least education. It is not clear why the effects in 2004 are so large but it is something that we think deserves further investigation. Overall the effect of the law is to reduce the probability of a pay cut by 1.1 percentage points which is statistically significant at the 5% level. The average probability of a pay cut for the 1957 cohort over the sample period is 0.342. Therefore a coefficient of -0.011 implies about a 3.2% reduction. When translated into our 2SLS estimate, this implies that an extra year of education reduces the probability of receiving a pay cut by about 10.6 percent. However, it is important to keep in mind that there is a lot of heterogeneity across years with positive coefficients in some years and negative ones in others.

The increased probability of the higher educated receiving a pay cut between 2007 and 2011 (as shown by the 2009 estimate) is consistent with evidence that the Great Recession had a bigger impact on those with more education, in particular bankers and those working in the financial industry.²⁰ Therefore, it appears that, in the Great Recession, more education led to a higher probability of taking a real earnings cut.²¹ This highlights how the effects of education on pay cuts can be very heterogeneous and suggests that the average effects over the entire period should be treated with caution.

²⁰This is not surprising given that in 2009 financial institutions reacted to the Lehman brother's scandal by cutting pay.

²¹ We have looked at year-by-year wage changes to examine this further and found that those affected by the law change were more likely to experience a pay cut between 2008 and 2009 but there is no evidence of adverse effects of education in the years after 2009.

Table 3.7: Effect of the Law on the Probability of a Real Weekly Pay Cut (over 5-Year Period)

| Year (age of 1957 cohort) | Effect | Year (age of 1957 cohort) | Effect |
|---------------------------|-------------------|---------------------------|---------------------|
| 1985 (27) | -0.0154 (0.0107) | 1999 (41) | -0.0207 (0.012) |
| 1986 (28) | -0.0253 (0.0132) | 2000 (42) | -0.025 (0.0214) |
| 1987 (29) | -0.0001 (0.0169) | 2001 (43) | 0.0014 (0.0193) |
| 1988 (30) | -0.0123 (0.0103) | 2002 (44) | 0.0088 (0.0147) |
| 1989 (31) | -0.0196* (0.0096) | 2003 (45) | -0.0164 (0.03) |
| 1990 (32) | -0.0291 (0.0215) | 2004 (46) | -0.0623*** (0.0164) |
| 1991 (33) | -0.0001 (0.0158) | 2005 (47) | -0.0117 (0.0201) |
| 1992 (34) | -0.03* (0.0143) | 2006 (48) | -0.045 (0.0332) |
| 1993 (35) | -0.0047 (0.0154) | 2007 (49) | -0.0074 (0.018) |
| 1994 (36) | 0.0172* (0.0077) | 2008 (50) | -0.0004 (0.0159) |
| 1995 (37) | 0.038* (0.0179) | 2009 (51) | 0.043* (0.0197) |
| 1996 (38) | 0.0256 (0.022) | 2010 (52) | 0.0275 (0.0214) |
| 1997 (39) | -0.0116 (0.0099) | 2011 (53) | 0.0167 (0.015) |
| 1998 (40) | -0.0073 (0.0165) | | |
| Young (aged less than 40) | -0.0082* (0.0037) | n = 157,710 | |
| Older (aged 40+) | -0.0094 (0.0063) | n = 195,484 | |
| All | -0.011** (0.0045) | n = 353,194 | |

Local linear regression estimates using cohorts born between 1952 and 1962 with the 1957 cohort omitted. All regressions done at the cohort with weights equal to the number of observations per cohort. Standard errors in parenthesis. Significance level: *** at .01, ** at .05 and * at .10

3.6 Robustness Checks

In this section we perform some robustness checks for each of our outcome measures. The estimates are reported in the Appendix.

3.6.1 Bandwidth and Model Specification

First, we look at the effects of changing the bandwidth and the model specification. To compare with our baseline estimates that do local linear regression using a bandwidth of 5, we examine the effect of bandwidths of 7 and 10 using a range of different specifications. Since the fit can vary greatly between specifications, we use the Akaike Information Criterion (AIC) to pick the model that fits best in each case. With the bandwidth of 5, the standard errors get very high when we use high order polynomials. Therefore, for this bandwidth, we only report a global quadratic in addition to the local linear specification. For bandwidths 7 and 10, we report local quadratic estimates that allow a different quadratic function each side of the discontinuity and also global quadratic, cubic, and quartic specifications. In each case, the estimate favoured by the AIC criterion is reported in bold.

The estimates are in Tables 3.8 to 3.12. In general, we find that, once one

chooses the specification with the best fit for each bandwidth, the estimates are quite robust to specification. One finding is that the effect of the law on $\log(\text{earnings})$ is stronger with the bandwidth of 10 – here we find statistically significant estimates of about 3% for the young and 2% for the older men. On average the effect of the law is about 2% which implies a return to education of about 7%. The estimates for the standard deviation, for earnings cyclicalities, and for pay cuts are all similar across bandwidths. We conclude that our findings are quite robust to the choice of bandwidth or specification.

3.6.2 Excluding Scotland

We also analyse the effect of excluding people who live in Scotland as the law only affected those born in England and Wales. Because we don't have information on country of birth, we cannot tell whether persons living in Scotland were actually born there or in England or Wales. However, removing Scottish residents still constitutes a useful robustness check. The estimates are in Table 3.13. Once again, the estimates are very similar to our baseline results.

3.6.3 Standard Errors Clustered at the Individual Level

We have conducted all our analysis at the cohort level and weighted by the number of observations in each cohort to control for any heteroskedasticity. An alternative approach to inference is not to group by cohort but to instead treat deviation from the local linear or polynomial fit as specification error and report robust standard errors (Chamberlain, 1994). In situations where we pool years and so have repeated observations on individuals, we implement this method by clustering by individual. As a check on our earlier estimates, we report these standard errors in Tables 3.14 to 3.16. Given that earnings cyclicalities is measured at the cohort level, it is not possible to generate standard errors at the individual level for this outcome. We find that the level of statistical significance is generally similar with this approach compared to our earlier approach. All our findings are robust to using this alternative method of inference.

3.7 Conclusion

This paper finds that education leads to a decrease in lifetime earnings volatility, particularly for younger men. We look at the effects of education on earnings uncertainty via the effects on earnings volatility, real pay cuts and earnings cyclical-ity. Across all three mechanisms the results point towards benefits from education through sheltering one from the adverse effects of earnings shocks. The effects vary over the life cycle with education leading to benefits in terms of reduced earnings volatility and cyclical-ity at younger ages but with no discernible benefits for persons aged over 40. Our findings for real wage cuts are quite heterogenous with the effects varying in magnitude and sign across different years. On average, however, more educated workers appear less likely to experience real wage cuts. The estimates are robust across many specifications. Overall, our paper identifies another, and largely unexplored, avenue through which education leads to positive rewards in the labour market.

3.8 References

Angrist, Joshua D., and Alan B. Krueger (1992). The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples. *Journal of the American Statistical Association* 87 (418) 32836.

Ammermueller, A., A. Kuckulenz and T. Zwick (2009). "Aggregate unemployment decreases individual returns to education" *Economics of Education Review*, 28 (2009), pp. 217226

Banks, James and Blundell, Richard and Brugiavini, Agar, (2001). "Risk Pooling, Precautionary Saving and Consumption Growth," *Review of Economic Studies*, Wiley Blackwell, vol. 68(4), pages 757-79, October.

Bils, Mark J. (1985), Real Wages over the Business Cycle: Evidence from Panel Data, *Journal of Political Economy* 93, 666-689.

Blundell, Richard, Howard Reed and Thomas M. Stoker (2003), "Interpreting Aggregate Wage Growth: The Role of Labor Market Participation". *The American Economic Review*, Vol. 93, No. 4. (Nov., 2003), pp. 1114-1131.

Blundell, Richard, Luigi Pistaferri, and Ian Preston (2008). "Consumption Inequality and Partial Insurance." *American Economic Review*, 98(5): 1887-1921.

Blundell, Richard, Claire Crawford and Wenchao Jin (2014). "What Can Wages and Employment Tell Us about the UK's Productivity Puzzle?". *The Economic Journal*. Volume 124, Issue 576, pages 377407, May 2014

Blundell, Richard, Michael Graber and Magne Mogstad (2015). "Labor Income Dynamics and the Insurance from Taxes, Transfers, and the Family". *Journal of Public Economics*. Volume 127, July 2015, Pages 5873

Buscha, Franz and Matt Dickson. (2015) "The wage returns to education over the life-cycle: Heterogeneity and the role of experience", 2015. Working Paper.

Chamberlain, G., (1994). Quantile Regression, Censoring, and the Structure of Wages. In: Sims, C.A. (Ed.), *Advances in Econometrics, Sixth World Congress*, vol. 1. Cambridge University Press, Cambridge.

Chen, Stacey. (2008). "Estimating the Variance of Wages in the Presence of Selection and Unobserved Heterogeneity". *Review of Economics and Statistics* 2008 90:2, 275-289

Cunha, Flavio, James Heckman and Salvador Navarro (2005), "Separating Uncertainty from Heterogeneity in Life Cycle Earnings", *Oxford Economic Papers*, 57(2), 191-261.

Clark, Damon, and Heather Royer (2013), "The Effect of Education on Adult Mortality and Health: Evidence from Britain." *American Economic Review*, 103(6): 2087-2120. (October 2013)

Devereux, Paul J. and Robert A. Hart (2010). "Forced to be Rich? Returns to Compulsory Schooling in Britain". *The Economic Journal*. Volume 120, Issue 549, pages 1345-1364, (December 2010)

Dickson, M. (2013), *The Causal Effect of Education on Wages Revisited*. *Oxford Bulletin of Economics and Statistics*, 75: 477-498.

Elsby, Michael W., Bart Hobijn, Aysegul Sahin (2010), "The Labor Market in the Great Recession". Published: *Brookings Papers on Economic Activity, Economic Studies Program*, The Brookings Institution, vol. 41(1 (Spring)), pages 1-69

French, Eric. (2005), The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour. *The Review of Economic Studies* 72.2: 395-427.

Gregg, P., Machin, S. and Fernandez-Salgado, M. (2014), Real Wages and Unemployment in the Big Squeeze. *The Economic Journal*, 124: 408-432

Grenet, Julien. (2013), "Is Extending Compulsory Schooling Alone Enough to Raise Earnings? Evidence from French and British Compulsory Schooling Laws". *The Scandinavian Journal of Economics*. Volume 115, Issue 1, pages 176-210, January 2013

Hahn, J., Todd, P., and Van der Klaauw, W. (2001), Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design. *Econometrica*, 69(1), 201-209.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante. (2014). "Consumption and Labor Supply with Partial Insurance: An Analytical Framework." *American Economic Review*, 104(7): 2075-2126.

Hines, James. R., Hilary Hoynes and Alan. B Krueger, (2001) "Another Look at Whether a Rising Tide Lifts All Boats" 2001. NBER Working Paper No 8412. August

Hoynes, Hilary. (2002). "The Employment, Earnings, and Income of Less Skilled Workers Over the Business Cycle". Published: Blank, Rebecca and David Card (eds.) *Finding Jobs: Work and Welfare Reform*. Russell Sage Foundation.

Imbens, G. W., & Angrist, J. D.. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica*, 62(2), 467-475.

Imbens, Guido W., and Thomas Lemieux. (2008). "Regression discontinuity

designs: A guide to practice". *Journal of Econometrics*, Volume 142, Issue 2, February 2008, Pages 615-635

Liu, Kai, Magne Mogstad and Kjell Salvanes. (2015). "Education, Earnings Dynamics, and Life-Cycle Inequality". Working Paper

Low, Hamish, Costas Meghir, and Luigi Pistaferri. (2010). "Wage Risk and Employment Risk over the Life Cycle." *American Economic Review*, 100(4): 1432-67.

Machin, S., Salvanes, K. G. and Pelkonen, P. (2012), Education and Mobility. *Journal of the European Economic Association*, 10: 417-450.

Meghir, C. and Pistaferri, L. (2004), Income Variance Dynamics and Heterogeneity. *Econometrica*, 72: 132.

Nickell, S. , Jones, P. and Quintini, G. (2002), A Picture of Job Insecurity Facing British Men. *The Economic Journal*, 112: 127.

Oreopoulos, Philip. (2006). "Estimating Average and Local Average Treatment Effects of Education when Compulsory Schooling Laws Really Matter." *American Economic Review*, 96(1): 152-175.

Oreopoulos, Philip, and Kjell G. Salvanes. (2011). "Priceless: The Nonpecuniary Benefits of Schooling." *Journal of Economic Perspectives*, 25(1): 159-84.

Mincer, Jacob. (1991). "Education and Unemployment". NBER Working Paper No 3838. September.

Keane, Michael, and Anwar Prasad (1993), Skill Levels and the Cyclical Variability of Employment, Hours, and Wages, *International Monetary Fund Staff Papers* 40 (1993), 711-743.

Solon, G., Barsky, R., & Parker, J. A.. (1994). Measuring the Cyclicalities of Real Wages: How Important is Composition Bias. *The Quarterly Journal of Economics*, 109(1), 125

3.9 Appendix

First Stage Effect of the 1973 Schooling Reform

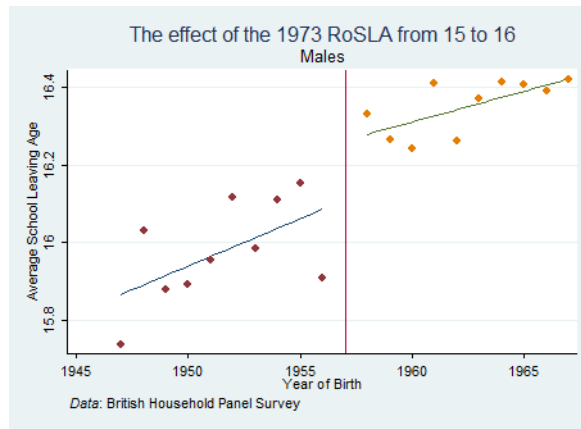


Figure 3.2: First stage effect using BHPS

Log Weekly Pay

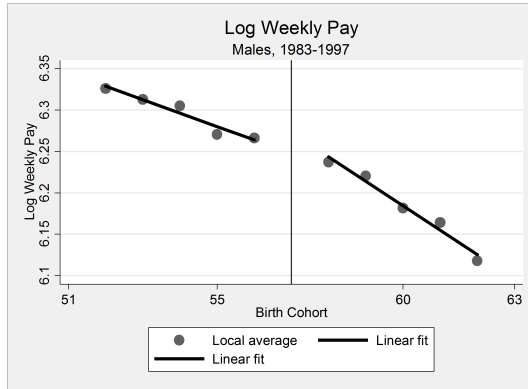


Figure 3.3: Young

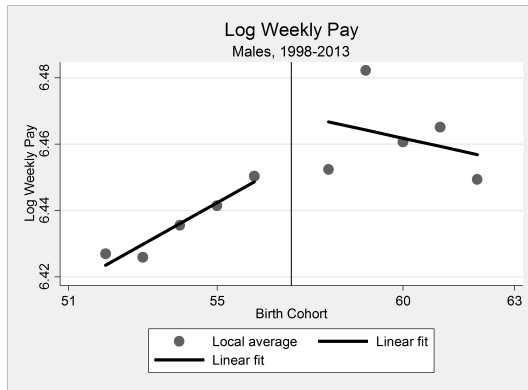


Figure 3.4: Older

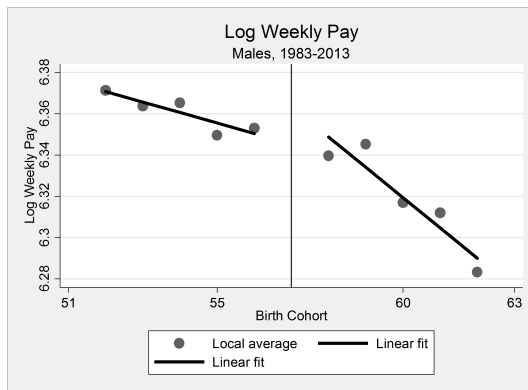


Figure 3.5: All

Standard Deviation of Log Weekly Pay

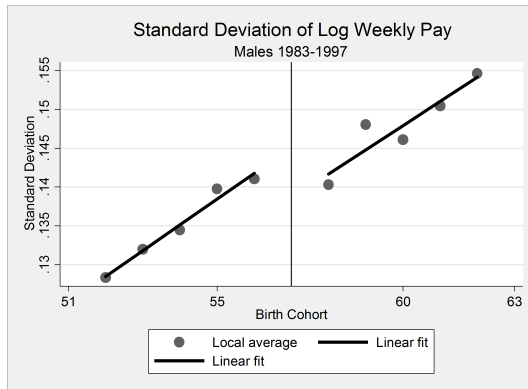


Figure 3.6: Young

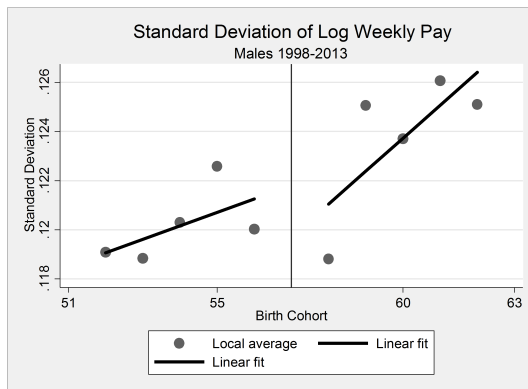


Figure 3.7: Older

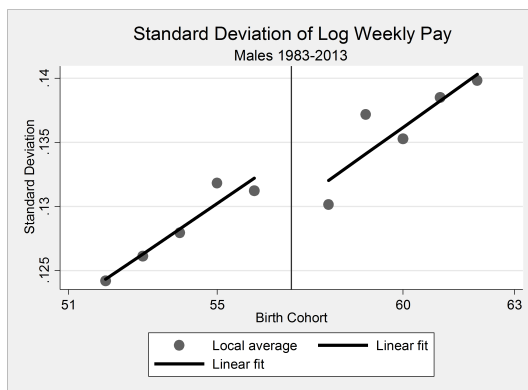


Figure 3.8: All

Earnings Cyclicality

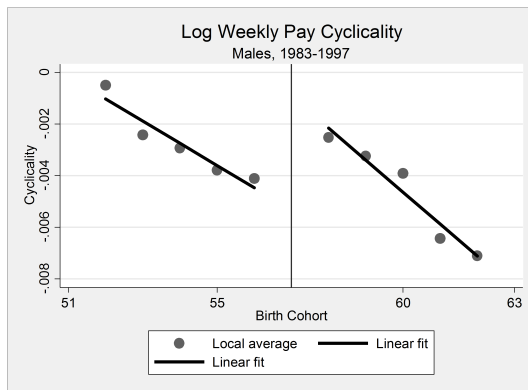


Figure 3.9: Young

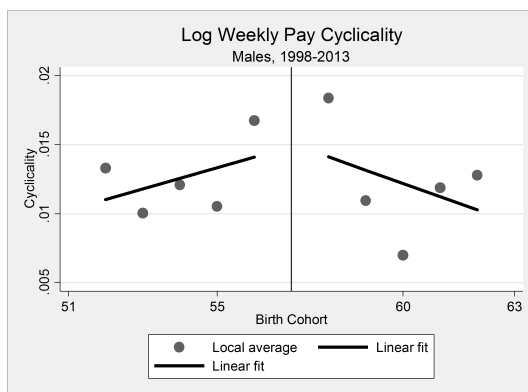


Figure 3.10: Older

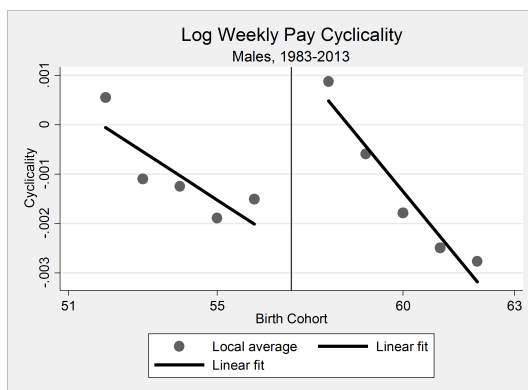


Figure 3.11: All

Regional Earnings Cyclicity

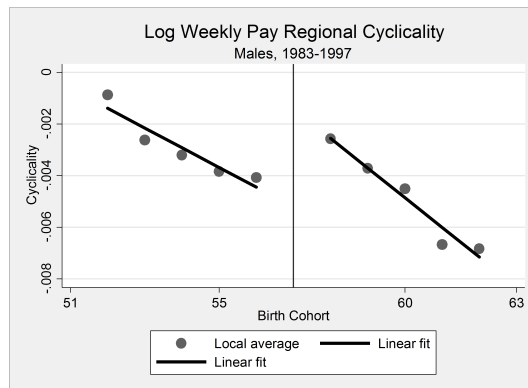


Figure 3.12: Young

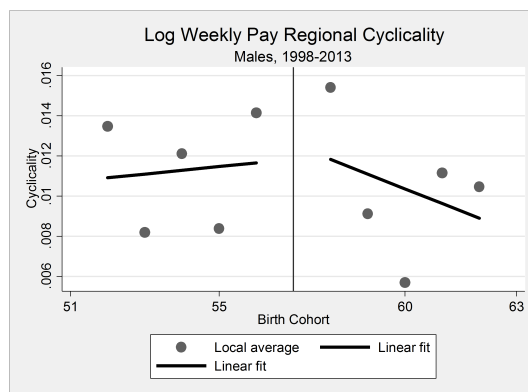


Figure 3.13: Older

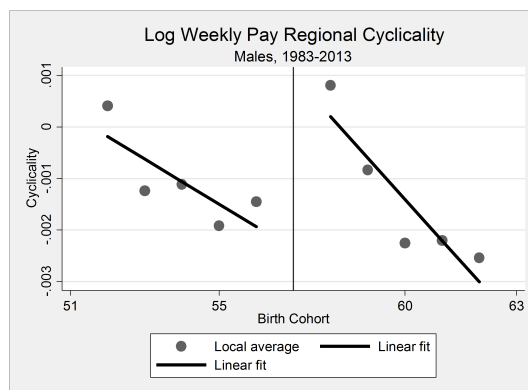


Figure 3.14: All

Real Pay Cuts over 5-Year Period

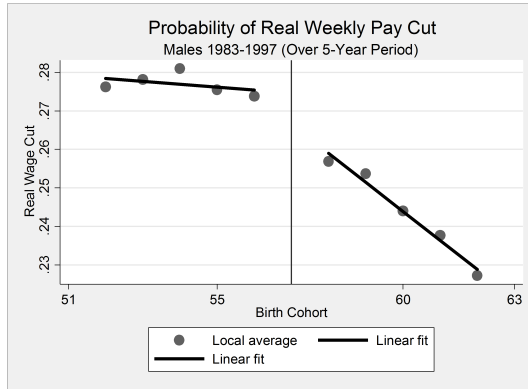


Figure 3.15: Young

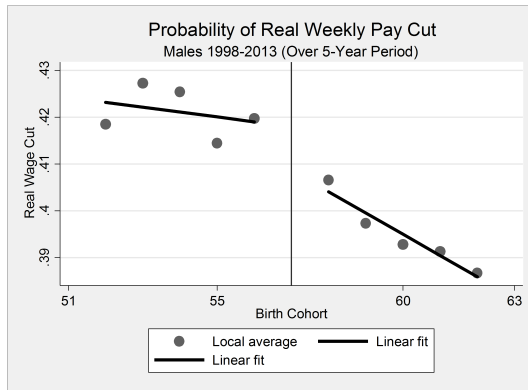


Figure 3.16: Older

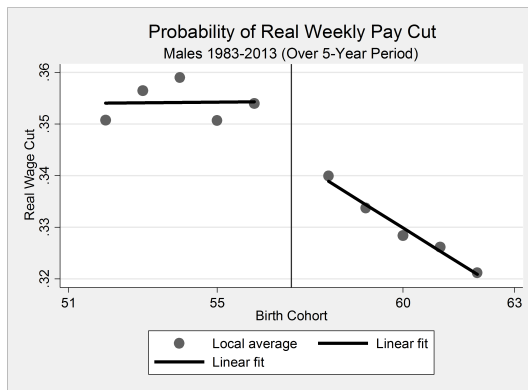


Figure 3.17: All

Varying Bandwidth and Specification

Table 3.8: Effect of the Law on Log Weekly Pay

| | Young | | Old | | All | |
|------------------------------|-----------------|----------|-----------------|----------|---------------|----------|
| Bandwidth = 5 | | | | | | |
| Linear (Different Slopes) | 0.0252* | (0.0123) | 0.0143 | (0.0159) | 0.018 | (0.0121) |
| Global Quadratic | 0.0251* | (0.0112) | 0.0139 | (0.0151) | 0.0179 | (0.0112) |
| <i>Sample Size</i> | 302,527 | | 271,971 | | 574,498 | |
| Bandwidth = 7 | | | | | | |
| Linear (Different Slopes) | 0.0015 | (0.0144) | 0.021* | (0.0112) | 0.0037 | (0.0121) |
| Quadratic (Different Slopes) | 0.0456* | (0.0219) | -0.0004 | (0.0211) | 0.0271 | (0.021) |
| Global Quadratic | 0.0016 | (0.0153) | 0.0209* | (0.0115) | 0.0038 | (0.0125) |
| Global Cubic | 0.0377* | (0.0185) | 0.0092 | (0.0175) | 0.0254 | (0.0176) |
| Global Quartic | 0.0369** | (0.0159) | 0.0074 | (0.0174) | 0.0237 | (0.0155) |
| <i>Sample Size</i> | 417,611 | | 377,869 | | 795,480 | |
| Bandwidth = 10 | | | | | | |
| Linear (Different Slopes) | -0.019 | (0.0121) | 0.0234** | (0.0098) | -0.0128 | (0.0101) |
| Quadratic (Different Slopes) | 0.0332** | (0.014) | 0.0103 | (0.0157) | 0.0233 | (0.0137) |
| Global Quadratic | -0.0185 | (0.0127) | 0.0225** | (0.0089) | -0.0127 | (0.0103) |
| Global Cubic | 0.0193 | (0.0135) | 0.0138 | (0.0129) | 0.0148 | (0.0121) |
| Global Quartic | 0.0198 | (0.013) | 0.0149 | (0.0132) | 0.0144 | (0.0123) |
| <i>Sample Size</i> | 578,171 | | 525,933 | | 1,104,104 | |

All regressions done at the cohort level with weights equal to the number of observations per cohort. Standard errors in parenthesis. Figures in bold denote the preferred model according to the Akaike Information Criterion. Significance level: *** at .01, ** at .05 and * at .10

Table 3.9: Effect of the Law on the Standard Deviation of Weekly Pay

| | Young | | Old | | All | |
|------------------------------|------------------|----------|----------------|----------|----------------|----------|
| Bandwidth = 5 | | | | | | |
| Linear (Different Slopes) | -0.0065** | (0.0027) | -0.0021 | (0.0028) | -0.0042 | (0.0026) |
| Global Quadratic | -0.0065** | (0.0027) | -0.0021 | (0.0029) | -0.0042 | (0.0026) |
| <i>Sample Size</i> | 197,741 | | 175,943 | | 373,684 | |
| Bandwidth = 7 | | | | | | |
| Linear (Different Slopes) | -0.0035 | (0.0026) | -0.0025 | (0.0022) | -0.0025 | (0.0022) |
| Quadratic (Different Slopes) | -0.0086* | (0.0041) | -0.0016 | (0.0045) | -0.0055 | (0.004) |
| Global Quadratic | -0.0035 | (0.0025) | -0.0025 | (0.0022) | -0.0024 | (0.0021) |
| Global Cubic | -0.0072* | (0.0036) | -0.0014 | (0.0035) | -0.0045 | (0.0032) |
| Global Quartic | -0.0073** | (0.0031) | -0.0016 | (0.0036) | -0.0046 | (0.003) |
| <i>Sample Size</i> | 273,186 | | 244,299 | | 517,485 | |
| Bandwidth = 10 | | | | | | |
| Linear (Different Slopes) | 0.0001 | (0.0022) | -0.0018 | (0.0022) | 0.0005 | (0.0021) |
| Quadratic (Different Slopes) | -0.0074** | (0.0032) | -0.0016 | (0.0041) | -0.0049 | (0.0034) |
| Global Quadratic | -0.0001 | (0.0022) | -0.0017 | (0.0023) | 0.0004 | (0.0021) |
| Global Cubic | -0.0054* | (0.0027) | -0.0014 | (0.0034) | -0.0033 | (0.0028) |
| Global Quartic | -0.0052* | (0.0028) | -0.0013 | (0.0035) | -0.0033 | (0.0029) |
| <i>Sample Size</i> | 378,017 | | 337,411 | | 715,428 | |

All regressions done at the cohort level with weights equal to the number of observations per cohort. Year of birth equal to 1957 is omitted. Standard errors in parenthesis. Figures in bold denote the preferred model according to the Akaike Information Criterion. Significance level: *** at .01, ** at .05 and * at .10

Table 3.10: Effect of the Law on Weekly Pay Cyclicity

| | Young | | Old | | All | |
|------------------------------|------------------|----------|----------------|----------|------------------|----------|
| Bandwidth = 5 | | | | | | |
| Linear (Different Slopes) | 0.0044*** | (0.0008) | 0.0002 | (0.0056) | 0.0039*** | (0.0008) |
| Global Quadratic | 0.0044*** | (0.0008) | 0.0003 | (0.0059) | 0.0039*** | (0.0009) |
| <i>Sample Size</i> | 223,974 | | 202,412 | | 441,890 | |
| Bandwidth = 7 | | | | | | |
| Linear (Different Slopes) | 0.0033*** | (0.0008) | 0.0035 | (0.0052) | 0.0032*** | (0.0006) |
| Quadratic (Different Slopes) | 0.0046** | (0.0014) | -0.0003 | (0.0077) | 0.0045*** | (0.001) |
| Global Quadratic | 0.0033*** | (0.0007) | 0.0038 | (0.0049) | 0.0032*** | (0.0006) |
| Global Cubic | 0.0044*** | (0.0011) | 0.0017 | (0.0078) | 0.0042*** | (0.0009) |
| Global Quartic | 0.0044*** | (0.0011) | 0.0009 | (0.0077) | 0.0042*** | (0.001) |
| <i>Sample Size</i> | 308,167 | | 280,359 | | 610,332 | |
| Bandwidth = 10 | | | | | | |
| Linear (Different Slopes) | 0.0008 | (0.0012) | 0.0043 | (0.0049) | 0.0008 | (0.0012) |
| Quadratic (Different Slopes) | 0.006*** | (0.0012) | -0.0017 | (0.0083) | 0.0059*** | (0.0011) |
| Global Quadratic | 0.0008 | (0.0012) | 0.0045 | (0.0048) | 0.0008 | (0.0012) |
| Global Cubic | 0.0047*** | (0.0012) | -0.0003 | (0.0069) | 0.0048*** | (0.0011) |
| Global Quartic | 0.0049*** | (0.0008) | -0.0008 | (0.0071) | 0.0047*** | (0.0008) |
| <i>Sample Size</i> | 425,838 | | 387,590 | | 844,682 | |

The contemporaneous national rate of unemployment for April is used as the cyclical indicator. All regressions done at the cohort level with weights equal to the number of observations per cohort. Standard errors in parenthesis. Figures in bold denote the preferred model according to the Akaike Information Criterion. Significance level: *** at .01, ** at .05 and * at .10

Table 3.11: Effect of the Law on Regional Weekly Pay Cyclicity

| | Young | | Old | | All | |
|------------------------------|------------------|----------|----------------|----------|------------------|----------|
| Linear (Different Slopes) | 0.0038*** | (0.0007) | 0.0007 | (0.0053) | 0.0034** | (0.001) |
| Global Quadratic | 0.0038*** | (0.0008) | 0.0008 | (0.0054) | 0.0034*** | (0.001) |
| <i>Sample Size</i> | 223,974 | | 202,412 | | 441,890 | |
| Bandwidth = 7 | | | | | | |
| Linear (Different Slopes) | 0.003*** | (0.0006) | 0.0034 | (0.0046) | 0.003*** | (0.0007) |
| Quadratic (Different Slopes) | 0.004*** | (0.0012) | -0.0003 | 0.0074 | 0.0038** | (0.0012) |
| Global Quadratic | 0.003*** | (0.0006) | 0.0036 | (0.0043) | 0.003*** | (0.0007) |
| Global Cubic | 0.0039*** | (0.0009) | 0.0013 | (0.0068) | 0.0036*** | (0.0011) |
| Global Quartic | 0.0039*** | (0.001) | 0.0008 | (0.007) | 0.0036** | (0.0011) |
| <i>Sample Size</i> | 308,167 | | 280,359 | | 610,332 | |
| Bandwidth = 10 | | | | | | |
| Linear (Different Slopes) | 0.0007 | (0.0011) | 0.0024 | (0.0038) | 0.0007 | (0.0011) |
| Quadratic (Different Slopes) | 0.0055*** | (0.0011) | 0.0003 | (0.0069) | 0.0053*** | (0.0011) |
| Global Quadratic | 0.0007 | (0.0011) | 0.0013 | (0.0025) | 0.0008 | (0.0011) |
| Global Cubic | 0.0044*** | (0.0011) | 0.0006 | (0.0056) | 0.0044*** | (0.001) |
| Global Quartic | 0.0045*** | (0.0007) | 0.0006 | (0.0058) | 0.0043*** | (0.0009) |
| <i>Sample Size</i> | 425,838 | | 387,590 | | 844,682 | |

The contemporaneous regional rate of unemployment for April is used as the cyclical indicator. All regressions done at the cohort level with weights equal to the number of observations per cohort. Standard errors in parenthesis. Figures in bold denote the preferred model according to the Akaike Information Criterion. Significance level: *** at .01, ** at .05 and * at .10

Table 3.12: Effect of the Law on Probability of Real Weekly Pay Cut over 5-Year Period

| | Young | | Old | | All | |
|------------------------------|------------------|----------|-----------------|----------|-------------------|----------|
| Bandwidth = 5 | | | | | | |
| Linear (Different Slopes) | -0.0082* | (0.0037) | -0.0093 | (0.0063) | -0.011** | (0.0045) |
| Global Quadratic | -0.0082** | (0.0025) | -0.0094 | (0.0064) | -0.011** | (0.0043) |
| <i>Sample Size</i> | 157,710 | | 195,484 | | 353,194 | |
| Bandwidth = 7 | | | | | | |
| Linear (Different Slopes) | -0.0132* | (0.0065) | -0.0196** | (0.0077) | -0.0218*** | (0.0059) |
| Quadratic (Different Slopes) | -0.0078 | (0.0118) | 0.0063 | (0.0096) | 0.0009 | (0.0074) |
| Global Quadratic | -0.0132* | (0.0069) | -0.0199** | (0.0072) | -0.0218*** | (0.0061) |
| Global Cubic | -0.01 | (0.011) | -0.0005 | (0.0076) | -0.0049 | (0.0063) |
| Global Quartic | -0.01 | (0.0088) | -0.0001 | (0.0078) | -0.0053 | (0.0059) |
| <i>Sample Size</i> | 215,787 | | 270,848 | | 486,635 | |
| Bandwidth = 10 | | | | | | |
| Linear (Different Slopes) | -0.009* | (0.005) | -0.0188** | (0.0067) | -0.0225*** | (0.0041) |
| Quadratic (Different Slopes) | -0.0111 | (0.0082) | -0.0138 | (0.0108) | -0.0151* | (0.0072) |
| Global Quadratic | -0.0092* | (0.0052) | -0.0194*** | (0.006) | -0.0224*** | (0.0045) |
| Global Cubic | -0.012 | (0.0077) | -0.0167* | (0.009) | -0.0182** | (0.0066) |
| Global Quartic | -0.0112 | (0.0076) | -0.0176* | (0.009) | -0.0184*** | (0.0062) |
| <i>Sample Size</i> | 294,567 | | 374,496 | | 669,063 | |

All regressions done at the cohort level weighted by the number of observations per cohort. Standard errors in parenthesis. Figures in bold denote the preferred model according to the Akaike Information Criterion. Significance level: *** at .01, ** at .05 and * at .10

Effect of Omitting Scottish Residents

Table 3.13: Effect of the Law after Omitting Scottish Residents

| Sample | Baseline | | n | Omit Scotland | | n |
|------------------------------------|-----------|----------|---------|---------------|----------|---------|
| <i>Log Weekly Pay</i> | | | | | | |
| Young (aged 27 to 39) | 0.0252* | (0.0123) | 302,527 | 0.0183 | (0.0123) | 274,430 |
| Older (40+) | 0.0143 | (0.0159) | 271,971 | 0.0115 | (0.0153) | 246,772 |
| All | 0.018 | (0.0121) | 574,498 | 0.0129 | (0.0115) | 521,202 |
| <i>Standard Deviation</i> | | | | | | |
| Young (aged 27 to 39) | -0.0065** | (0.0027) | 197,741 | -0.0075** | (0.0030) | 179,648 |
| Older (40+) | -0.0021 | (0.0028) | 175,943 | -0.0008 | (0.0032) | 159,325 |
| All | -0.0042 | (0.0026) | 373,684 | -0.0041 | (0.003) | 338,973 |
| <i>Earnings Cyclicity</i> | | | | | | |
| Young (aged 27 to 39) | 0.0044*** | (0.0008) | 223,974 | 0.0046*** | (0.0011) | 203,244 |
| Older (40+) | 0.0002 | (0.0056) | 202,412 | 0.0014 | (0.0057) | 183,351 |
| All | 0.0039*** | (0.0008) | 441,890 | 0.0042*** | (0.001) | 400,636 |
| <i>Regional Earnings Cyclicity</i> | | | | | | |
| Young (aged 27 to 39) | 0.0038*** | (0.0007) | 223,974 | 0.0036*** | (0.0009) | 202,722 |
| Older (40+) | 0.0002 | (0.0056) | 202,412 | 0.0018 | (0.0055) | 182,839 |
| All | 0.0039** | (0.0008) | 441,890 | 0.0034** | (0.0011) | 399,567 |
| <i>Real Pay Cut</i> | | | | | | |
| Young (aged 27 to 39) | -0.0082* | (0.0037) | 157,710 | -0.0083 | (0.0065) | 143,233 |
| Older (40+) | -0.0093 | (0.0063) | 195,484 | -0.0061 | (0.0062) | 176,995 |
| All | -0.011** | (0.0045) | 353,194 | -0.0094 | (0.0053) | 320,228 |

Local linear regression estimates using cohorts born between 1952 and 1962 with the 1957 cohort omitted. All regressions done at the cohort level with weights equal to the number of observations per cohort. Standard errors in parenthesis. Significance level: *** at .01, ** at .05 and * at .10

Effect of Using Robust Standard Errors

Table 3.14: Effect of the Law on Log Weekly Pay

| Year (age of 1957 cohort) | Effect | Year (age of 1957 cohort) | Effect |
|---------------------------|--------------------|---------------------------|------------------|
| 1983 (25) | 0.0179 (0.011) | 1999 (41) | 0.007 (0.017) |
| 1984 (26) | 0.0266** (0.0116) | 2000 (42) | 0.0251 (0.0173) |
| 1985 (27) | 0.0213* (0.0121) | 2001 (43) | 0.0165 (0.018) |
| 1986 (28) | 0.0285** (0.0123) | 2002 (44) | 0.001 (0.0183) |
| 1987 (29) | 0.0261** (0.013) | 2003 (45) | 0.0172 (0.0182) |
| 1988 (30) | 0.0106 (0.0133) | 2004 (46) | 0.0054 (0.0187) |
| 1989 (31) | 0.0292** (0.0136) | 2005 (47) | 0.0226 (0.0191) |
| 1990 (32) | 0.0301** (0.0137) | 2006 (48) | 0.0239 (0.0194) |
| 1991 (33) | 0.0354** (0.0139) | 2007 (49) | 0.0375* (0.0225) |
| 1992 (34) | 0.0468*** (0.0146) | 2008 (50) | 0.0255 (0.0229) |
| 1993 (35) | 0.0457*** (0.0152) | 2009 (51) | 0.0317 (0.0207) |
| 1994 (36) | 0.0426*** (0.0153) | 2010 (52) | -0.0143 (0.0213) |
| 1995 (37) | 0.0275* (0.0162) | 2011 (53) | 0.0104 (0.0216) |
| 1996 (38) | 0.0213 (0.0166) | 2012 (54) | 0.0052 (0.0228) |
| 1997 (39) | 0.0176 (0.0169) | 2013 (55) | -0.0089 (0.024) |
| 1998 (40) | -0.0029 (0.0168) | | |
| Young (aged less than 40) | 0.0252*** (0.0095) | n=302,527 | |
| Older (aged 40+) | 0.0143 (0.0141) | n=271,971 | |
| All | 0.0181* (0.0105) | n=574,498 | |

Robust standard errors clustered at the individual level are in parenthesis. Significance level: *** at .01, ** at .05 and * at .10

Table 3.15: Effect of the Law on the Standard Deviation of Log Weekly Pay

| Year (age of 1957 cohort) | Effect | Year (age of 1957 cohort) | Effect |
|---------------------------|---------------------|---------------------------|--------------------|
| 1985 (27) | -0.0111*** (0.0042) | 1999 (41) | -0.0055 (0.0045) |
| 1986 (28) | -0.0118*** (0.0043) | 2000 (42) | -0.0068** (0.0047) |
| 1987 (29) | -0.0081* (0.0042) | 2001 (43) | -0.0042 (0.0047) |
| 1988 (30) | -0.0111** (0.0043) | 2002 (44) | -0.0043 (0.0047) |
| 1989 (31) | -0.0102** (0.0042) | 2003 (45) | 0.0001 (0.0048) |
| 1990 (32) | -0.0068 (0.0042) | 2004 (46) | 0.0034 (0.0047) |
| 1991 (33) | -0.005 (0.0043) | 2005 (47) | 0.001 (0.005) |
| 1992 (34) | -0.0015 (0.0043) | 2006 (48) | -0.0036 (0.0054) |
| 1993 (35) | -0.0042 (0.0046) | 2007 (49) | -0.0088 (0.0056) |
| 1994 (36) | -0.005 (0.0046) | 2008 (50) | -0.0016 (0.0058) |
| 1995 (37) | -0.0047 (0.0046) | 2009 (51) | -0.0009 (0.0057) |
| 1996 (38) | -0.0059 (0.0046) | 2010 (52) | 0.0032 (0.0052) |
| 1997 (39) | -0.0044 (0.0046) | 2011 (53) | 0.006 (0.0054) |
| 1998 (40) | -0.006 (0.0046) | | |
| Young (aged less than 40) | -0.0065** (0.0026) | n = 197,741 | |
| Older (aged 40+) | -0.0021 (0.0028) | n = 175,943 | |
| All | -0.0042** (0.0021) | n = 373,684 | |

Standard deviations measured in the 5-year period centred around the listed year with at least 4 observations per window. Local linear regression estimates using cohorts born between 1952 and 1962 with the 1957 cohort omitted. Robust standard errors clustered at the individual level are in parenthesis. Significance level: *** at .01, ** at .05 and * at .10

Table 3.16: Effect of the Law on the Probability of a Real Weekly Pay Cut (over 5-Year Period)

| Year (age of 1957 cohort) | Effect | Year (age of 1957 cohort) | Effect |
|---------------------------|-------------------|---------------------------|---------------------|
| 1985 (27) | -0.0154 (0.0157) | 1999 (41) | -0.0207 (0.019) |
| 1986 (28) | -0.0253* (0.015) | 2000 (42) | -0.025 (0.0181) |
| 1987 (29) | -0.0001 (0.0155) | 2001 (43) | 0.0014 (0.0188) |
| 1988 (30) | -0.0123 (0.0165) | 2002 (44) | 0.0088 (0.0192) |
| 1989 (31) | -0.0196* (0.0171) | 2003 (45) | -0.0164 (0.0199) |
| 1990 (32) | -0.0291 (0.0177) | 2004 (46) | -0.0623*** (0.0202) |
| 1991 (33) | -0.0001 (0.0179) | 2005 (47) | -0.0117 (0.0228) |
| 1992 (34) | -0.03* (0.0176) | 2006 (48) | -0.045* (0.0235) |
| 1993 (35) | -0.0047 (0.0178) | 2007 (49) | -0.0074 (0.0208) |
| 1994 (36) | 0.0172 (0.018) | 2008 (50) | -0.0004 (0.0214) |
| 1995 (37) | 0.038** (0.0186) | 2009 (51) | 0.043* (0.0229) |
| 1996 (38) | 0.0256 (0.0184) | 2010 (52) | 0.0275 (0.023) |
| 1997 (39) | -0.0116 (0.0182) | 2011 (53) | 0.0167 (0.019) |
| 1998 (40) | -0.0073 (0.0186) | | |
| Young (aged less than 40) | -0.0082 (0.0066) | n = 157,710 | |
| Older (aged 40+) | -0.0094 (0.0067) | n = 195,484 | |
| All | -0.011** (0.0051) | n = 353,194 | |

Local linear regression estimates using cohorts born between 1952 and 1962 with the 1957 cohort omitted. Robust standard errors clustered at the individual level are in parenthesis. Significance level: *** at .01, ** at .05 and * at .10

Chapter 4

To Go to College or Not? The Role of Ability, Family Background and Risk

The earnings premium associated with going to college has risen substantially over the past few decades yet still less than half of students do not go to college. Why is this? Are the reported estimates missing something? Surely if returns are so large, estimated to be between 30% and 50% (Cunha et al (2011)), it would make sense for everyone to go. Perhaps individuals suspect that the average return may conceal the fact that there is a distribution of returns and that it may be that very high returns are driven by those at the top of the distribution with meagre returns for the majority. It may also be that individuals do not base their decision purely on monetary returns but take in to account many other factors such as how risky a college investment is, how likely it is that they will succeed in college, whether their parents approve of their decision or the cost of acquiring the education. There are thus many factors determining whether one should attend college or not. The opportunity to earn money and experience straight after high school is a very tempting factor which many teenagers may be unwilling to forgo. There is also the psychic cost to education (Carneiro, Hansen, and Heckman (2003), Cunha, Heckman, and Navarro (2005, 2006), Cunha and Heckman (2006), Hai and Heckman (2015)); although many teenagers may be aware that attending college will increase their lifetime earnings, the disutility from attending college may be enough to deter them from undertaking such an endeavour. Those with lower cognitive ability may find that the course-work is too difficult while those with lower non-cognitive ability may not have the

determination, motivation or self-application to complete the course.¹ Heckman, Stixrud and Urzua (2006) find that both cognitive and non-cognitive skills have a large impact on both wages and the decision to attend college.

Furthermore, family background may influence the college decision. Family background could affect college attendance via many channels including family income, parental education, genetic traits, child investment, quality of the neighbourhood, or tastes for education. There is a large literature examining credit constraints but the majority find little evidence that such constraints exist (Keane and Wolpin (1997), Caneiro and Heckman (2002), Cameron and Taber (2004)). However, Bellely and Lochner (2007) using the NLSY find that for more recent cohorts credit constraints are binding. Abbot et al (2016) also find that parental wealth is significant in a regression of college attendance; however they urge caution in interpreting this as evidence of credit constraints since parental wealth may be correlated with psychic costs of schooling. The evidence of whether parental education has any effect on childrens' education is mixed; some studies such as Chevalier (2003) and Oreopoulos et al (2006) find that there is a causal effect of parental education on child's education while the evidence by Black, Devereux and Salvanes (2005) and Lee, Roys and Seshardi (2015) suggests that the observed correlation is due to selection rather than any causal effect.² Family background may also affect the decision to go college due to the culture towards education which is created within the household. An individual may grow up in a household where the parents may not value education and may instil this dislike of education in to the child. Parental interest in the child's education may play an important role in encouraging individuals to attend college; children who do not have the support, encouragement or interest of their parents in their education may become disengaged from school and subsequently become unlikely to attain good grades and go to college.

Finally, risk may play a big part in this decision.³ Human capital like any other

¹Policies aimed at pushing all students in to college regardless of their ability may be ill-warranted if such students go on to drop out. Eckstein and Wolpin (1999) find that those who drop out of high school have lower ability and motivation, lower expectations of the rewards from staying on and a lower consumption value of school attendance.

²See Bjorklund and Salvanes (2011) for a summary and evaluation of recent literature on family background and education.

³There were a handful of papers looking at this topic in the 1970s and 1980s including Levhari and Weiss (1974), Williams (1979), Nickell (1979), Eaton and Rosen (1980), Kodde (1986). More recent papers examin-

investment is subject to the perils of risk. An individual deciding whether to invest in education faces uncertainty concerning whether they will graduate on time or at all, the type of job they will get, their future earnings and what fraction of time will be spent in employment. Income levels among observationally similar people may differ due to luck, social connectedness, illness, promotions, ability, different training opportunities or motivation and therefore there is a wide range of potential earnings outcomes which may be realised. To the extent that the college earnings distribution is significantly skewed to the right, a typical college graduate will earn significantly less than average earnings and consequently risk averse individuals may be willing to trade high earnings for low risk and may decide not to go to college. Given that there does not exist any market that insures against low returns to education this may explain part of the reason why many students do not progress to further education despite the perceived benefits.⁴

On the other hand, it is well documented that those with more education are less likely to experience spells of unemployment (Mincer, 1991) and this will tend to make college more attractive since if higher educated individuals face lower unemployment rates they will be in receipt of income for a larger fraction of their lives.⁵ However, in order to rigorously examine whether this is indeed the case, it is necessary to have a life cycle model since the dynamic effects of unemployment may differ across education levels due to varying levels of human capital depreciation, costs of foregone work experience and the impact on future wages and employment prospects due to scarring and atrophy. Additionally, the system of unemployment insurance that exists will also have a bearing on the outcome; generous unemployment benefits may help to negate the adverse affects of unemployment.

There are thus many factors which may impact the decision to go to college

ing this topic include Harmon, Hogan, and Walker (2003), Carneiro, Hansen, and Heckman (2003), Belzil and Hansen (2004), Hartog and Diaz-Serrano (2007) and Brown, Fang and Gomes (2012).

⁴The evidence on whether education increases wage risk is mixed; Cunha and Heckman (2007) find the variance of earnings to be larger for college graduates. On the other hand, Chen (2002) finds that once known heterogeneity is accounted for, risk does not rise with education while Meghir and Pistaferri (2006) and Abbot et al (2013) find no difference in wage risk across education levels. However, Delaney and Devereux (2016) using a compulsory school leaving age reform find evidence that those with more schooling experience less wage volatility over the life cycle.

⁵Education may lead to lower rates of unemployment if those with more education have higher levels of human capital, are less likely to be fired due to higher training costs and have the option to downgrade occupation in times of recession.

and it is paramount that any study attempting to understand the decision to go to college and the subsequent returns which are obtained take into account as many pertinent factors as possible. In this paper I estimate a structural life cycle model with savings, labour supply, ability, human capital accumulation and depreciation, employment risk and wage risk. In addition, similar to Keane and Wolpin (1997) who have a dynamic model of schooling, work, and occupational choice, I explicitly model the education decision. Their model is quite different to mine as they do not allow for unemployment risk, for persistent wage shocks nor do they include savings in the model.⁶ The inclusion of labour supply is important for two reasons. Firstly, labour supply is the utilisation of human capital and so directly impacts the returns. Secondly, endogenous labour supply decisions result in earnings fluctuations so without modelling labour supply explicitly the measure of risk would be upward biased. The inclusion of savings is important as it provides a channel through which individuals can self insure; moreover, without savings the effect of labour supply on achieving consumption smoothing would be greatly exaggerated. The model is similar to Low, Meghir and Pistaferri (2010) who estimate a life cycle model with consumption, labour supply, job mobility, employment risk and wage risk. They are interested in the insurance value of different welfare programs and do not look at the decision to go to college nor do they include information related to family background or ability. The model also borrows from Blundell et al (2013) who examine the effects of in-work benefits on female labour supply and human capital accumulation by allowing for education specific returns to experience to capture the dynamic effects of labour force participation and unemployment. Their paper focuses solely on females and they do not allow for employment frictions.

To the best of my knowledge, this is the first paper to have an education choice model with early measures of both cognitive and non-cognitive skills, parental background information and a rich earnings process capturing human capital accumulation, labour market risk and allowing for savings. Related papers looking at risk in education include Altonji (1993), Chen (2002), Abbot et al (2013), Athreya and Kartik (2013) and Brown et al (2013). While each of these papers allow for risk to

⁶While they do let unobserved heterogeneity enter the model by allowing for 4 types of individuals they do not explicitly model the affect of cognitive and non-cognitive ability on college decisions.

affect college returns, the majority of these papers do not incorporate rich measures of cognitive and non-cognitive abilities, allow for employment risk or estimate an earnings process with endogenous human capital accumulation. Moreover, Brown et al (2013) are the only paper to allow for employment risk in addition to wage risk.⁷ However, Brown et al (2013) are only interested in the ex-post risk-adjusted return to college and do not model the education decision or allow for heterogeneity in individual traits. The paper is also related to the literature that looks at the effect of cognitive ability on returns to education including Griliches et al (1972), Blackburn and Neumark (1993), Murnane et al (1995), Cameron and Heckman (1998), (2001), Taber (2001), Caneiro and Heckman (2002) and Hendricks and Schoellmann (2009). Finally, the paper is also related to the work by Heckman et al (2006), Blanden et al (2007), Conti et al (2010), Heckman et al (2014), and Hai and Heckman (2015) by allowing for non-cognitive skills to play a role in the college decision. However, these papers are less interested in the effect of risk. The paper that is closest in spirit to this paper is Navarro (2013) which allows for savings, psychic costs and specifically models the education decision. The paper is interested in the amount of uncertainty in college earnings which is known in advance. However, the paper is different to this paper as it does not model labour supply or allow for employment risk and thus implicitly assumes that lifetime labour supply is education invariant.

In the paper, I find that risk, grants, cognitive and non-cognitive ability have a substantial effect on the decision to go to college. In particular, I find that a decrease of 10% in the variance of college wages leads to a 2.87% increase in participation and an increase in grants by 1,000 leads college participation to rise by almost 2.14%. However, the biggest impact on college participation comes from policies which alter individual traits *before* college. Giving the high school graduates the same distribution of cognitive ability as college graduates leads college attendance to rise by just roughly 20%. This is driven both by differences in returns to college conditional on ability and by the larger psychic costs faced by those with low ability. This suggests that policies aimed at early childhood investment are key to increasing

⁷Nickell (1979) was the one of the earliest papers to allow for employment risk to impact returns to education.

college attendance.

The paper proceeds as follows: section 2 describes the data, section 3 contains some reduced form analysis, section 4 describes the model, section 5 discusses the estimation, section 6 conducts some policy analysis and section 7 concludes.

4.1 Data

I use the National Child Development Study 1958 (NCDS) which follows a sample of approximately 17,000 individuals born in Great Britain in a week in March in 1958. The NCDS contains rich data on family background, education, work experience, earnings and test scores. I limit the sample to males only.⁸ The information on cognitive ability relates to reading, mathematics, general ability and motor ability tests taken at different ages. I use the mathematics test score at age 11 as my measure of cognitive ability. The test comprised of 40 questions relating to numerical and geometric ability. I take the total score and dichotomise it in to a high or low variable. The dataset also contains comprehensive non-cognitive measures. The school teacher is asked to fill out the Bristol Social Adjustment Guide (BSAG) which is a range of questions relating to the child's behaviour in school at ages 7 and 11. The questions can be grouped together to pinpoint what form of behavioural disturbance is most relevant for the child, for example, unforthcomingness, hostility, anxiety, restlessness and inconsequential behaviour. I use the total score on all questions at age 11 and dichotomise the combined score in to a binary variable to form my measure of non-cognitive ability. The depth of questions related to behaviours provide an advantage over many other studies which examine non-cognitive skills using, for example, the NLSY which has limited questions related to non-cognitive skills.⁹ In addition, another important advantage of using questions which are asked to the teacher results in less anchoring issues/reference bias than if the student answered the question directly and less subjective bias than if the parent answered the questions.

Questions relating to the mother and father's highest level of education are also

⁸A credible analysis of female college participation would require an enhanced model allowing for fertility decisions and part-time work.

⁹Hai and Heckman (2015) use 3 questions including whether had under age sex, whether stole something worth more than 50 dollars or whether attacked someone.

asked. I combine the highest level of education for the mother and father and then dichotomise the variable. As alluded to already, parental education is an important factor in education decisions. This may be due to innate ability which is passed on to the child, the availability of financial resources to pay for education, or the attitude towards education which is created within the household. There are questions which elicit how interested parents are in their child's education. Teachers are asked how interested they think the mother is in the child's education and also asked about the father's interest. On top of this, parents are asked directly if they wish the child had been able to leave school at fifteen. This question was asked since the 1958 cohort was the first cohort which faced the new minimum school leaving age of 16. I use these 3 questions to form an index relating to the parental interest in the child's education and split the index in to a high or low binary variable. It can be argued that parental interest in the child's education will affect the child's decision to go to college but should not impact wages conditional on cognitive ability, non-cognitive ability and parental education. Therefore this variable provides an exogenous source of variation in the model to help identify the returns to college.

One caveat to using the cohort studies is that information is only collected at several points in the life cycle: at birth and at ages 7, 11, 16, 23, 33, 42, 50 and 55.¹⁰ Therefore, I impute earnings for the years when the individual is not surveyed. The British Household Panel Survey (BHPS) and the General Household Survey (GHS) are two other UK datasets which overlap with the life cycle of the NCDS cohort. The BHPS started in 1991 with almost 5,500 households and 10,300 individuals in the UK. I use the first 18 waves of the BHPS covering the period 1991-2008 and use variables such as region, mother's highest education level, father's highest education level, education, and sex to predict earnings in the NCDS. The GHS is a cross sectional survey which started in 1971 and surveys approximately 9,000 households and about 16,000 adults in the UK. I use the GHS to impute earnings before 1991. The variables I use to do this include region, father's social class, education, and sex. Having created the life cycle earnings profile of the individual I trim the top and bottom 1% of earnings in each year and deflate to 2012 prices

¹⁰At the moment there is work to use the individual's social security number to infer earnings for the years unobserved.

using the UK Retail Price Index. To obtain labour market experience over the life cycle I use retrospective questions in the NCDS which ask about previous periods of employment and non-employment. This allows me to create a life cycle profile of the individual's work experience and labour market history.

The early test score measures are a great advantage of using this dataset since one of the main criticisms of papers looking at returns to education is the endogeneity of education due to the fact that those with higher ability are more likely to receive higher education but it is likely they would earn more anyhow regardless of whether or not they obtained more education. Most studies in the literature try to address this problem by using a clever instrument (see Card 1995 for a review), but as Imbens and Angrist (1994) have shown, most of the time this will just lead to identification of a local average treatment effect, i.e., the treatment effect for those most sensitive to a change in the instrument which may not be policy relevant. Other papers in the literature such as Cameron and Heckman (1998), (2001) and Cameron and Taber (2004) use the NLSY which contains test scores but these are taken from ages 15 to 18 and so maybe confounded with education and likely will not pick up raw ability. Therefore using the NCDS provides me with a unique advantage to plausibly control for ability bias.

4.2 Reduced Form Analysis

Before delving in to the model, it is useful to start with some reduced form analysis to gain some insight in the importance of certain parameters both for the college decision itself and for earnings. To start with I show some simple correlations between cognitive ability, non-cognitive ability, parental education and parental interest. Figure 1 contains graphs depicting the effect of parental education on each of the other variables. As expected, there is a positive gradient for each of the variables; those in the top tercile of the parental education distribution display higher cognitive and non-cognitive ability and also tend to have parents who have a higher interest in their education.¹¹ Those from the top tercile of the parental education distribution are a lot more likely to have higher cognitive and non-cognitive ability and greater

¹¹In the figures parental interest refers to the sum of the responses to both father and mother interest in the child's education and whether the parent's want the child to leave school at age 15 to form a parental interest score. More details are provided in the appendix.

parental interest than those in the second tercile whereas the differential between the second and bottom tercile is not as substantial for each outcome.

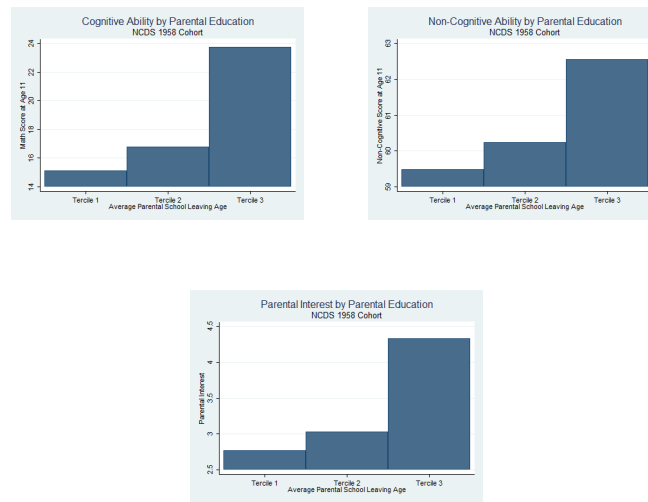


Figure 4.1: Cognitive ability, non-cognitive ability and parental interest by parental education

Figure 2 shows the correlation between cognitive ability, non-cognitive ability and parental interest. Those in the bottom tercile of cognitive ability display lower non-cognitive ability and are also less likely to have parents with a strong interest in education. The differential in non-cognitive ability between the bottom and middle tercile is a lot bigger than that between the middle and top tercile and so it seems that those in the lower end of the cognitive ability distribution display lower non-cognitive skills but the differential on other parts of the distribution is not large.

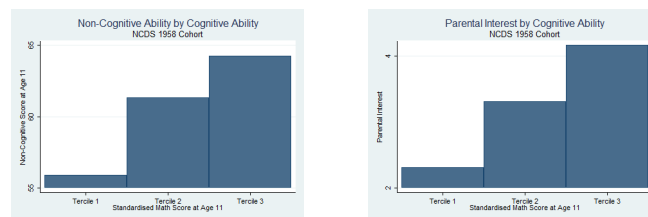


Figure 4.2: Non-cognitive ability and parental interest by cognitive ability

Lastly, figure 3 displays the relationship between non-cognitive ability and parental interest. There is a positive relationship between the two variables with those in the higher non-cognitive ability tercile more likely to have parents interested in education.

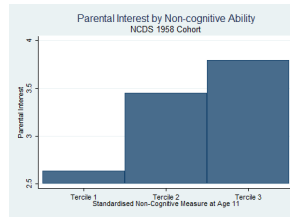


Figure 4.3: Parental interest by non-cognitive ability

Next, I turn to the relationship between each variable and college attendance. Figure 4 displays this information. As expected there is a positive relationship between each variable and college attendance. Individuals with high cognitive ability, non-cognitive ability, parental education and parental interest are more likely to go to college and the effect is monotonic across the terciles of each distribution. Cognitive ability in particular seems to have a large effect on college attendance with those from the top tercile more than 4 times as likely to go to college as those who are in the bottom tercile of the ability distribution.

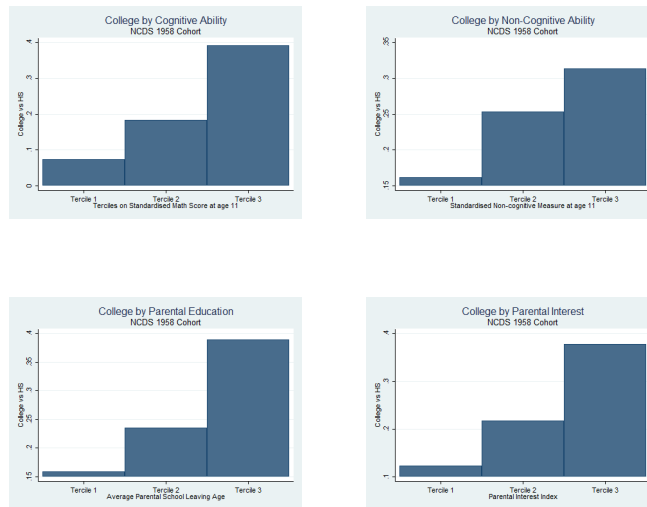


Figure 4.4: College by cognitive ability, non-cognitive ability, parental education and parental interest

Table 1 shows a regression of log earnings at ages 33 and 42 on a dummy for college (with the base category being high school graduate), average years of parental education, cognitive and non-cognitive ability. Since the cohort studies are measured only at certain points in the life cycle, I chose to focus on the two ages when males are most likely to be in the labour market. This also has the advantage of showing how standard returns to education differ if earnings are measured

at different points in the life cycle. I first show the return without any controls. It should be noted that the low sample size as compared to the original sample size of 17,000 individuals is due to the fact that I only focus on males which reduced the sample by half and I also condition on at least a high school degree. Given that a non-trivial amount of individuals born in 1958 leave school without any qualifications this again reduces the sample size by over 30 percent. The remainder is due to missing observations on any of the variables in the model. Across both ages college graduates receive a significant increase in earnings: 27% at age 33 and almost 33% at age 42. The higher returns at older ages is consistent with findings by Bhuller, Mogstad and Salvanes (2013) who study the life cycle bias present in measures of returns to education. This also shows the importance of looking over the whole life cycle which the structural model that I present in the next section will take into account. In columns 2 and 4, I show how selection on ability and family background may play a role in the large returns by including parental education, cognitive and non-cognitive ability in the regression. The returns to college fall by almost 42% for at age 33 and by almost 34% at age forty two. This is quite a staggering decrease in returns and shows the importance of controlling for individual characteristics and in particular cognitive ability. The estimates lend credence to the so called "ability bias" that may be affecting many reported estimates of returns to college. It is clear that those who have higher ability are more likely to go to college but would receive higher earnings independent of their education status. While I cannot control for other unobserved determinants which may be affecting both earnings and college, in an effort to gain some insight in to how much unobserved heterogeneity if left, I add a host of other variables to the regression including birth order, number of siblings, whether smoked at age 16, father's social class at age 16, birth weight and other ability measures such as reading at age 11, math at age 7, copying design test at age 7 and non-cognitive measures at age seven. Surprisingly, the coefficient on college barely changes at all once with the addition of these other variables. While this may suggest that controlling for measures such as parental education, math test and non-cognitive measures at age 11 does a good job at soaking up the effect of omitted variable bias it may still be the case that there are other important unob-

servables which may be omitted. Nonetheless, I think this exercise illustrates how important controlling for cognitive ability is when measuring returns to education.

Table 4.1: Regression of Log Earnings at Ages 33 and 42

| | Log(earn) at 33 | Log(earn) at 33 | Log(earn) at 33 | Log(earn) at 42 | Log(earn) at 42 | Log(earn) at 42 |
|---------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| College versus HS | 0.2369*** (0.0296) | 0.1376*** (0.0315) | 0.1347*** (0.0312) | 0.3148*** (0.0437) | 0.2092*** (0.0453) | 0.2123*** (0.0460) |
| Father Age Left School | | 0.0260** (0.0101) | 0.0159 (0.0110) | | 0.0282* (0.0151) | 0.0186 (0.0160) |
| Mother Age Left School | | 0.0010 (0.0126) | -0.0010 (0.0127) | | -0.0045 (0.0159) | -0.0082 (0.0162) |
| Math at 11 | | 0.1136*** (0.0165) | 0.0760*** (0.0253) | | 0.1259*** (0.0237) | 0.1083*** (0.0322) |
| Non-Cognitive at 11 | | 0.0210 (0.0162) | 0.0151 (0.0172) | | 0.0390 (0.0281) | 0.0221 (0.0285) |
| Reading at 11 | | | 0.0097 (0.0254) | | | -0.0236 (0.0299) |
| Non-Cognitive at 7 | | | 0.0225 (0.0165) | | | 0.0551** (0.0248) |
| Math at 7 | | | 0.0347** (0.0173) | | | 0.0057 (0.0232) |
| Copying Design at 7 | | | 0.0031 (0.0138) | | | 0.0240 (0.0209) |
| Number of Siblings | | | -0.0160 (0.0099) | | | 0.0036 (0.0156) |
| Birth Order | | | -0.0013 (0.0150) | | | -0.0357 (0.0224) |
| Smokes at 16 | | | 0.0511* (0.0300) | | | 0.0367 (0.0456) |
| Father Social Class at 16 | | | 0.0895*** (0.0305) | | | 0.1060** (0.0425) |
| Birth Weight | | | 0.0192 (0.0142) | | | 0.0116 (0.0213) |
| Observations | 1110 | 1110 | 1110 | 939 | 939 | 939 |

Standard errors in parentheses. Significance levels: * 10%, ** 5%, *** 1%. The math, reading, copying design and non-cognitive variables represent the standardised measure of each variable. The birthweight variable is also standardised. Smokes at 16 is a binary variable denoting whether the individual smoked or not at age sixteen. Father's social class is a binary variable denoting whether the father worked in a skilled or unskilled occupation when the individual was sixteen.

4.3 Model

4.3.1 Set-up

Individuals complete high school/A-levels at age 18 then decide whether to enter the labour market or go to college in the next period – at age 19. The decision depends on expected benefits and costs including idiosyncratic tastes for education. College lasts for 3 years after which individuals enter the labour market. Each period an individual is subject to an unemployment shock if employed and a job offer arrival if unemployed. After observing these shocks he must decide whether it is optimal

to work or not. Individuals retire at age 65 and face a mandatory spell of retirement of 10 years at the end of life where they consume their savings and receive the state pension. The date of death is known with certainty and there is no bequest motive.

4.3.2 Utility

The agent wishes to maximise the present discounted value of lifetime utility subject to a budget constraint

$$\max_{c,P} E_{t_0} \sum_{t=t_0}^T \beta^{t-t_0} U(c_{st}, P_{st})$$

The utility function is of CRRA form and is non-separable in consumption and leisure

$$U(c_{st}, P_{st}) = \frac{c_{st}^{(1-\gamma)} (\exp(\eta_s P_{st}))}{(1-\gamma)}$$

γ is the coefficient of risk aversion, η_s represents the disutility from working and P_{st} is an indicator variable denoting whether the individual is employed or not. s denotes education, all other parameters of the model are education specific (the discount rate and coefficient of relative risk aversion are the only ones not) so from now on I drop the s subscript.

4.3.3 Budget Constraint

While in college assets are a function of the current interest rate plus the amount of grants received.¹²

$$a_{t+1} = (1+r)a_t + g(t) - c_t$$

¹²In the UK in 1976 local education authorities paid all fees which amounted to 375 pound per year. While the amount of grants was means tested and therefore individuals did not receive the full level of grants (although all students received at least 50 per year) since I do not have data on parental transfers I assume that all individuals received the full level of grants. Therefore, I am assuming that the amount of transfers given by parents whose income was above the threshold for grant eligibility, was equal in value to the amount of the grant.

When the individual has left education assets accumulate depending on the interest rate and whether or not they are employed. If they are employed they get an income (y) but if they are not in employment they just receive unemployment benefits (UI).

$$a_{t+1} = (1 + r)a_t + P_t y_t + (1 - P_t)UI - c_t$$

I do not allow individuals to borrow as in the UK in 1976 it wasn't very common for individuals to borrow particularly since there was no fees.

$$a_{t+1} \geq 0$$

4.3.4 Earnings Process

The earnings process is education specific and composed of a permanent and transitory component:

$$\ln y_{it} = Y + \beta_1 Cog_i + \beta_2 NonCog_i + \beta_3 Exp_{it} + \beta_4 Exp_{it}^2 + u_{it} + v_{it}$$

$$v_{it} = \rho v_{it-1} + \zeta_{it}$$

The unobserved income shocks $u_{it} \sim N(0, \sigma_{u_t}^2)$ and $\zeta_{it} \sim N(0, \sigma_{\zeta_t}^2)$ are independent and serially uncorrelated. I assume the transitory component u_{it} is just measurement error and only focus on the permanent component v_{it} in estimating the model. Blundell et al (2008) show that the transitory component can be easily smoothed over time and so I focus on the permanent part which has the biggest welfare effects. Moreover, in the model I allow for job destruction which usually constitutes the main part of transitory shocks and I have trimmed the top and bottom 1 percent of earnings which is largely affected by measurement error.

Endogenous experience accumulation depends on labour force participation and the rate of depreciation. Experience depreciates in each period at the rate δ . However, an individual can increase his stock of experience by participating in the

labour market.

$$Exp_{ist} = Exp_{ist-1} * (1 - \delta) + P_{ist}$$

4.3.5 Value Functions

Let I denote the agent's information set which is composed of time, education, assets, experience, the permanent wage shock, cognitive and non-cognitive ability, parental education and parental interest in education. Labour market frictions enter the model via job destruction and job offers. In each period the individual may lose their job, the probability of which depends on their education level. Similarly, each period when an individual does not have a job they will receive a job offer which depends on their level of education. I assume that both the probability of job destruction and job offers are constant over the life cycle.

The value functions conditional on employment status are as follows:

$$V_t^e(I) = \max_c [U(c_{it}, P_{it} = 1) + \beta(\pi E_t V_{t+1}^u(I) + (1 - \pi) E_t \max(V_{t+1}^e(I), V_{t+1}^u(I)))]$$

$$V_t^u(I) = \max_c [U(c_{it}, P_{it} = 0) + \beta((1 - \omega) E_t V_{t+1}^u(I) + \omega E_t \max(V_{t+1}^e(I), V_{t+1}^u(I)))]$$

where π is the probability of job destruction and ω denotes the probability of a job offer, e denotes employment and u denotes unemployment.

4.3.6 Education Decision

Agents make their college choice by comparing the expected present discounted utility of leaving school with just a high school degree against the value from choosing to enter college which depends on the costs of college C_i . This decision is made when the agents have just completed high school and depends on the initial joint distribution of assets, cognitive ability, non-cognitive ability, parental education and the parent's interest in their education. Parental education and parental interest do not enter the earnings process and so

act as exclusion restrictions which help to identify the model parameters. The intuition is that conditional on cognitive ability, non-cognitive ability and parental education; parental interest only acts to affect the individual's decision to go to college but has no effect on earnings or utility other than via its effect on education thus providing exogenous variation in the college decision.

$$Ed = \max((EV_{hs}|I18), (EV_{cg}|18, I18) - C_i)$$

$$C_i = \beta_0 + \beta_1 cognitive_i + \beta_2 noncog_i + \beta_3 pareduc_i + \beta_4 parinterest_i + \varepsilon_i$$

Here ε comes from a normal distribution with mean zero and variance to be estimated within the model, $I18$ denotes the individuals' information at age 18 which includes assets, ability, parental education and parental interest and C_i denotes the costs of schooling.

4.4 Estimation

I use a two step estimation approach. In the first step, I calibrate a set of parameters using estimates from the literature. In the second step, I use indirect inference (Gourieroux, Monfort and Renault, (1993)) to estimate all the remaining parameters.

$$\hat{\theta} = \arg \min_{\theta} (\hat{M}_d - \tilde{M}_s(\theta))' W (\hat{M}_d - \tilde{M}_s(\theta))$$

In practice, I match a mix of coefficients from regressions as well as data moments in order to estimate the parameters. A full list of moments and coefficients used in the estimation is provided in the appendix. The aim of the algorithm is to minimise the distance between the data moments and the moments got from simulating the model. I start with an initial guess for the parameters and compute the distance between the model and data moments and then keep updating the parameters until the distance between the data and simulated moments has converged to close proximity. I use a diagonal weighting matrix with the inverse of the variance of moments on the diagonal. Altonji and Segal (1996) have shown that using the

efficient variance-covariance weighting matrix leads to biased estimates in finite samples. More details on the computation is contained in the appendix.

4.4.1 Calibration

I calibrate the coefficient of risk aversion γ to be 1.56 which is similar to findings by Atanasio and Weber (1995) and Blundell, Browning and Meghir (1994). I set the interest rate to 0.015 and the discount rate equal 0.98 implying that the individuals have some degree of impatience. I calibrate unemployment benefits to be 3728.40 for those aged 25 and over and 2925 if under age 25. This corresponds to the annual UK unemployment benefits in 2012 prices which for simplicity I assume can be received for the full period of 1 year. I set the pension income equal to 5727.80 which is equivalent to what a person with 30 qualifying years (years in which national insurance contributions were paid) at flat rate of 110.15 per week would obtain. I set the depreciation rate to 10% similar to what Fan et al (2015) have found using a structural model with labour supply and endogenous experience and also Schmeider et al (2014) who use a regression discontinuity design of unemployment duration in Germany.

4.4.2 Parameter Estimates

Table 4.2: Model Estimates with standard errors in parenthesis

| | High School | College |
|--------------------------------|---------------------|---------------------|
| Earnings | | |
| Intercept | 9.47 (0.0026) | 9.72 (0.0033) |
| Experience | 0.075 (0.003) | 0.095 (0.0022) |
| Experience Squared | -0.0015 (0.0093) | -0.0023 (0.0068) |
| Cognitive Ability | 0.123 (0.031) | 0.20 (0.0313) |
| Non-Cognitive Ability | 0.0786 (0.0311) | 0.0339 (0.0302) |
| AR Parameter | 0.87 (0.0367) | 0.914 (0.0111) |
| Variance of Permanent Shocks | 0.06 (0.0196) | 0.0436 (0.0093) |
| Labour | | |
| Tastes for Work | 0.23 (0.0318) | 0.2354 (0.0557) |
| Probability of Job Offer | 0.3459 (0.0393) | 0.5087 (0.0638) |
| Probability of Job Destruction | 0.0137 (0.0072) | 0.0129 (0.0131) |

Table 4.3: Psychic Cost Equation

| | |
|----------------------------------------|---------------------|
| Schooling Cost | |
| Intercept | 0.6369 (0.0278) |
| Cognitive Ability | -0.1461 (0.0283) |
| Non-Cognitive Ability | -0.0524 (0.0217) |
| Parental Education | -0.0561 (0.0384) |
| Parental Interest | -0.0697 (0.0431) |
| Standard Deviation of Unobserved Taste | 0.3627 (0.0261) |

Unsurprisingly, I find that college graduates start out in the labour market commanding a higher level of earnings than what high school graduates earn when they enter the labour market for the first time. Not only do college graduates receive a higher level of earnings at the beginning of their work lives but the growth in their earnings is also faster resulting in steeper earnings profiles for college graduates. However, the earnings profiles flatten out a lot quicker for the college graduates.

An interesting feature of the model estimates is the finding that the return to cognitive ability for college graduates (20%) is nearly double the return for high school graduates (12.3%). This implies that college and ability are complementary with those with higher ability receiving larger returns. Conversely, the return to non-cognitive ability is 7.86% for high school graduates but only 3.39% for college graduates with the latter estimate not significantly different from zero. The higher return to non-cognitive skills for high school graduates may be due to the fact that the jobs high school graduates obtain rely more on soft skills. However, it is still puzzling that the non-cognitive skills do not seem to pay off for college graduates. Hai and Heckman (2015) using the NLSY97 and controlling for schooling find that returns to non-cognitive skills are not significantly different from zero.

Turning to the risk parameters, I find that the job destruction rate is only slightly higher for high school graduates, 1.37% versus 1.29%, which may be due to the fact that I am looking at males with at least a high school degree born in 1958 whom tended to have a high labour market attachment. However, college graduates receive job offers at the rate of 50.87% while it's only 34.59% for high school graduates. Therefore, as expected college graduates do receive less employment risk than their high school counterparts. However, the overall variance of the earnings is slightly higher for college graduates consistent with the idea that higher returns come at the expense of higher risk. Although the variance of permanent shocks is smaller for those with college the persistence parameter of 0.914 is large enough to make the overall earnings risk larger for college graduates.

I find that those with greater cognitive and non-cognitive ability, parents with higher education and parents who are more interested in the child's education face a lower psychic cost of going to college. Similar to Navarro (2011), I find that cognitive ability is the main determinant of the psychic cost function. The larger returns to cognitive ability for college graduates combined with the lower psychic costs for those with high cognitive ability imply a large degree of college sorting by ability.

4.4.3 Compensating Variation

I now use the parameter estimates to calculate a return to college education. Since I am working with a utility framework it is not reasonable to just look at the difference in lifetime earnings. Therefore in order to find the return to college which takes a comprehensive account of all parameters of the model I measure a type of compensating variation. This got by finding the constant fraction of per-period consumption a college graduate would have to give up in order to make him indifferent between college and high school:

$$EV_{hs} = \frac{E_0 \sum_t \beta^{t-1} ((1 - CE) C_{cg} \exp(\eta_{cg} P_{cg}))^{1-\gamma}}{1 - \gamma}$$

$$CE = 1 - \left(\frac{EV_{hs}}{EV_{cg}} \right)^{\frac{1}{1-\gamma}}$$

Table 4 displays the compensating variation for a range of subgroups. The overall compensating variation is 18.8% which means that a college graduate would have to be compensated by almost one fifth of his per-period consumption to make him indifferent between high school and college. There is quite a lot of heterogeneity in the compensating variation with those with higher ability having the largest compensating variation of 27.7% and those with the lowest ability needing only 11.2%. The compensating variation differential between high and low non-cognitive ability is not as large as the payoff to non-cognitive skills is larger in the high school labour market than the college labour market although the psychic costs to college are lessened for those with greater non-cognitive ability.

Table 4.4: NCDS 1958 Cohort Compensating Variation

| | Compensating Variation | PDV(Y) |
|----------------------------|------------------------|--------|
| Overall | 0.188 | 0.28 |
| Low Cognitive Ability | 0.112 | 0.244 |
| High Cognitive Ability | 0.277 | 0.322 |
| Low Non-Cognitive Ability | 0.184 | 0.304 |
| High Non-Cognitive Ability | 0.192 | 0.259 |
| Low Parental Education | 0.172 | 0.274 |
| High Parental Education | 0.222 | 0.292 |
| Low Parental Interest | 0.168 | 0.274 |
| High Parental Interest | 0.204 | 0.285 |

Table 5 displays the proportion attending college conditional on ability and family background variables. Consistent, with the fact that the largest returns to college are obtained by those with higher cognitive ability, I also find that the biggest proportion attending college is those who are high ability with just over 35% of the high cognitive ability group going to college. On the other hand, just under 17% of the low cognitive ability group attend college. Not surprisingly, those with higher parental interest, higher parental education and higher non-cognitive ability are far more likely to attend college than those in the low groups.

4.4.4 Heterogeneity in College Attendance

Table 4.5: NCDS 1958 Cohort College Attendance

| | College Proportion |
|----------------------------|--------------------|
| Overall | 0.2531 |
| Low Cognitive Ability | 0.168 |
| High Cognitive Ability | 0.352 |
| Low Non-Cognitive Ability | 0.208 |
| High Non-Cognitive Ability | 0.292 |
| Low Parental Education | 0.216 |
| High Parental Education | 0.327 |
| Low Parental Interest | 0.20 |
| High Parental Interest | 0.295 |

4.5 Policy Experiments

One advantage of using a structural model is that it is possible to do policy experiments. I conduct several policy experiments to gain insight into the main determinants of college participation. Firstly, I look at the effect of a 1,000 pound increase in grants.¹³ I find that such an increase in grants leads college attendance to increase by 2.14%. This is similar to Deardon, Fitzsimons and Wyness (2011) who use variation in the level of real grants in the UK and find that a 1,000 increase in grants leads to a 2.6% increase in college attendance.

Decreasing the variance of the college permanent wage shock by 10% leads to an increase of 2.87% in college participation.¹⁴ This is quite a large effect and suggests that risk may be an important reason for the low college participation rate, particularly for those

¹³There is a large literature looking at the effect of changing tuition on college attendance (Kane (1994), Dynarski (2002), Avery and Hoxby (2004) and Epple et al (2006), Blanden and Machin (2004)) and the evidence is quite mixed.

¹⁴I also look at the effect of increasing the variance of the permanent shocks by 10% and find a decrease in college participation of 2.72%. Abbot et al (2016) analysing the same policy experiment using US data find a decrease of 2.5%.

from less well off backgrounds whose parents do not have a buffer stock of savings to insure against low returns. From an intergenerational perspective this may be a reason for the low education mobility found in the UK (Blanden and Machin (2013)); if risk is deterring students from less affluent backgrounds from going to college this will tend to be perpetuated until steps are taken to alleviate this risk.

Changing the job destruction rate and the job offer rate have negligible effects on participation perhaps due to the already high labour market attachment of this sample of males.

A 20 % increase in unemployment insurance leads to a modest decrease in participation since unemployment insurance favours high school graduates to a larger extent as they represent a larger share of predisplacement income and also they are more likely to be unemployed and so this increase helps raise the returns from being a high school graduate. This result highlights the difficulty in balancing the act of reducing poverty but at the same time providing incentives for college attendance.

Table 4.6: NCDS 1958 Cohort Counterfactuals

| | % Δ College Proportion |
|---------------------------------------------|-------------------------------|
| Increase Grants by 20% (1,000) | 2.14 |
| Decrease College Variance of ζ by 10% | 2.87 |
| Decrease College Job Destruction by 10% | 0.04 |
| Increase College Job Offer by 10% | 0.12 |
| Increase Unemployment Insurance by 20% | -0.32 |

Looking at the effect of the initial distribution of traits which individuals enter the labour market with, it is evident that there is a large return to altering some of these traits. For example, if high school graduates had the same distribution of parental education as college graduates then college attendance would increase by around 7%. While it is hard to change the parents education in any short term setting, one thing which could be targeted is the parent's interest in the child's education. This could be implemented through the use of mentoring programs for parents concerning the merits of education, disseminating information etc. As an upper bound, I look at the effect of increasing the parental interest of those high school graduates who have low parental interest, roughly 52%, and find that this would cause college participation to increase by almost 9 percent. Avvisati et al (2014) examine a field experiment in Paris which attempts to increase the level of parental involvement in the child's education. Since the focus is on student's in 6th grade, examining the effects on college attendance is not possible for now; however their results showing that the

program led to improved levels of literacy and better outcomes in terms of both attitudes and behaviour in school suggests that this may also have an effect on future college attendance.¹⁵ The main advantage of a policy like this is that cost-benefit analysis makes such a policy particularly appealing. Providing mentoring programs to parents once a week and distributing information leaflets are a cost effective way to increase college attendance.

Lastly, I look at the impact of changing the initial high school distribution of cognitive and non-cognitive ability to match that of college graduates. I find that the return from changing the cognitive ability is quite substantial resulting in an increase in college participation of over 20%. This is driven both by the larger cognitive ability returns for college graduates and the lower psychic costs faced by high cognitive ability individuals. The effect from changing the non-cognitive distribution is an increase in college attendance of 4.4% which is not as large due to the lower returns to non-cognitive skills for college graduates.

Table 4.7: NCDS 1958 Cohort Counterfactuals

| | % Δ College Proportion |
|--------------------------------------------|-------------------------------|
| Set HS Parent Educ Dist = Coll Dist | 7.07 |
| Set HS Parent Interest Dist = Coll Dist | 5.81 |
| Set Parental Interest to High for Everyone | 8.81 |
| Set HS Cognitive Ability Dist = Coll Dist | 20.43 |
| Set HS Non-Cognitive Dist = Coll Dist | 4.39 |

While I find that wage risk and grants have non-negligible effects on college participation; similar to Keane and Wolpin (1997), I find that in order to substantially increase college participation it is best to focus on improving the traits with which students enter college. This suggests policies aimed at building the foundation of skills at early ages will have the biggest return.¹⁶ There have been many experimental studies which provide evidence that interventions at early ages can lead to a increased cognitive and non-cognitive skills, for example, Doyle et al (2016) find evidence that an early childhood intervention in Dublin led to significant effects on childhood cognitive development and the effect opened up at age 18 months and continued up until the last wave of data collected at age five while Fryer, Levitt and List (2015) find evidence that parental investment has a significant effect on non-cognitive development.

¹⁵The treatment increased average grades on all subjects by .08 standard deviations on tests scores by teacher, however there was no effect on standardised tests which maybe due to reasons of subjectivity of the teachers or effort by the students since they know the teacher measured tests will be sent home to their parents whereas the standardised tests will not.

¹⁶There is evidence that cognitive ability is malleable up until the age of eight (Heckman (2005))

4.6 Conclusion

Getting more individuals to go to college has been the aim of government policy for quite sometime. But where should policies be targeted? In this paper I estimate a rich structural model incorporating savings, labour supply, endogenous human capital accumulation, wage risk, employment risk and psychic costs of college attendance. I estimate the model using British data containing early age measures of both cognitive and non-cognitive ability as well as unique information on parental background. I find evidence that returns to ability are higher in the college labour market highlighting the complementary that exists between education and ability. I also find evidence that those with higher cognitive ability face substantially lower psychic costs of attending college and this combined with the fact that the returns to ability are larger for college graduates leads to a large degree of college sorting by ability. This is consistent with reduced form evidence I present which illustrates that failure to control for ability may lead to an upward bias in returns to college of around 40 percent.

I use the model to conduct some policy analysis and find evidence that policies such as decreasing the degree of wage risk facing a college graduate or increasing the amount of grants awarded will lead to an increase in college attendance of between 2 to 3%. However, the biggest return is got from policies which focus on altering the environment of the student *before* she enters college. As an upper bound, I find that policies which get parents of all high school graduates to have a high interest in their child's education would lead to almost a 9 percent increase in college attendance. This could be implemented via mentoring programs, dissemination of information about the school and so on, policies which may be quite cost effective. Lastly, I look at counterfactuals examining the effect of changing cognitive and non-cognitive skills of the student. Specifically, I find that if high school graduates faced the same non-cognitive skill distribution as college graduates then the proportion going to college would increase by almost 4.4% while doing the same experiment for cognitive skills would lead college attendance to increase by 20 percent. This is due to the higher returns to cognitive skills in the college labour market and to the lower psychic costs faced by those with higher cognitive skills. This suggest that policies focusing on raising the cognitive skills of students at early ages will prove most fruitful which is consistent with the findings by Heckman and others which show that the biggest returns to policy is through early childhood investment.

References

Avvisati F., Gurgand M., Guyon N., Maurin E. (2014). "Getting parents involved: A field experiment in deprived schools" *Review of Economic Studies*, 81 (1) , pp. 57-83.

Abbott, B., Gallipoli, G., Meghir, C., and Violante, G. L. (2013). Education policy and intergenerational transfers in equilibrium. National Bureau of Economic Research Working Paper No. w18782

Altonji, Joseph G. "The Demand for and Return to Education When Education Outcomes are Uncertain" (1993). *Journal of Labor Economics*. Vol. 11, No. 1, Part 1: Essays in Honor of Jacob Mincer (Jan., 1993), pp. 48-83

Angrist, Joshua D., Imbens, Guido W. (1994). "Identification and Estimation of Local Average Treatment Effects", *Econometrica*, Vol. 62, No. 2, pp. 467-475. March.

Arcidiacono, P. (2004). Ability sorting and the returns to college major, *Journal of Econometrics*, 121(1-2), 343375.

Bhuller, Manudeep, Mogstad, Magne, Salvanes, Kjell G., (2011). "Life-Cycle Bias and the Returns to Schooling in Current and Lifetime Earnings," IZA Discussion Papers 5788, Institute for the Study of Labor (IZA).

Black, Sandra E., Devereux Paul J., and Salvanes Kjell G. "Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital." *The American Economic Review* 95.1 (2005): 437-49. Web.

Blanden, J. and Machin, S. (2013), Educational Inequality and The Expansion of United Kingdom Higher Education. *Scott J Polit Econ*, 60: 597598.

Blundell, R., Graber, Michael., Mogstad, Magne. (2012). Labor Income Dynamics and the Insurance from Taxes, Transfers and the Family. Working Paper.

Blundell, R., Costa Dias, Monica., Meghir, Costas., Shaw, Jonathan. (2013). "Female Labour Supply, Human Capital and Welfare Reform," Cowles Foundation Discussion Papers 1892, Cowles Foundation for Research in Economics, Yale University.

Bound and Alan B. Krueger The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right? *Journal of Labor Economics* Vol. 9, No. 1 (Jan., 1991), pp. 1-24

Cameron, S., and J.J. Heckman (1998), Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males , *The Journal of Political Economy*, 106, 262-333.

Cameron, S., and J.J. Heckman (2001), The dynamics of educational attainment for black, hispanic and white males, *The Journal of Political Economy*, 109, 455-499.

Cameron. Stephen V., Taber, Christopher (2004). "Estimation of Educational Borrowing Constraints Using Returns to Schooling" *Journal of Political Economy*, Vol. 112, No. 1, pp. 132-182, February.

Card, D. (1995). Earnings, Schooling, and Ability Revisited. *Research in Labor Economics*, edited by Solomon Polachek. Greenwich, CT: JAI.

Chen, Stacey (2008). "Estimating the Variance of Wages in the Presence of Selection and Unobserved Heterogeneity," *Review of Economics and Statistics*, 90 , 275-289.

Cunha, F., Heckman, J.J., Navarro, S. (2005). Separating uncertainty from heterogeneity in life cycle earnings, The 2004 Hicks lecture. *Oxford Economic Papers* 57 (2), 191-261. April.

Dearden, L. Fitzsimons, E. Wyness, G. (2011) The impact of tuition fees and support on university participation in the UK". CEE Discussion Papers 0126, Centre for the Economics of Education, LSE.

Delaney, Judith M. and Devereux, Paul J. (2016). "Education, Uncertainty and Business Cycles". Working Paper.

Doyle and the PFL Evaluation Team. (2016). "Did *Preparing for Life* Improve School Readiness? Preparing for Life Early Childhood Intervention Final Report"
<http://geary.ucd.ie/preparingforlife>

Eckstein, Z. and Wolpin, K. I. (1999), Why Youths Drop out of High School: The Impact of Preferences, Opportunities, and Abilities. *Econometrica*, 67: 1295-1339.

Fan, X., Seshardi, A. and C. Taber (2015). "Estimation of a Life Cycle Model with Human Capital, Labor Supply and Retirement," Working Paper.

Hai, Rong and Heckman, James J. (2015), A Dynamic Model of Health, Education, and Wealth with Credit Constraints and Rational Addiction. Becker Friedman Institute for Research in Economics Working Paper No. 2688384.

Heckman, J.J, Lochner, L.J., Todd, P.E. (2008). "Earnings Functions and Rates of Return". *Journal of Human Capital* , Vol. 2, No. 1 , pp. 1-31. Spring

Heckman, J.J, Vytlacil E.J. (2001). "Identifying the role of cognitive ability in explaining the level of and change in the return to schooling". *The Review of Economics and Statistics*, Vol. LXXXIII. February

Keane, Michael P., Wolpin, Kenneth I. (1997). "The Career Decisions of Young Men" *Journal of Political Economy*, Vol. 105, No.3 pp. 473-522, June.

Lillard, L. A., Weiss, Y. (1977). "Components of Variation in Panel Earnings Data: American Scientists 1960-70." *Nat. Bur. Econ. Res. Working Paper no. 121*, rev. Stanford, Calif.: Center Econ. Analysis Human Behavior and Soc. Inst.

Mincer, J. (1991). Education and unemployment (No. w3838). National Bureau of

Economic Research.

Meghir, Costas., Pistaferri, Luigi. (2011). "Earnings, Consumption and Life Cycle Choices," Handbook of Labor Economics, Elsevier.

Navarro, Salvador. (2011). "Using Observed Choices to Infer Agent's Information: Reconsidering the Importance of Borrowing Constraints, Uncertainty and Preferences in College Attendance". CIBC Working Paper Series.

Nickell, S. (1979). "Education and Lifetime Patterns of Unemployment". Journal of Political Economy, Vol. 87, No. 5, pp. S117-S131. October.

Olson, Lawrence., Shefrin, H.M., White, Halbert. (1979). "Optimal Investment in Schooling When Incomes Are Risky". Journal of Political Economy , Vol. 87, No. 3, pp. 522-539. June.

Padula, Mario, and Luigi Pistaferri (2005). "Education, Employment and Wage Risk," Working Paper.

Rosen, S. and Willis, R. (1979). "Education and self-selection," Journal of Political Economy 87(5), S7-36.

Weiss, Yoram. (1972) "The Risk Element in Occupational and Educational Choices" Journal of Political Economy, Vol. 80, No. 6 , pp. 1203-1213, November-December.

Appendix

Institutional Setting

The NCDS is the best dataset available to answer this question since it covers almost the entire life cycle of the individual - at present data is available up until age 55 and in addition contains rich measures of ability and family background. While the situation facing the 1958 cohort may not be applicable to other cohorts or indeed other countries, it is nevertheless one of the only datasets which contains such rich individual information and spanning such a long period of time and therefore I think the NCDS is the best dataset to answer this question.

A little bit about the set up facing the NCDS 1958 cohort. At this time, tuition fees amounted to 375 but were paid in full by the local education authority regardless of the individual's parental income. In addition, grants existed to help fund living expenses but they were means tested according to parental income. The requirement to get in to university was a minimum of 2 A-levels which meant that not everyone who wished to go to college could attend. It can be argued that by not modeling the supply side, the results may be misleading since there may have been a capacity constraint. While it is hard to obtain information on the amount of individuals who would like to have gone to college but did not get in due to rejection or due to the fact that they were discouraged from applying because of capacity constraints, I think this is more of an issue for earlier cohorts. The Robbins Report in 1963 recommended a large expansion of UK universities "on the principle that there should be a place in higher education for every student with the appropriate qualifications and motivation" and as a result in the 1960s 8 new universities were established¹⁷, some other institutions converted to university status and the Open university was opened which facilitated distance learning. Also, Chao Fu (2012) finds that increasing the supply of colleges has very limited effect on college attendance.

Model Fit

NCDS 1958

¹⁷These universities are sometimes referred to as "plate glass" universities due to the glass exterior; examples include Warwick university and Sussex university

Table A1: Comparing Simulated and Data Moments - NCDS 1958

| | Data | Model |
|-----------------------------------------------------------------------------------|---------|---------|
| Proportion Going to College | 0.2526 | 0.253 |
| Variance of College Distribution | 0.187 | 0.189 |
| HS Coeff on Lagged Yt-1 in Regression of Yt on Yt-1 | 0.869 | 0.872 |
| CG Coeff on Lagged Yt-1 in Regression of Yt on Yt-1 | 0.911 | 0.909 |
| HS Coeff on Exp in Reg of Yt on exp, exp^2 , cog and noncog ability | 0.064 | 0.067 |
| CG Coeff on Exp in Reg of Yt on exp, exp^2 , cog and noncog ability | 0.098 | 0.0831 |
| HS Coeff on Exp Sq in Reg of Yt on exp, exp^2 , cog and noncog ability | -0.0013 | -0.0012 |
| CG Coeff on Exp Sq in Reg of Yt on exp, exp^2 , cog and noncog ability | -0.002 | -0.0017 |
| HS Coeff on Cog Ability in Reg of Yt on exp, exp^2 , cog and noncog ability | 0.117 | 0.122 |
| CG Coeff on Cog Ability in Reg of Yt on exp, exp^2 , cog and noncog ability | 0.191 | 0.20 |
| HS Coeff on Non-cog Ability in Reg of Yt on exp, exp^2 , cog and noncog ability | 0.072 | 0.068 |
| CG Coeff on Non-cog Ability in Reg of Yt on exp, exp^2 , cog and noncog ability | 0.037 | 0.04 |
| HS Variance on Residual from Reg of Yt on exp, exp^2 , cog and noncog ability | 0.246 | 0.242 |
| CG Variance on Residual from Reg of Yt on exp, exp^2 , cog and noncog ability | 0.254 | 0.251 |
| HS Mean Unemployment to Employment Transition | 0.256 | 0.346 |
| CG Mean Unemployment to Employment Transition | 0.40 | 0.609 |
| HS Mean Employment to Unemployment Transition | 0.016 | 0.024 |
| CG Mean Employment to Unemployment Transition | 0.011 | 0.041 |