UCD GEARY INSTITUTE
DISCUSSION PAPER
SERIES

# Sheepskin or Prozac:

# The Causal Effect of Education on Mental Health

**Draft 4.3**

Dr. Arnaud Chevalier

Department of Economics, University of Kent, Canterbury, CT2 7NP, England

Associate at the Centre for Economics of Education, London School of Economics,

UCD Geary Institute, University College Dublin & Institute for the Study of Labour, IZA, Bonn

a.chavalier@kent.ac.uk Tel: +44 (0) 1227 827412, Fax: +44 (0) 1227 827850

Dr. Leon Feinstein

Institute of Education, Centre for Research on the Wider Benefits of Learning, 20 Bedford Way, London WC1H OAL, England

Associate at the Centre for the Economics of Education, London School of Economics, England

lf@bg.ioe.ac.uk

## 08 June 2007

**Abstract**:
Mental illness is associated with large costs to individuals and society. Education improves various health outcomes but little work has been done on mental illness. To obtain unbiased estimates of the effect of education on mental health, we rely on a rich longitudinal dataset that contains health information from childhood to adulthood and thus allow us to control for fixed effects in mental health.

We measure two health outcomes: malaise score and depression and estimate the extensive and intensive margins of education on mental health using various estimators. For all estimators, accounting for the endogeneity of education augments its protecting effect on mental health. We find that the effect of education is greater at mid-level of qualifications, for women and for individuals at greater risk of mental illness. The effects of education are observed at all ages, additionally education also reduces the transition to depression. These results suggest substantial returns to education in term of improved mental health.

**Key words**: Returns to education, mental health
**JEL**: I12, I29

**Acknowledgment**:
The data was made available by the Data Archive at Essex. The project was partially funded by the Evidence Based Policy Fund from H. M. Treasury. The authors thank participant at LoWER (LSE), COST (Amsterdam), RES (Nottingham), Econometric Evaluation of Public Policies (Paris) and seminars at Sheffield University, CEP, Portsmouth University, Université de Rouen, Louvain la Neuve and Zurich as well as Tarja Viitanen, Gauthier Lanot and the late Mark Berger for comments and suggestions that have enhanced early version of this paper. We are also indebted to Christian Hansen for providing the codes to estimate IV quantile regressions. All remaining errors are solely ours.

## Introduction

The prevalence of depression has increased considerably in the last decades so much so that the World Health Organisation (1999) rated depression as one of the main three causes of disability and morbidity in the developed world. Antidepressant drugs have at the same time become household names as their prescription more than doubled in the last decade, and the stigma associated with mental health have dwindled. In 2000, one in seven visits to a General Practitioner in England was related to depression and 2.6 million individuals were diagnosed as depressed. US evidence is similar with 6.3% of adults – 12.5 million individuals – suffering from major depression in 1999 (Dickey and Blumberg, 2002)

Poor mental health generates an annual loss of 110 million working days in the UK. Adding the cost of treatment and mortality, Thomas and Morris (2003) estimate a total annual cost of depression in the UK reaching £9 billion. This cost nevertheless omits the reduced quality of life of suffered and of their relatives. Because of these large costs, partially borne by society, policies reducing the risk of mental illness may thus have large private and social returns.

Education and other learning interventions have been found to have a positive effect on various health outcomes (see Grossman 2000 and 2005 for extensive surveys). Education directly affects heath outcomes by making individuals more able to process information and thereafter more health conscious (allocative efficiency) or by improving the efficiency of treatment. For example, the more educated are more prompt to seek diagnosis and more diligent in following treatment (Goldman and Lakdawalla, 2001 or Goldman and Smith, 2002). The correlation between education and health could also originate from three alternative reasons: first, a third factor such as genetic endowments, parental background or discount rate, affecting both education attainment and health (Fuchs, 1982). Second, the correlation reflects reverse causality from health to education or third, stems from non-classical measurement error, where qualification is correlated with the error term.

The literature has estimated large effects of education on several health outcomes but has mostly ignored mental health. Estimating the causal effect of education on mental health is an important policy issue, not only as a way to reduce the social costs of depression but also as an additional return to educational investment. As for other health outcomes and maybe even more importantly for mental health, the first hurdle comes from measurement problems. Self-reported measures are problematic as individuals may i) misreport their status due to stigma (Backer et al., 2005) or ii) lack of awareness of their mental health status. Measurement error will bias the estimates of the returns to education if it is not independent of the individual's educational attainment. In this paper, we rely on indirect measures of mental health based on a general questionnaire and consider the effect of education on both the intensive and extensive margins of mental health by defining a malaise score as well as an indicator for depression. We discuss the reliability of these measures of mental health and also provide evidence using alternative markers of mental health.

The British National Child Development Survey has a longitudinal structure which is used to control for potential unobserved time-invariant characteristics affecting mental health. We observe mental health outcomes over two decades and thus estimate the effect of education as the individual ages, but also whether education reduces the risk of transition to a depressed state. In the absence of experiments where education has some exogenous component, we rely on instrumental variables to estimate a causal effect of education. We provide a variety of estimates based on linear and probit model but also count and quantile regression to estimate the effect of education at various points of the mental health distribution.

Tentatively evidences on the channel by which education affects mental health are also provided. We consider the following channels: (1) economic factors i.e. income, employment or working conditions (Westergaard-Nielsen, et al. 2005), (2) family relations (Wilson and Oswald, 2005). These additional controls are also likely to be endogenous to the education decision and are thus ignored in our favoured reduced form model.

Overall, we find that education reduces the risk of poor mental health. The impact is observed for all ages and at all points of the distribution. Education also reduces the risk of becoming depressed. The effects are larger for women than for men, and at mid-level of qualification. These conclusions are not altered when assuming the endogeneity of education and in most cases the protecting effect of education on mental health increases (in absolute value) when assuming the endogeneity of education. The benefits of education are not solely due to income, family or work effects and remain even after accounting for these factors (with the caveats that these factors are potentially not exogenous). These results suggest that investing in education has long term causal benefits on mental health. Due to the high costs of poor mental health, education policies thus have large additional returns.

The outline of the paper is the following. The first section provides a rapid overview of the literature. This is followed by a discussion on the various identifying strategies including instrumental variables in the context of probit, zero inflated negative binomial and quantile regression. The British data used for the empirical work is described in section 3 and the results on depression are discussed in section 4. The following section compares the results over the life time of the individuals whilst section 6 presents the estimates for the continuous measure of mental health. Robustness checks using alternative measures of mental health are discussed in section 7 and we conclude with some policy recommendations.

## 1 Literature

There is scant evidence on the effect of education on depression. The concern of most epidemiological and etiological research in the field has been with the influence of genetics (e.g. Zubenko et al., 2003), gene-environment interactions (e.g. Silberg et al., 2001), personality style (e.g. Kendler et al., 1993), prior history (e.g. Lewinsohn et al., 1988), adverse life events (e.g. Mazure et al., 2000), age and gender. Relative to these important features, education may operate only at the margin. However, Kubzansky et al. (1999) have noted that more qualified individuals are significantly less at risk of bad mental health outcomes or to suffer from long-term stress.

Like for other health outcomes, the effect of education may be direct or through its impact on a third variable. For example, a higher occupational grade is associated with greater income, more control over the working life, and with more varied and challenging work and thus reduced morbidity (Marmot et al., 1991) but also higher levels of stress. Additionally, education increases earnings, reduces the probabilities of unemployment and divorce (Jalovaara, 2002), factors affecting the risk of mental illness. Thus the relationship between educational success and mental health is complex and likely to be non-linear. Moreover, the causality between education and mental health could be reversed; Currie and Stabile (2004) or Heckman et al. (2006) provide evidence of the impact of non-cognitive skills such as attention, self-esteem and locus of control on educational attainment.

More importantly, the literature is concerned that the observed correlation between health and education does not reflect a causal relationship. For example, individuals with a higher discount rate would invest less in their education and less in their health. Not accounting for self-selection would bias the estimate of education on health upwards. However, focusing on reduced form models, the causality of education on non-mental health outcomes is clearly established; after the pioneering work of Berger and Leigh (1989), most recent research has used educational reforms to identify the causal effect of education on birth weight (Currie and Moretti, 2003), mortality (Lleras-Muney, 2005) and smoking behaviour (Kenkel *et al.*, 2006), all finding a positive effect of education on the health outcome of interest. Moreover, contrary to the self-selection hypothesis the estimates of education on health outcomes are not reduced when assuming the endogeneity of the educational decision (Grossman, 2005) either because the instruments identify a local average treatment effect rather than the mean effect. Alternatively, this may indicate that unobservable characteristics are positively correlated with educational attainment and negatively with mental health; individuals with a greater propensity for education also have more risk of bad mental health in the future.

## 2 The identification problem

Estimating the effect of education on adult depression is a typical problem of programme evaluation. Here the treatment is defined as the level of education. It can be seen as a multiple treatment, varying in intensity (years of education) or quality (qualification). To simplify the presentation, and the empirical analysis, we concentrate on the simpler case of a single treatment, which in the following section will be defined as having at least O-levels as the highest qualification[1].

We assume that an indicator of mental health measured without error exists. We introduce the following notations of mental health status (dichotomous or continuous): $Y_1$ and $Y_0$ associated with the higher and lower level of education respectively. D is an indicator of the educational level attained. The parameter of interest is the average treatment effect on the treated (ATET).

$$ATET = E(Y_1 - Y_0 / X, D = 1) = E(Y_1 / X, D = 1) - E(Y_0 / X, D = 1) \qquad (1)$$

Without relying on an experimental set-up, it is not possible to assume that self-selection into the treatment is unimportant ($E(Y_0 / X, D = 1) \neq \mathrm{E}(\mathrm{Y}_0 / X, D = 0)$). Heckman et al. (1997) decompose the selection bias into three components: $B_1$ is the bias that occurs due to lack of common support (educated and non-educated individuals have different observable characteristics), $B_2$ arises from different distributions of X between the two populations on the common support, and $B_3$ is due to differences in outcomes that remain even after conditioning on observables and being on a region of common support (due to selection on unobservable characteristics). In the absence of experimental data we rely on various identifying assumption to recover the value of the counterfactual outcome.

[1] O-levels are a national examination taken at 16, which determine whether the pupil can follow the academic curriculum and counts towards admission to university.

## 2.1 Homogeneous returns model

We assume the following simple model of the determinants of mental health for individual i at period t ($Y_{it}$), where S represents the level of education, $X_1$ is a set of individual characteristics from periods concomitant or pre-dating the schooling decision, whilst $X_2$ is a set of individual characteristics posterior to the schooling decision that may potentially dependent on S.

$$Y_{it} = \beta_0 + \beta_1 S_{it} + \beta_2 X_{1i} + \beta_3 X_{2it} + \mu_i + \nu_t + \varepsilon_{it} \qquad (2)$$

The parameter of interest is $\beta_1$. If $X_2$ is excluded, $\beta_1$ represents the total effect of education on mental health. When $X_2$ is included, $\beta_1$ estimates the direct impact of education on mental health after accounting for the alternative channels of influence (assuming that S and $X_2$ are orthogonal). Both parameters are of interest to the policy maker. In the first case, we measure the overall social returns to education. In the second, the mechanisms by which education impacts on mental health are highlighted. However, the estimated coefficients $\beta_1$ and $\beta_3$ will be biased if the variables $X_2$ are correlated with unobservable characteristics explaining the choice of education and the health outcome of interest. Hence we favour the reduced form model.

If a third characteristic affects both educational attainment and adult depression, these models lead to biased estimates of the effect of education. However, if this third characteristic is fixed over time (i.e. genetic), early measures of mental health will absorb this effect. The estimated model is then:

$$Y_{it} = \beta_0 + \beta_1 S_{it} + \beta_2 X_{1t} + \rho Y_{i0} + \mu_i + \nu_t + \varepsilon_{it} \qquad (3)$$

Alternatively, modelling changes in mental health over time eliminates the bias due to time-invariant unobservable characteristics. All covariates are in levels as childhood characteristics are constant and only a few individuals gain qualifications as adults.

$$\Delta Y_{it} = \beta_0 + \beta_1 S_i + \beta_2 X_{1i} + \eta_{it} \qquad (4)$$

Since the focus is on adult mental health, there is no simultaneity in the timing of both outcomes and any remaining bias can only stem from unobservable variables affecting the education decision at time t but mental health at time t+1.

### *2.2 Instrumental variables*

Instrumental variables can be used to recover a consistent estimate of $\beta_1$ in the presence of an endogenous schooling decision. Assuming that S can be estimated linearly, we estimate the following model, were *z* is a vector of variables correlated with education decision but orthogonal to mental health:

$$S = z\beta + \gamma X_1 + v_2 . \tag{5a}$$

$$Y = \beta_0 + \beta_1 \hat{S} + \beta_2 X_1 + \beta_3 X_2 + \varepsilon_2 \tag{5b}$$

The choice of instruments is discussed in the data section. This identifying strategy eliminates the bias-$B_3$. Unobservables increasing the risk of mental health problems are likely to be negatively correlated with education (Currie and Stabile, 2004), hence the IV estimates are expected to be lower (in absolute value) than the estimates assuming the exogeneity of the education decision.

### *2.3 Zero Inflated Negative Binomial*

Most of the literature on mental health has focused on a dichotomous outcome flagging individuals with the most severe mental health problems. These individuals are the most likely to face the highest cost of mental illness, but it is also informative to estimate the effect of education on the intensive margin of mental health. While linear models are reported, malaise score can only take a limited set of values and can be thus viewed as a count variable. Due to the over-dispersion of the score, note that the variance of the score is at least three times larger than the mean (see Annex 1), the count is modelled using a negative binomial model which relies on a Gamma distribution[2]. Moreover, the distribution is characterised by an over-representation of zeros which may be deterministic or due to chance. The malaise score is thus modelled using a Zero Inflated Negative Binomial (ZINB) which can be formally written as:

$$\begin{cases} f(y_i \mid x_i) = \psi_0 + (1-\psi_0)g(y_i \mid x_i) & \text{if } \mathrm{y_i} = 0 \\ f(y_i \mid x_i) = (1-\psi_0)g(y_i \mid x_i) & \text{if } \mathrm{y_i} > 0 \end{cases} \tag{6}$$

where $g(t) = {\phi(t)}\big/{1-\Phi(t)}$; zeros are over-represented as they occur with an additional probability $\psi_0$ over and above all other outcomes[3]. Wooldbridge (2002) shows that when one of the variables (say S) is endogenous, it is possible to obtain consistent estimate of $\beta_1$ by including $\hat{\upsilon}_2$ (from 5b) in the ZINB model. A t-test on the coefficient for $\hat{\upsilon}_2$ can be used as a test for the exogeneity of $X_2$.

### 2.4 Quantile regression

Linear models are informative to describe the relationship of interest at the mean, but when the effects are non homogenous, it is also informative to estimate the impact of the covariates throughout the distribution using quantile regression (Koenker and Bassett, 1978). Moreover, Chernozhukov and Hansen (2005) have proposed a method to estimate quantile regressions when one of the covariates is endogenous. This section draws heavily on their work.

Koenker and Bassett (1978) define the Conditional Quantile Function $Q_Y(\tau / s, x)$, when assuming that all covariates are exogenous. Y can be defined as: $Y = S'\alpha(U) + X'\beta(U)$, where U is a random component. $Q_Y$ satisfies $P[Y \le Q_Y(\tau / S, X)/ S, X] = \tau$. Chernozhukov and Hansen (2005) show that when an instrument *Z* for the endogenous variable *S* exist the Structural Quantile Function is defined as, $S_Y(\tau / s, x) = s'\alpha(\tau) + x'\beta(\tau)$ such that $P[Y \le S_Y(\tau / S, X)/ Z, X] = \tau$. Their solution is robust to any functional form of the underlying process, but the identification imposes the rank similarity of the dependent variables across the values of the exogenous covariates.

[2] Whilst count data are usually used to model process that occurs through time, fertility decision, number of incidents, this is not the case here. Thus, the usual limitation of the negative binomial model being compatible with duration dependence in the underlying process (the Poisson assumes it is fixed) or individual heterogeneity is not present here.
[3] Φ and Φ are the probability density and cumulative density functions of a normal distribution respectively.

Assuming that individual A has some unobserved characteristics reducing her risk of mental illness compared to individual B, then at all levels of education, A will have a lower mental health score than B, even if the difference in score may change with education.

## 3. Data and definitions

The empirical analysis is conducted using the National Child Development Study (NCDS). The NCDS is a longitudinal study of all the children born in Britain during a given week of March 1958 (14,746 observations). These children and their parents were followed up at age 7, 11 and 16. Each wave also included medical and school questionnaires. Members of the NCDS cohort as well as their families were interviewed at age 23, 33 and 42, when 11,419 individuals were surveyed[4]. These adult questionnaires also include a health supplement used to calculate the malaise score and depression. The malaise score was designed to identify depression in non-clinical settings and has been found to be a good indicator of depression (Rutter et al., 1970). It is calculated from responses to 24 questions on various aspects of well-being and somatic conditions (See Annex 1)[5]. Following the psychological literature, individuals scoring 8 or more positive symptoms are coded as depressed. Since the questions do not directly ask about mental health, the malaise score is likely to be less biased by stigma and misreporting than other, more direct measures. In fact, the malaise score tends to over-predict clinical depression (Meltzer et al., 1995), but is well correlated with diagnoses from clinical interviews (Stansfeld and Marmot, 1992) or contact with mental health services (Lindelow et al., 1997).

[4] To create the final sample, we drop individuals with no information on maternal characteristics throughout childhood and appropriate education and health characteristics, leaving us with a sample of 6,666 individuals at age 42. Almost all of the missings observations are due to missing observations on parents. However, the difference in adult depression rate between the final sample and the omitted individuals is never significant.

[5] A principal component analysis reveals that the components of the malaise score are essentially orthogonal to each others. The first component account for 20% of the variance, and eight factors are needed to reach 50% of explained variance. We therefore follow the literature and give equal weight to each item of the score.

Annex 1 reports the answers to each component of the malaise score for the three adult waves. The most common predicament is to admit to “often worry about things” with at least a third of individuals giving this answer. The evolution of malaise score and depression over time is plotted respectively in Figure 1A and 1B. Malaise scores are skewed with a large proportion of individuals having a score of 0 and a fine tail of individuals with high score. The most noticeable pattern is that by age 42, the distribution becomes less skewed with a drop in the 0-frequency of 15 points and a fatter tail such that the probability of depression doubles and triples for women and men respectively; at each age, men are about 4 percentage points less depressed than women. This increase could be consistent with a reduction in the stigma associated with depression in the last decade. However, since we do not rely on a self assessment of depression and there is no change in the measurement of the score, the increase in score is likely be a mixture of age and a time effects[6]. For individuals in their forties the average malaise score increases by a full point. The main factors responsible for the increase in the scores are: feeling tired (+78%), feeling miserable (+76%), difficulty sleeping (+81%), wake up early (+90%) and worry about health (+137%). Additionally, the probability of having had a previous breakdown increases three-fold between the twenties and forties. These statistics are broadly in line with the psychiatric and morbidity survey (Singleton et al., 2001).

The correlation in depression (malaise score) between any two periods ranges from 0.29 to 0.36 (0.43 to 0.59); no gender differences are observed in these correlations. These high correlations highlight the persistence of malaise score over time and suggest that the malaise score truly captures some permanent features of individuals’ mental health. The malaise score includes a set of question that are directly related to somatic rather than psychological trauma, which could generate a correlation between score and occupation, if for example, manual workers are more

[6] There is no consensus in the psychological literature on the age profile of depression, see Jorm (2000) for a review. According to the psychiatric morbidity survey, the prevalence of depression in GB amongst the 35/44 age group increased by 3 percentage points between 1993 and 2000. Moreover, in a given wave, individuals in the 35/44 age group are almost 2 percentage points more likely to be diagnosed as depressed than those aged 25/34.

likely to suffer these somatic ailments. This would bias upwards the estimated effect of education on mental health. Rodgers et al. (1999) report that "the 15-item psychological sub-scale is no better in discriminating those with significant psychiatric morbidity or those who use mental health services, than the full scale" (p339) but robustness checks using the non-somatic scale, as well as each individual ailment are reported in the final section.

Annex 2 provides evidence of the validity of the malaise inventory as a measure of mental health problems, by comparing it with a self-reported measure of currently feeling sad, low or depressed. Table A2-1 reports malaise scores and depression status for each category of the direct mental health measure. Individuals who describe themselves as depressed most days have a malaise score of 11 and 75% are classified as depressed; these statistics are significantly different for individuals who report to never have been depressed. In fact, a clear monotonic relation is observed between self assessment of depression and malaise score. In Table A2-2, we assess whether all items in the malaise score are associated with depression and calculate for individuals who admit to the ailment and those that do not the proportion reporting to be depressed occasionally or most days. Even for ailments that a priori are not related to mental health the between group differences in scores are always significant. For example, 16.5% of individuals reporting backache admit to be depressed but only 9.8% of those without back problems are depressed. The full malaise score thus appears to be an appropriate measure of the current mental health of individuals.

Figure 2 plots the main relation of interest and reports the level of mental illness in terms of the highest qualification[7][8]. At all ages, a steep education gradient is observed, especially for women, which tails off with higher qualification. For both genders, the difference in the probability of depression between the most qualified and least qualified individuals is four folds. The relation

[7] The pattern when using malaise score rather than depression is also similar.

[8] CSE and equivalent are mostly non academic qualifications attained at age 16. O-levels is an academic secondary education qualification also obtained at age 16. A-levels is the end of secondary education qualification.

is monotonic but for male graduates aged 42 who have a higher risk than A-level holders. This may be an indication of job related stress in occupations requiring a degree as the career progresses. For all levels of education, men are less likely to be depressed but the gender depression gap decreases with age and qualification, so much so that by age 42 the gender gap in depression between university graduates is only 1 percentage point.

Summary statistics on the control variables included in (2) are summarised separately by gender in Table 1. Education is measured by the highest qualification attained rather than years of education. Women are less likely than men to have no qualification (17% vs. 20%) but men are 4 percentage points more likely to have a university degree. Since malaise score is not available in the children's wave, we control for time fixed effects in mental health using the externalising and internalising scores of children at age 11, as well as an indicator of character provided by the parents at age 16. The age 11 indicators reflect the child's dominant behavioural responses to an unobserved emotional disturbance (Rutter, 1967); externalisers are acting up or behaving in a way linked to conduct disorder, whilst internalisers tend to go silent and are prone to depression (Feinstein and Bynner, 2005). These behavioural patterns are measured at age 11 by the parents[9]. At age 16, parents assess the behaviour of their child using 18 questions. Each question gauges the intensity of a characteristic (on a 3-point scale) such as does the child appears irritable, miserable, destroy things. There are important gender differences in external score but more limited one in internal and character scores. Additionally, at age 16, we can control for bad mental health. Girls are twice as likely as boys (3% vs 1.5%) to have missed school for more than a week due to nervous problems, but boys are more likely to have seen a specialist doctor for emotional problems. It is important to control for these early symptoms of bad mental health as they affect both current educational attainment and future mental health. Moreover, early test scores proxy the child ability

[9] The internalising score is measured as the sum of the following: child prefers doing things alone, child miserable, child worries about many things, child upset by new situation, whilst externalising score is the sum of: difficulties settling into anything, child destroys things, child squirmy, child irritable, child fights, child disobedient. The scores are normalised.

which will affect her school performance but also her mental development; low ability at a young age is correlated with psychological disorder (Beck, 1987). Reading and Math scores at age 7 are thus included as proxy for ability with girls being better readers but worse in Maths than boys.

For controls on family and school limited gender differences are found. The family characteristics include parental age at the child's birth, parental education, their interest in the child's education as perceived by the child's teacher, and the paternal social class (based on occupation), as well as indicators of whether the family experienced financial difficulties when the child was age 11 and 16. At age 16, we include an indicator of whether the child lives with parental figures as well as measures of the quality of these relationships in the view of the child. Whilst 73% of children get on well with their mother, the proportion getting on well with their father is only 67% for boys and 63% for daughters.

The final set of environmental factors concerns neighbourhood and the peers as they may have long-term effects on mental health[10]. To capture them the type of school attended at age 16 and the proportion, broken down by gender, of pupils in that school who remained in post-compulsory education are included, so as to broadly measure the "quality" of peers.

## 4 Results

The introduction advocated that economists should investigate mental health since it impacts on productivity. We first assess this statement; these estimates are not meant to be causal effects but simply stress the correlations between mental health and labour market outcomes (participation and wages). The top panel of Table 2 reports results for depression whilst the bottom panel uses normalised malaise score as a measure of mental health. Individuals with worst mental health are significantly less likely to be participating in the labour force; increasing the malaise score by one standard deviation, reduces participation at age 42 by 3 to 6 percentage points and

[10] Evidence on neighbourhood effects on depression only concern adults; Propper et al. (2004) using the BHPS and relying on a precise measure of neighbourhood (500 individuals) do not find systematic effect of neighbourhood characteristics (Disadvantage, mobility, age, ethnicity and urban-ness ) on mental health or mental health transitions.

being depressed has an even larger impact. There is also some evidence that this correlation increases with age. Conditioning on participation, depression has no significant effect on wages whilst higher malaise score leads to a small but significant reduction in wages; individuals with even faint mental health problems are less productive[11]. It is thus important to analyse both the extensive and intensive margins.

### 4.1 The basic model

As in the rest of the literature, we first focus on the effect of depression. Since the beneficial effect of education may be additive over the individual's life time, we first model the determinant of mental health at age 42; the marginal effects of the covariates on the probability of depression are reported in Table 3. The first two columns report, separately by gender, the estimate of qualification on depression in the base model when education is assumed to be exogenous.

There is persistence in the propensity to depression over the life-time with measures of mental health at age 11 and 16 both significantly correlated with adult depression. Internalising girls and externalising boys are at greater risk of adult depression, as are boys who have had treatment for emotional problems. This is consistent with the assumption that some of the unobserved characteristics leading to depression are time-independent. Whether teenagers live in an intact family does not affect their risk of adult depression but the quality of the relationship with the parents does matter. Getting on well with the opposite sex parent reduces adult depression by four to five percentage points. Finally, for women, financial hardship significantly increases the probability of adult depression; experiencing financial difficulties at age 16 increases depression by 7.5 percentage points from a baseline of 16%.

Education reduces the risk of depression for both genders. For women, having a CSE reduces the risk of depression by three percentage points, each additional qualification up to A-levels reduces this probability by a further two percentage points. Higher education does not lead to

[11] This contradicts evidence from Westergaard-Nielsen et al. (2005) who reports substantial negative wage impact of depression, but the difference may be due to self-selection bias which is not accounted for in the present study.

a significant further reduction in depression compared to A-levels. For men, the pattern is different. Having a CSE does not lead to a significant reduction in adult depression, but reaching O-levels leads, as for women, to a reduction of about five percentage points. For men, the effect of education is non-monotonous, and whilst A-level leads to a further reduction in the probability of depression; university graduates are at the same level of risk as individuals with O-levels.

Compared to the base line, education has substantial effects on the probability of being depressed, reducing the average risk by 50% for the highest qualifications. However, getting on well with the parent of the opposite sex, and for women, not having experienced financial hardship at age 16 have effects equivalent to having gained A-levels rather than no-qualification. While the effects of education on depression are significant, family characteristics when growing up are at least as important.

As discussed previously, these estimates assuming the endogeneity of education are potentially biased[12]. Since all the children are born in the same week, it is impossible to find change in legislation that could be used to identify the effect of education. Instead, the identification comes from two sets of instruments. First, the teacher's expectations concerning the schooling of the child or, to be more precise, the answer to the following question "in your opinion would staying-on past compulsory schooling benefit the child?" Figure 3A reports this statistic by highest qualification attained, which varies monotonically from less than 10% for individuals with no qualification to more than 70% for individuals with A-levels or higher. The teacher view is clearly related to academic achievement and there is no evidence that after controlling for child mental health and ability, the teacher view is directly correlated to adult mental health since this opinion was never directly communicated to the child[13].

The second set of instruments is based on proxying the discount rate of pupils. The discount rate is correlated with the decision to invest in post-compulsory schooling and individuals

[12] In regressions excluding previous mental health measures (not reported here), the estimated effects of qualification was between 0.5 and 2 percentage points higher than in the reported model.
[13] The teacher view is never a determinant of adult depression in regressions also controlling for highest qualification.

with a greater preference for the present will invest the least. Smoking is used to proxy the discount rate; smokers have low time preference, so that the more a pupil smoked at age 16, the higher her discount rate[14]. Figure 3B reports the proportion of smokers by highest qualification whilst 3C plots the average smoking intensity of smokers by qualification. For both measures of smoking, we find a significant gradient by highest qualification achieved. The proportion of smokers drops from 40% for individuals with no qualification to 27 for individuals with O-levels, whilst the number of cigarettes smoked per week for men with these two qualification levels are 38 and 25 respectively. As an instrument, we use the number of cigarettes smoked per week at age 16, recoding non-smokers to 0.

The discount rate is correlated with health investment (Fuchs, 1982) and is not an appropriate instrument for general health. However, since preventive measures are not as readily available for mental health we argue that the discount rate is independent of mental health. Moreover, Duncan and Rees (2005) conclude that after accounting for the endogeneity of smoking there is "little evidence that smoking intensity is related to the depressive sympomatology of smokers".

Finally, the instruments do not identify reaching a given qualification but only remaining in academic track for which O-levels or equivalent is a required qualification (see Figures 3A and 3C). O-levels are an important qualification in the British system which entitles to remain in the academic track post compulsory education and study for A-levels which are required to attend university. Additionally, the exogenous estimates suggest that for men, the effect of qualification is non-linear and that O-levels is the first qualification that significantly reduces depression. We re-estimate the exogenous model when defining education as having at least O-levels (Column 3 and 4 of Table 3). Having at least O-levels reduces depression by about 5 percentage points for both men and women. Allowing for the endogeneity of education, the reduction in depression due to

[14] Smokers choose occupations with lower wage profile (Munasinghe and Sicherman, 2000) but smoking is assumed to be independent of wages, thus can be used as an instrument for education in a wage regression (Evans and Montgomery, 1994).

education is further strengthened. These results refute the self-selection hypothesis and, as for other health outcomes, accounting for the endogeneity of education increases its effects (Grossman, 2005). It could thus be that measurement error in mental health are correlated with qualification or that individuals investing in education post-16 would be at greater risk of depression in the absence of this investment; for example individuals with a greater ability or taste for analysing problems are more likely to invest in education but may also be more self-critical which leads to a greater risk of depression[15]. Alternatively, the instruments mostly identify individual with a lower taste for schooling for which the preventive effect of education on mental health may be larger than in the general population (Local Average Treatment Effect).

The instruments are individually and jointly significant in the first stage equation, and pass the rule of thumb for weak instruments with F-test of joint significance around 200[16]. The over-identification also shows that the instruments are valid[17] however, the exogeneity of education can only be rejected for women. The estimates assuming exogeneity of education provide thus a lower bound of the effect of education on depression. Using a continuous measure of mental health leads to similar conclusions (estimates can be found in Table 4 panel B); in the endogenous model the estimates increase 3-folds for women and by 50% for men.

### *4.2 Routes through which education may affect depression*

The reduced form is now extended to include additional current characteristics and assess whether the effect of education is direct or through these variables. This vector of current characteristics (at the age when depression is measured) is broken down into two components: job

---

[15] For Feinstein and Bynner (2005) children with greater internalising behaviour are more likely to stay on in school and more prone to depression as adults since internalising behaviour is associated with low self-esteem and poor peer relations.

[16] Bound et al. (1995) recommend that a F-test on the joint significance of the instruments in the first stage equation be greater than 10, in order not to suffer from weak instruments problems.

[17] Estimates using the two instruments separately are in the same range as those reported in this model.

and family[18]. Including these covariates takes out the partial correlation between schooling and depression that stems from them. Careful interpretation must be made since these characteristics are measured after the education decision was taken and can be considered endogenous. The base specification provides an estimate of the total effect of education on depression whilst these subsequent specifications are intended mostly to indicate the possible process by which education affects mental health.

Results are reported in Table 4 for the base model and 3 models including additional variables. Panel A reports estimates when mental health is measured by depression whilst panel B focuses on the malaise score. Model 1 adds work characteristics including income, labour market experience, current employment status to the base model. The effect of education on mental health is reduced by 20% to 40%, depending on gender and whether education is assumed exogenous. It can also be noted, that for men we cannot reject the hypothesis that the IV estimates are not significantly different from the estimates assuming exogeneity of education. This is true in all subsequent models.

Model 2 controls for current family characteristics. The issue here is that education may alter fertility decisions (leading to a reduction in number of children, older age at first birth) and marital situation (assortative mating) which potentially affects the depression probabilities (Wilson and Oswald, 2005). The family characteristics include marital status, number of children, age when first child was born, age of the youngest child and an indicator for the death of a child. The estimated effects of education on depression remain close to those obtained in the base model. The effect of education on depression does not work through the current family characteristics.

Finally, Model 3 includes both sets of additional variables. In the exogenous model, the estimates on education are about 50% smaller than in the base model and remain significant only

[18] Current job characteristics include the following: wage (in logarithmic form), labour market experience since 16, whether currently working, whether works full-time or part-time and a measure of self-employment. The family component reports marital status, whether had any children, number of children, age at first birth and age of the youngest child.

for men. However, the estimates assuming the exogeneity of education are only about a third smaller than in the base model, and for women remain significant. The effect of education on mental health is not solely due to an income, labour market or family effects but remain substantial even after accounting for these characteristics, indicating that the impact of education on depression could be direct: education per se reducing the risk of bad mental health.

As highlighted previously, it is also interesting to estimate the effect of education on malaise score and not only on depression. The effects here are somehow more limited than in the depression model; having a qualification above O-levels reduces the malaise score by 0.15 standard deviation for both men and women. Assuming the endogeneity of education increases this effect by 50% for men and three fold for women, as in the model using depression as the dependent variable. Adding work and family covariates has a similar impact on the estimates as in the depression model. In all models, the IV estimates are larger (in absolute value) than those obtained assuming the exogeneity of education. The IV estimates are significant only for women, for men the point estimates are also larger but imprecise. Education has thus a direct effect on mental health both at the intensive and extensive margin[19].

## 5 Depression over time

Tables 5 assess the evolution of the education effect over time. More specifically, we focus on depression at age 23, 33 and 42 for models with and without contemporary controls[20]. For women the effect of education in the base models assuming exogeneity of education is remarkably persistent over time. Estimates are much larger when the endogeneity of education is assumed and increase with age as the base line depression rates increases from 9% to 17% over the period. As

[19] Since several models reject the endogeneity of education, the assumption that the selection into education is based on observable characteristics may be acceptable and the effect of education on mental health was also estimated using propensity score matching. Gaining O-levels reduces the risk of depression by 5% to 6% or decreases malaise score by 0.13 to 0.30 of a standard deviation. The estimates obtained by matching are comparable to those obtained when imposing a functional form. Thus imposing a functional form to the estimator does not lead to a substantial bias and only parametric estimates are presented subsequently.

[20] Estimates were also computed using malaise score and propensity score matching estimator. As was the case at age 42, the results were consistent with those presented.

was already observed at age 42, the estimates are dampened but remain significantly different from 0 when work or family characteristics are included. However, when these covariates are included the endogenous model is generally rejected due to the lack of precision of the estimates at age 23 and 33. For men, the effect of education on depression appears U-shaped in age, with estimates at age 33 not being significantly different from 0. Adding contemporary controls reduces these estimates by about 30% but at age 23 and 42, they remain significant. Estimates assuming the endogeneity of education are much larger but the endogeneity of education can only be accepted at age 23. For men, education seems to have a large protecting role to play mostly in the early years of adulthood.

Finally, to assess whether our results are driven by unobserved heterogeneity, depression transition models are estimated at age 33 and age 42 (Table 6). Since most individuals depressed in one wave are still depressed 10 years later, the only transition observed is to become depressed. Keeping only individuals who were not depressed in the previous wave, we estimate a probit model on the probability of having become depressed.

For women each additional qualification significantly reduces the risk of becoming depressed. The patterns are similar in both transitions with a peak for achieving A-levels, but the effect of qualification almost doubles between the first and the second transitions (the base line also doubles from 5% of women becoming depressed at age 33 to 11% at age 42). Surprisingly, when defining education as having O-levels or more, no significant effect is found either in the endogenous or exogenous models; this could indicate that the grouping of no qualification and CSE is not appropriate, and biases the estimated effect of having upper qualifications downwards. Note nevertheless than as in the stock models presented above, the estimated effect of education assuming endogeneity is four times larger than the exogenous effect. Since, the exogeneity of education cannot be rejected, one could conclude that the first model is more informative, and each additional qualification reduces the risk of becoming depressed over a 10 year period.

For men, no effect of education on depression transitions is found for the first transition, maybe because of the small proportion of depressed men in these two waves (only 2% of men become depressed between 23 and 33). Nine percent of men become depressed between the age of 33 and 42 and having O-levels or A-levels significantly reduces this risk by 3 to 3.5 percentage point. Having any qualification at O-levels or above significantly reduces the risk of becoming depressed by the age of 42 by between 3 to 4 percentage points; again estimates assuming the endogeneity of education are imprecise. Education appears to improve individual's mental health level through out their life-time but also reduces the risk of becoming depressed over time.

## 6 Mental health as a continuous measure

The analysis has mostly focused on the tail of the mental health distribution, estimating the probability that education reduces the malaise score so that the individual's score passes the adhoc threshold defining depression. Focusing on the intensive margin may largely underestimate the overall impact of education. So far, results using a continuous measure of mental health have broadly been consistent with those obtained for depression. However, a linear model maybe inappropriate as malaise score only takes a few values and can be considered as count data (counting numbers of ailments). Also as noted in Figure 1A, the malaise score distribution is heavily skewed, suggesting that estimates at the mean may not be informative. Moreover, the effect of education on mental health may be non-homogenous, it is thus of interest to estimate them at different points of the mental health distribution using quantile regression. For both models the endogeneity of the education decision can also be assumed.

### 6.1: Count data

The distribution of the malaise score is skewed with an over-representation of zeros and a thin long tail of individuals with high score. As explained above, this type of distribution can be estimated by a zero inflated negative binomial model, results of which are reported in Table 7 by gender and age. The results are similar to those presented above, with each qualification reducing the malaise score. For both gender and all ages, the effect of qualification is monotonous (except females at age 42) with more qualifications leading to a greater reduction of the score. Achieving O-levels or above reduces the malaise score by about 0.2 points or between 8% and 10% at the mean, in the exogenous model. The effects are similar by gender and decline as individuals age which contradicts the pattern found in the depression model.

Assuming the endogeneity of education, the estimates are two to three times larger than those assuming exogeneity, and with the exception of men aged 42, the exogeneity of education can be rejected. The first-step residuals are always positive, as previously assumed, and individuals with a greater propensity for education are more at risk of worse mental health. These conclusions are broadly similar to those presented when estimating the extensive margin.

**6.2: Quantile regression**

The last aspect to investigate is whether the effect of education varies with the intensity of mental illness. In the exogenous model, for each qualification, the impact of education increases with the quantiles, so that the effect of a higher qualification is three times as large at the top quartile than at the bottom for women and much larger for men. For example having a university degree reduces the malaise score by 1.5 point for a woman at the 75$^{th}$ percentile but only by 0.5 point at the 25$^{th}$ percentile (Table 8). The effect of qualifications are clearly not homogenous: having a CSE improves mental health only at the top quartile and no qualification significantly improves the mental health of individuals at low level of mental illness. Contrary to evidence based on the mean, a U-shape in the effect of education is found for both men and women for all quartiles. For a given quartile, the qualification that offers the greatest reduction in mental illness risk is having A-levels. The second best qualification, as mental health protection is concerned is having a higher degree, while O-levels and especially CSE offers much lower reduction in risk.

When relying on the dichotomous indicator, O-levels or above, the conclusions are similar. One can notice that the shape of the effect across quartiles is steeper for men, especially at the top of the mental health distribution; so education appears to be particularly effective at protecting individuals with the highest risk of mental illness. Assuming the endogeneity of education, the estimates are always larger but in general become insignificant. The two significant coefficients are obtained for women at the 3rd quartile and at the median for men.

As in estimates based on the mean, there is no strong evidence that estimates assuming the exogeneity of education are biased. Focusing on the exogenous models, education does not have a homogenous effect on mental health and individuals higher up in the mental health risk distribution have more to gain from qualifications. Since these individuals are those at risk of more dramatic and expensive form of mental illness, the financial returns to education could be large. The qualification offering the greatest risk reduction is A-levels.

## 7 Additional robustness checks

In this section we conduct a series of tests to assert that the impact of education on mental health is not due to differences in occupational choice. Llena-Nozal et al. (2004) also using the NCDS and the malaise score as a measure of mental health, report that in a fixed effect model, occupation has no effect on adult mental health for men but that professional women score about 2 points less on the malaise score than unskilled workers. As stated previously, one may be concerned that the malaise score contains somatic ailment that may be more likely to be correlated with education and occupation, thus biasing the estimates of education upwards.

As seen in Annex 2, the somatic ailments, as defined by Rodgers et al (1999) are correlated with mental health. To assess whether these variables bias our results, the model is estimated separately for each item of the malaise score[21]. Only estimates for O-levels and degrees are reported in Figures 4A and 4B, separately for somatic and non-somatic ailment. Statistically

[21] Only results of the base model on the effect of qualifications on depressions at age 42 are reported as this simple model was found not to be significantly biased by endogeneity issues.

insignificant coefficients were recoded as 0 to simplify the figure. Assuming that somatic ailments are more likely in manual occupations, we should expect a larger impact of education on somatic ailments (denoted by a star on the graph). With the exception of backache (1*), there is no evidence that this is the case, and for a number of non-somatic ailment, similarly high impact of education is found, especially for women.

Additionally, a malaise score where all somatic ailments have been taken out is computed and used to replicate the linear model presented in Table 4. The results presented in Table 9 are similar to those obtained with the full score. Qualification has a U-shaped effect on mental health, the estimates are somehow larger for women than for men and for women the endogeneity of education cannot be rejected. These estimates are almost identical to those obtained when the full malaise score is used, apart from male in the endogenous model where both point estimates are different but insignificant. Thus, there is no evidence that using the full score – including somatic ailment - biases the effect of education upwards. These checks confirm the stability of the estimated impact of education on mental health which is not solely due to the inclusion of somatic ailments in the measure of mental health and can even be identified for most items individually. The effect of education is unlikely to be due to occupational choice differences.

## 8 Conclusion

We consistently find that education significantly reduces the risks of adult depression especially for women. The effect is non-linear and is larger at low to mid-levels of education. We estimate that having a secondary education qualification reduces the risk of adult depression (age 42) by 5 to 7 percentage points. The positive effect of education is present at all ages, possibly in a non-linear way, and remains even after accounting for work and family characteristics. The point estimates assuming the endogeneity of education are always higher than those that do not but for men the exogeneity of education is usually not rejected. The greater effects when assuming endogeneity are consistent with most of the literature on the effect of education on health and imply that unobservable characteristics explaining mental illness are positively correlated with education. The effect of education on mental health is direct. We also provide evidence that education reduces the risk of transition to depression. Additionally, education improves mental health even at low level of the malaise score, suggesting that the literature focusing on depression as an outcome underestimates the total effect of education on mental health. Finally, the estimates are not dependent on the inclusion of somatic ailments in the malaise score, so that the effect of education is unlikely to be solely through an occupational effect.

These results imply that policies increasing the education of individuals would have positive effects on their future mental health. It is difficult to calculate the cost benefit of such policies since the costs of depression, while assumed to be large have still not been satisfyingly identified and the costs of a policy increasing education can vary substantially. Here we conduct a simple calculation of the returns to education in term of improved mental health. Assuming that the costs of poor mental health in term of lost output are due solely to depressed individuals and in the most conservative case, that only women who originally had no qualification will see a reduction in their risk of adult depression at age 42 from 26% to 22%, that is a risk reduction of 15% for this population which represents 17% of the depressed individuals. We also assume that this reduction is permanent throughout the working life of the individual. Based on the estimated cost of £9 billion a year for the full population (Thomas and Morris, 2003) a policy increasing the education of females from no to basic qualification will reduce the total cost of depression for the population of interest by £230 million a year or £4.9 billion over the working life of these women, assuming a discount rate of 3.5%. Alternatively, if the probability of depression is not constant over the life time and is only about half as prevalent for the first 20 years of adulthood, the present value of such a policy would drop to £3.2 billion. These estimates can be considered under-estimates, as we have assumed no effect for men or other education group. This is an additional substantial return to policies increasing education for individuals with low level of achievement.

**Figure 1A: Distribution of malaise score over life time**

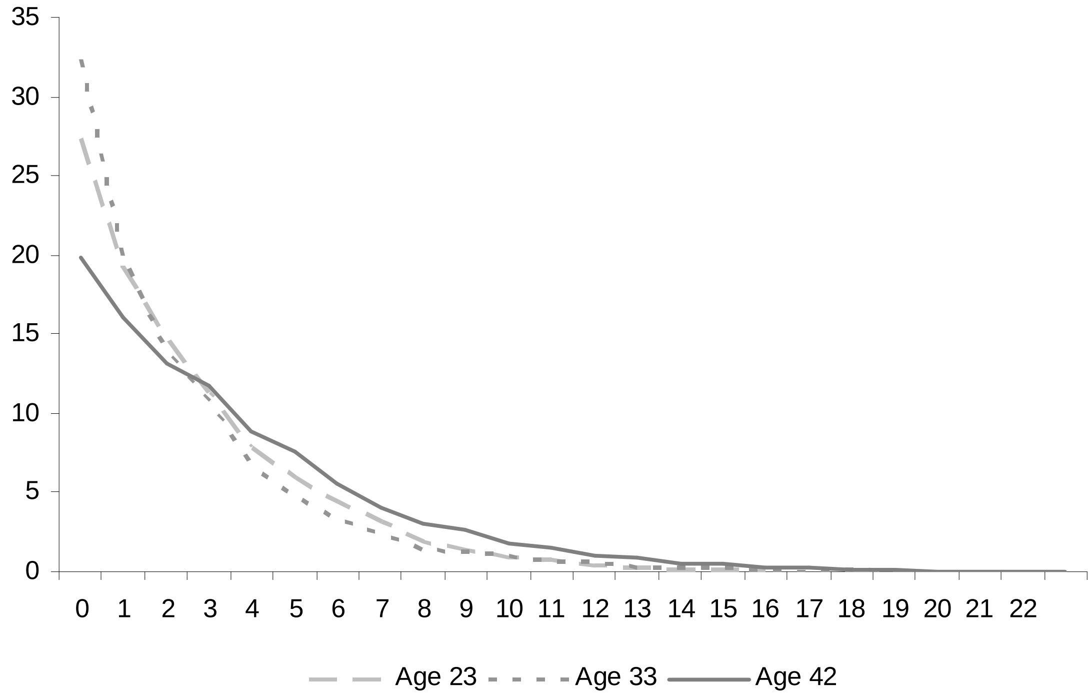



**Figure 1B: Depression over life time by gender**

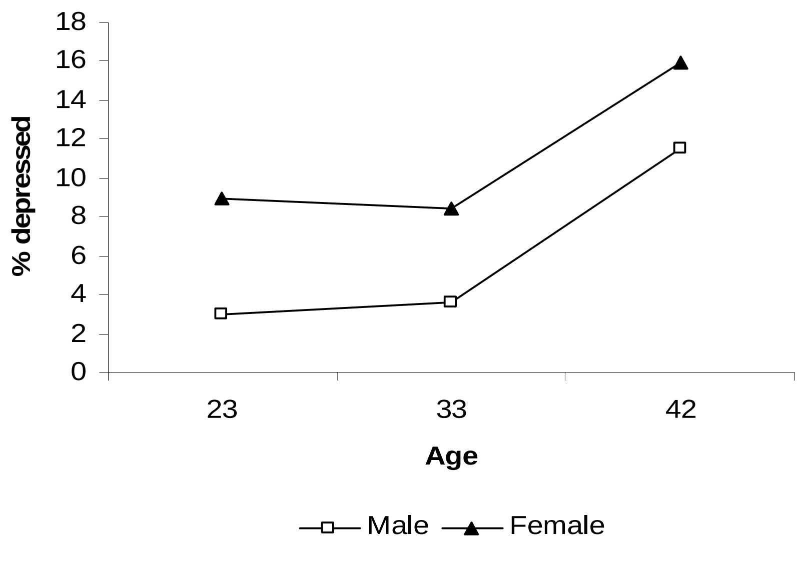
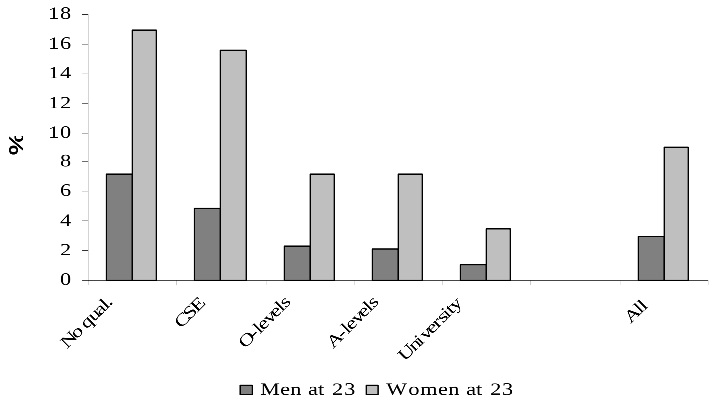
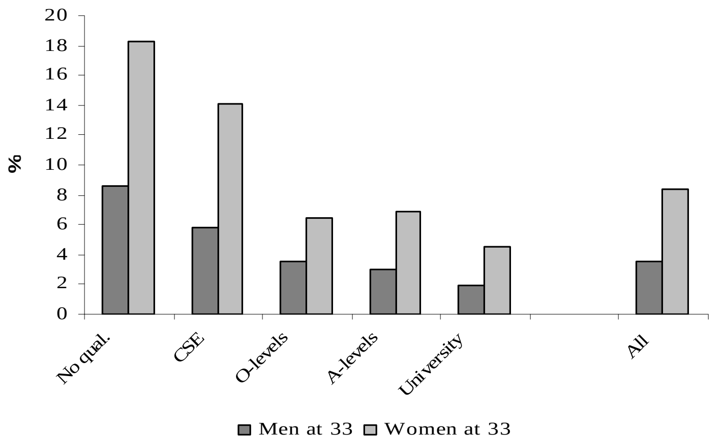


**Figure 2: Education level and depression at age 23, 33 and 42**

Panel A: Age 23




Panel B: Age 33




Panel C: Age 42

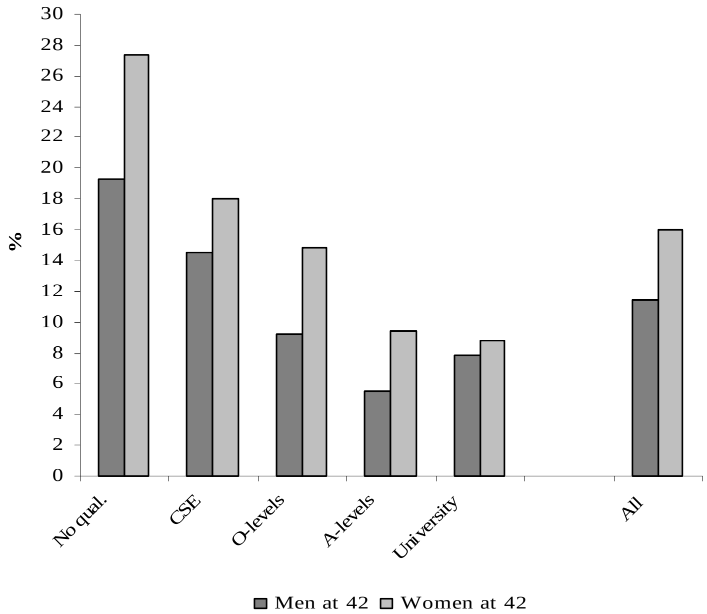


**Figure 3A: Proportion of pupils who would benefit from post-compulsory schooling, as assessed by teachers at age 16 by highest qualification at age 42.**

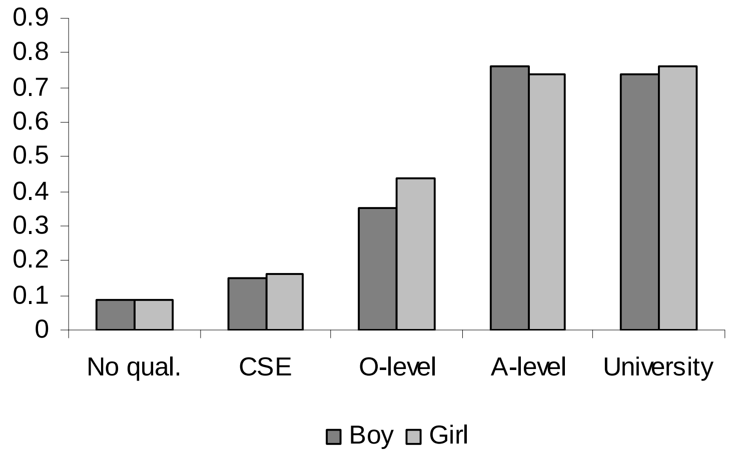



**Figure 3B: Proportion of smokers at 16 by highest qualification at age 42**

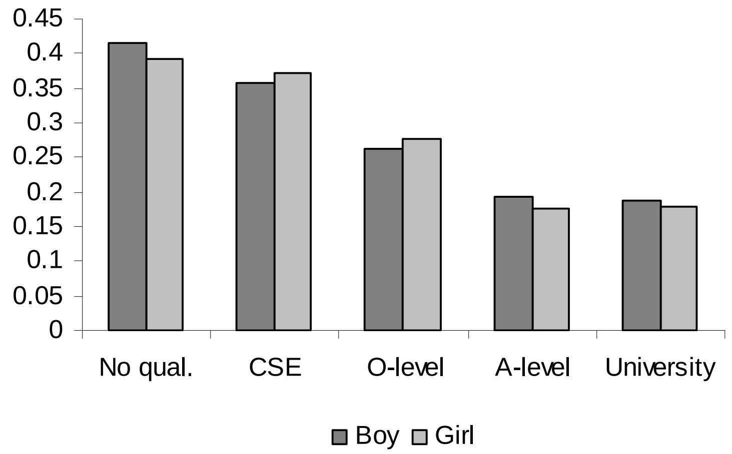



**Figure 3C: Intensity of smoking at 16 (cigarette smoked per week) by highest qualification at age 42**

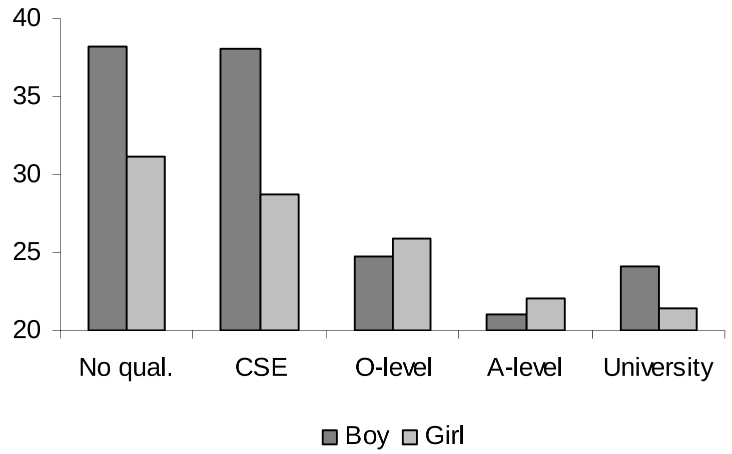


**Figure 4A: Estimate of the effect of qualification on the individual components of the malaise score: Men age 42**

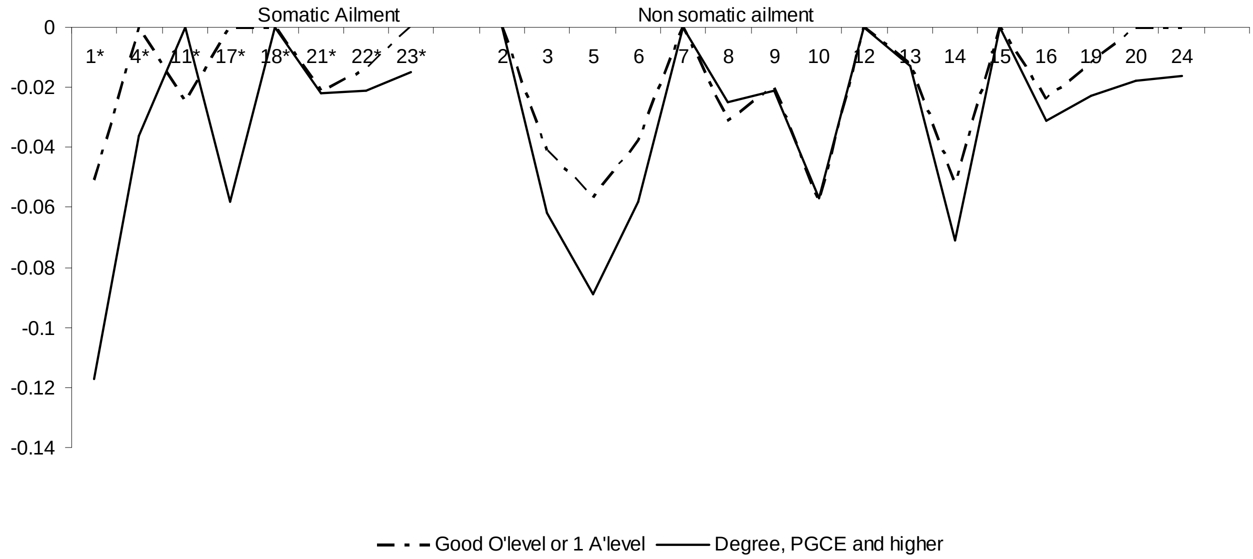
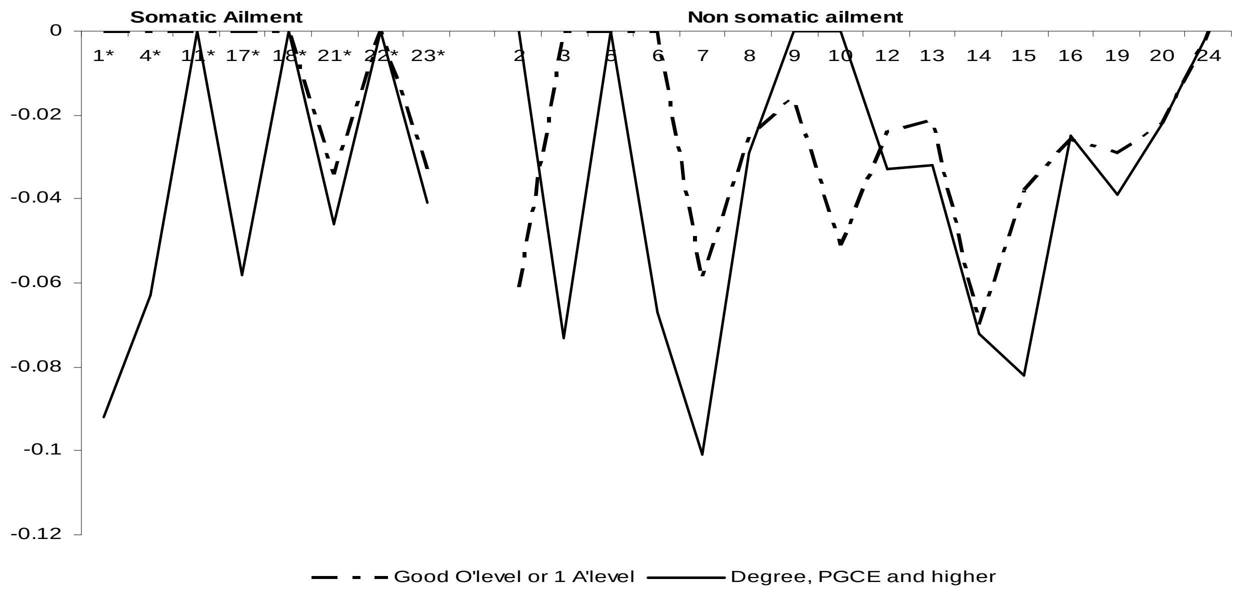



**Figure 4B: Estimate of the effect of qualification on the individual components of the malaise score: Women age 42**




Note: * denote somatic ailment. Numbers on the axis refer to ailments as defined in Table A-1

**Table 1: Summary statistics of selected variables**

| | Men | | Women | |
|---|---|---|---|---|
| Variable | Mean | Std. Dev. | Mean | Std. Dev. |
| Depression 42 | 0.115 | 0.319 | 0.160 | 0.367 |
| Depression 33 | 0.030 | 0.169 | 0.090 | 0.286 |
| Depression 23 | 0.035 | 0.185 | 0.083 | 0.277 |
| CSE | 0.166 | 0.372 | 0.141 | 0.348 |
| O-level | 0.377 | 0.485 | 0.447 | 0.497 |
| A-level | 0.083 | 0.275 | 0.096 | 0.295 |
| Higher education | 0.183 | 0.387 | 0.147 | 0.354 |
| dad age when born | 29.908 | 7.590 | 29.558 | 8.178 |
| mom age when born | 27.550 | 5.468 | 27.514 | 5.643 |
| dad age left education | 16.041 | 1.761 | 16.071 | 1.746 |
| mom age left eduation | 16.077 | 1.527 | 16.157 | 1.620 |
| internal score @ 11 | -0.019 | 0.990 | 0.018 | 1.009 |
| external score @ 11 | 0.149 | 1.024 | -0.144 | 0.954 |
| financial difficulties @ 11 | 0.088 | 0.284 | 0.100 | 0.300 |
| character @ 16 | -0.020 | 1.010 | 0.020 | 0.990 |
| Missed school for nervous problem @16 | 0.015 | 0.123 | 0.032 | 0.177 |
| Seen specialist for emotional problem @16 | 0.039 | 0.194 | 0.028 | 0.166 |
| Get well with Mother | 0.732 | 0.443 | 0.734 | 0.442 |
| Get well with Father | 0.673 | 0.469 | 0.637 | 0.481 |
| Live with mother | 0.955 | 0.207 | 0.957 | 0.204 |
| Live with father | 0.895 | 0.306 | 0.873 | 0.333 |
| financial difficulties @ 16 | 0.084 | 0.277 | 0.093 | 0.291 |
| % boy staying on | 60.048 | 24.721 | 55.398 | 20.052 |
| % girl staying on | 55.005 | 20.668 | 60.143 | 24.884 |
| Reading test 7 | -0.082 | 1.011 | 0.172 | 0.881 |
| Math test 7 | 0.059 | 0.978 | -0.028 | 0.963 |
| Type of school: Grammar | 0.098 | 0.297 | 0.117 | 0.322 |
| Type of school: Comprehensive | 0.520 | 0.483 | 0.512 | 0.493 |
| Type of school: Secondary | 0.184 | 0.389 | 0.195 | 0.396 |
| Type of school: Independent | 0.045 | 0.208 | 0.043 | 0.203 |
| Type of school: Other and missing | 0.152 | 0.337 | 0.086 | 0.203 |
| **Instruments** | | | | |
| Child would benefit from further schooling (teacher's view) | 0.372 | 0.483 | 0.418 | 0.493 |
| Smoking at age 16 | 0.289 | 0.453 | 0.287 | 0.452 |
| Numbers of cigarette smoked if>0 | 34.37 | 23.77 | 26.96 | 21.42 |
| Observations | 3271 | | 3395 | |

Note: Observations for depression at age 23 and 33 are respectively for men and women: 2837, 2820 and 3033 and 3073

**Table 2: Effect of mental health on labour market outcomes**

| | Women | | Men | |
|---|---|---|---|---|
| Panel A: depression | Labour force participation | Ln. wage | Labour force participation | Ln. wage |
| At age 23 | -0.100 | -0.222 | -0.001 | 0.211 |
| | (0.039) | (0.186) | (0.030) | (0.287) |
| At age 33 | -0.082 | -0.045 | -0.112 | 0.002 |
| | (0.034) | (0.052) | (0.034) | (0.061) |
| At age 42 | -0.178 | -0.016 | -0.154 | -0.034 |
| | (0.022) | (0.027) | (0.021) | (0.039) |
| Panel B: malaise | | | | |
| At age 23 | -0.009 | -0.037 | -0.004 | 0.021 |
| | (0.004) | (0.018) | (0.002) | (0.021) |
| At age 33 | -0.029 | -0.035 | -0.022 | -0.007 |
| | (0.009) | (0.014) | (0.003) | (0.011) |
| At age 42 | -0.062 | -0.096 | -0.035 | -0.020 |
| | (0.007) | (0.124) | (0.003) | (0.011) |

Note: Robust standard errors are reported into parentheses.
The specifications include highest qualification, test scores in Math and English at age 7, type of school. In the wage regression, we also control for experience and whether working part-time. In the participation equation, the additional controls are: marital status, number and age of the youngest child.
For the labour force participation model, marginal effect from a probit estimated at the mean values of the covariates are reported.

**Table 3: Determinants of depression at age 42- Marginal effects**

| | Probit | | Probit | | IV Probit | |
|---|---|---|---|---|---|---|
| | Female | Male | Female | Male | Female | Male |
| **Education** | | | | | | |
| CSE | -0.028 | -0.007 | | | | |
| | (0.019) | (0.016) | | | | |
| O levels | -0.046 | -0.046 | | | | |
| | (0.017) | (0.014) | | | | |
| A levels | -0.075 | -0.066 | | | | |
| | (0.019) | (0.014) | | | | |
| Higher education | -0.079 | -0.046 | | | | |
| | (0.018) | (0.016) | | | | |
| O-level or above | | | -0.042 | -0.051 | -0.194 | -0.094 |
| | | | (0.016) | (0.014) | (0.062) | (0.062) |
| Internal score @ 11 | 0.024 | 0.005 | 0.023 | 0.005 | 0.027 | 0.006 |
| | (0.006) | (0.005) | (0.006) | (0.005) | (0.007) | (0.005) |
| External score @ 11 | -0.008 | 0.009 | -0.007 | 0.009 | -0.015 | 0.009 |
| | (0.007) | (0.005) | (0.007) | (0.005) | (0.012) | (0.005) |
| Character score @16 | 0.028 | 0.011 | 0.029 | 0.011 | 0.031 | 0.011 |
| | (0.007) | (0.005) | (0.007) | (0.005) | (0.008) | (0.006) |
| Missed school for | 0.003 | -0.015 | 0.005 | -0.013 | -0.000 | -0.016 |
| nervous problems @16 | (0.032) | (0.036) | (0.032) | (0.036) | (0.044) | (0.035) |
| Has ever seen a specialist | 0.052 | 0.131 | 0.054 | 0.131 | 0.043 | 0.130 |
| for emotional prob @16 | (0.041) | (0.040) | (0.041) | (0.040) | (0.049) | (0.040) |
| Get well with dad @ 16 | -0.050 | -0.014 | -0.051 | -0.014 | -0.054 | -0.014 |
| | (0.017) | (0.016) | (0.017) | (0.016) | (0.018) | (0.016) |
| Get well with mum @16 | -0.003 | -0.047 | -0.003 | -0.048 | 0.003 | -0.046 |
| | (0.018) | (0.019) | (0.018) | (0.019) | (0.020) | (0.020) |
| Financial hardship @16 | 0.075 | 0.024 | 0.075 | 0.024 | 0.081 | 0.021 |
| | (0.025) | (0.021) | (0.025) | (0.021) | (0.029) | (0.021) |
| Observations | 3395 | 3271 | 3395 | 3271 | 3395 | 3271 |
| Pseudo $R^2$ | 0.082 | 0.085 | 0.080 | 0.084 | 0.078 | 0.078 |
| $1^{st}$ step regression | | | | | | |
| Teacher's view | | | | | 0.883 | 0.806 |
| | | | | | (0.067) | (0.069) |
| Cigarette smoked @16 | | | | | -0.006 | -0.008 |
| | | | | | (0.002) | (0.001) |
| Joint significance of | | | | | 203.7 | 191.7 |
| instrument : $\chi^2(2)$, p | | | | | p=0.00 | p=0.00 |
| Exogeneity test: | | | | | 7.96 | 0.559 |
| $\chi^2(1)$, p | | | | | P=0.01 | P=0.454 |
| Overidentification test[A]: | | | | | 1.56 | 0.034 |
| $\chi^2(1)$, p | | | | | p=0.21 | p=0.85 |

Note: Robust standard errors are reported into parentheses. Marginal effects estimated at the mean value of the covariates are reported.
The basic regression also includes controls for parental age and education; parental interest in child's schooling at age 11, paternal social class at 11, financial hardship at age 11 and whether living with natural parents at age 16, as well as test score at age seven in math and English, the type of school attended at age 16 and the proportion of pupils at the school attended who stayed on after compulsory education, region of residence at age 16 and at age 42.
Instruments for O levels and above include: teacher's view on whether child would benefit from more schooling and number of cigarettes smoked per week at age 16.
[A] Overidentification test reports the Hansen J, distributed as a $\chi^2(1)$. This statistic is calculated using a linear model estimated by GMM.

**Table 4: Determinants of mental health at age 42 (including additional characteristics)- Marginal effects**

| | Base model | | Model 1: Base + work | | Model 2: Base + family | | Model 3: Base + work + family | |
|---|---|---|---|---|---|---|---|---|
| **Panel A: Depression** | Female | Male | Female | Male | Female | Male | Female | Male |
| O-level and above (probit) | -0.042 (0.016) | -0.051 (0.014) | -0.027 (0.015) | -0.036 (0.013) | -0.038 (0.015) | -0.050 (0.013) | -0.016 (0.014) | -0.035 (0.013) |
| O-level and above (IV) | -0.194 (0.062) | -0.094 (0.062) | -0.172 (0.065) | -0.068 (0.053) | -0.180 (0.063) | -0.075 (0.053) | -0.154 (0.066) | -0.063 (0.055) |
| Work characteristics | | | Yes | Yes | | | Yes | Yes |
| Family Characteristics | | | | | Yes | Yes | Yes | Yes |
| Joint significance of instrument $\chi^2(2)$ | 203.7 | 191.7 | 185.53 | 175.16 | 197.17 | 183.26 | 179.15 | 165.65 |
| Exogeneity test: $\chi^2(1)$ | 7.96 | 0.56 | 5.91 | 0.76 | 6.37 | 0.64 | 4.46 | 0.68 |
| Probability | P=0.01 | P=0.31 | P=0.02 | P=0.38 | P=0.01 | p=0.42 | p=0.03 | p=0.41 |
| Over identification test $\chi^2(1)$ and Probability | 1.563 P=0.21 | 0.034 P=0.85 | 1.460 P=0.23 | 0.063 P=0.80 | 1/079 P=0.30 | 0.006 P=0.94 | 1.008 P=0.32 | 0.110 P=0.74 |
| **Panel B: Normalised Malaise score** | | | | | | | | |
| O-level and above (OLS) | -0.174 (0.045) | -0.151 (0.044) | -0.128 (0.045) | -0.091 (0.042) | -0.159 (0.045) | -0.151 (0.044) | -0.114 (0.045) | -0.087 (0.042) |
| O-level and above (IV) | -0.582 (0.167) | -0.188 (0.167) | -0.509 (0.174) | -0.137 (0.169) | -0.520 (0.168) | -0.199 (0.172) | -0.436 (0.176) | -0.127 (0.175) |
| Work characteristics | | | Yes | Yes | | | Yes | Yes |
| Family Characteristics | | | | | Yes | Yes | Yes | Yes |
| Exogeneity test: $\chi2(1)$ | 6.550 | 0.054 | 5.197 | 0.077 | 5.056 | 0.085 | 3.615 | 0.054 |
| Probability | P=0.01 | P=0.81 | P=0.02 | P=0.78 | P=0.02 | P=0.77 | P=0.06 | P=0.82 |
| Over identification test $\chi^2(1)$ and Probability | 5.337 P=0.02 | 0.003 P=0.95 | 5.48 P=0.02 | 0.351 P=0.55 | 3.715 P=0.05 | 0.029 P=0.86 | 3.820 P=0.05 | 0.444 P=0.50 |

Note: Robust standard errors are reported into parentheses. The specification includes the same covariates as specified in Table 1. For all models we observe 3395 women and 3271 men. Marginal effects estimated at the mean from a probit regression are reported in Panel A. In panel B, the normalised malaise score is the dependent variable and is modelled using linear regressions.

Model 1 adds log pay, and months of work experience as well as dummies for not working, self-employment, full-time work and pay missing to the base model.

Model 2 adds a set of dummies for marital status, number of children, age at first child, age of the youngest child and dummies for no child and death of a child to the base model.

Model 3 includes all the variables of the base model as well as the additional variables of Model 1 and 2.

Instruments for O levels and above include: teacher’s view on whether child would benefit from more schooling and for children attending a comprehensive schools, how this school was created.

**Table 5: Determinants of depression over time (including additional characteristics)-Marginal effects**

| | Model 0 | | Model 1 (Work) | | Model 2 (Family) | | Model 3 (All) | |
|---|---|---|---|---|---|---|---|---|
| | Female | Male | Female | Male | Female | Male | Female | Male |
| **Depression at age 23** | | | | | | | | |
| O-level and above (probit) | -0.035<br>(0.012) | -0.019<br>(0.008) | -0.031<br>(0.012) | -0.019<br>(0.008) | -0.031<br>(0.012) | -0.017<br>(0.008) | -0.029<br>(0.012) | -0.017<br>(0.008) |
| O-level and above (IV) | -0.030<br>(0.052) | -0.153*<br>(0.054) | -0.025<br>(0.059) | -0.181*<br>(0.063) | -0.001<br>(0.058) | -0.144*<br>(0.053) | -0.013<br>(0.063) | -0.174*<br>(0.063) |
| **Depression at age 33** | | | | | | | | |
| O-level and above (probit) | -0.041<br>(0.012) | -0.012<br>(0.007) | -0.036<br>(0.012) | -0.010<br>(0.006) | -0.034<br>(0.012) | -0.010<br>(0.006) | -0.031<br>(0.012) | -0.008<br>(0.006) |
| O-level and above (IV) | -0.074<br>(0.052) | -0.054<br>(0.034) | -0.072<br>(0.053) | -0.062<br>(0.035) | -0.142<br>(0.348) | -0.045<br>(0.035) | -0.035<br>(0.060) | -0.052<br>(0.036) |
| **Depression at age 42** | | | | | | | | |
| O-level and above (probit) | -0.043<br>(0.016) | -0.051<br>(0.014) | -0.027<br>(0.015) | -0.036<br>(0.013) | -0.038<br>(0.015) | -0.050<br>(0.013) | -0.016<br>(0.014) | -0.035<br>(0.013) |
| O-level and above (IV) | -0.184*<br>(0.062) | -0.094<br>(0.052) | -0.172*<br>(0.065) | -0.068<br>(0.053) | -0.180*<br>(0.062) | -0.075<br>(0.053) | -0.154*<br>(0.066) | -0.063<br>(0.055) |

Note: Robust standard errors are reported into parentheses. The specification in Model 0 includes the same covariates as specified in Table 1.
Model 1 adds log pay, and months of work experience as well as dummies for not working, self-employment, full-time work and pay missing.
Model 2 adds a set of dummies for marital status, number of children, age at first child, age of the youngest child and a dummy for no child.
Model 3 includes all the variables of the base model as well as the additional variables of Model 1, 2 and 3.
Instruments for O levels and above include: teacher's view on whether child would benefit from more schooling and for children attending a comprehensive schools, how this school was created.
* indicates a 5% statistical significance in the exogeneity test.

**Table 6: Effect education on depression transitions**

| | Women | | | Men | | |
|---|---|---|---|---|---|---|
| | Probit | Probit | IV | Probit | Probit | IV |
| **Age 23 - 33** | | | | | | |
| CSE | -0.017 | | | 0.007 | | |
| | (0.009) | | | (0.009) | | |
| O levels | -0.016 | | | 0.000 | | |
| | (0.009) | | | (0.006) | | |
| A levels | -0.030 | | | -0.008 | | |
| | (0.008) | | | (0.006) | | |
| Higher education | -0.021 | | | -0.005 | | |
| | (0.009) | | | (0.007) | | |
| O-level or above | | -0.011 | -0.045 | | -0.005 | -0.003 |
| | | (0.009) | (0.045) | | (0.005) | (0.029) |
| Observations | | 2716 | | | 2658 | |
| Pseudo $R^2$ | 0.084 | 0.078 | 0.077 | 0.100 | 0.096 | 0.096 |
| **Age 33 - 43** | | | | | | |
| CSE | -0.032 | | | 0.006 | | |
| | (0.017) | | | (0.016) | | |
| O levels | -0.032 | | | -0.030 | | |
| | (0.017) | | | (0.014) | | |
| A levels | -0.048 | | | -0.036 | | |
| | (0.018) | | | (0.016) | | |
| Higher education | -0.046 | | | -0.023 | | |
| | (0.019) | | | (0.017) | | |
| O-level or above | | -0.018 | -0.076 | | -0.036 | -0.075 |
| | | (0.015) | (0.066) | | (0.013) | (0.053) |
| Observations | | | | 2670 | | |
| Pseudo $R^2$ | 0.055 | 0.052 | 0.052 | 0.068 | 0.067 | 0.061 |

Note: keep only individuals who reported not being depressed in the first period

**Table 7: Effect of education on malaise score – zero inflated negative binomial**

| | Female | | | Male | | |
|---|---|---|---|---|---|---|
| | zinb | zinb | IV | zinb | zinb | IV |
| **Age 42** | | | | | | |
| CSE | -0.062 | | | -0.161 | | |
| | (0.054) | | | (0.065) | | |
| O-levels | -0.157 | | | -0.222 | | |
| | (0.047) | | | (0.059) | | |
| A-levels | -0.307 | | | -0.326 | | |
| | (0.070) | | | (0.086) | | |
| Higher ed. | -0.292 | | | -0.337 | | |
| | (0.063) | | | (0.076) | | |
| O-levels and | | -0.151 | -0.547 | | -0.168 | -0.252 |
| above | | (0.037) | (0.151) | | (0.047) | (0.185) |
| Residuals | | | 0.423 | | | 0.089 |
| | | | (0.157) | | | (0.190) |
| **Age 33** | | | | | | |
| CSE | -0.078 | | | -0.146 | | |
| | (0.073) | | | (0.082) | | |
| O-levels | -0.287 | | | -0.291 | | |
| | (0.063) | | | (0.077) | | |
| A-levels | -0.363 | | | -0.315 | | |
| | (0.093) | | | (0.109) | | |
| Higher ed. | -0.465 | | | -0.471 | | |
| | (0.089) | | | (0.103) | | |
| O-levels and | | -0.265 | -0.748 | | -0.236 | -0.804 |
| above | | (0.049) | (0.198) | | (0.059) | (0.242) |
| Residuals | | | 0.515 | | | 0.604 |
| | | | (0.205) | | | (0.249) |
| **Age 23** | | | | | | |
| CSE | 0.016 | | | -0.211 | | |
| | (0.056) | | | (0.073) | | |
| O-levels | -0.197 | | | -0.364 | | |
| | (0.049) | | | (0.066) | | |
| A-levels | -0.302 | | | -0.348 | | |
| | (0.079) | | | (0.102) | | |
| Higher ed. | -0.343 | | | -0.459 | | |
| | (0.071) | | | (0.092) | | |
| O-levels and | | -0.226 | -0.693 | | -0.270 | -0.901 |
| above | | (0.039) | (0.162) | | (0.053) | (0.219) |
| Residuals | | | 0.500 | | | 0.676 |
| | | | (0.167) | | | (0.225) |

Note: see note under Table 3 for specification. Standard errors in the IV models are obtained by boostrapping (500 replications)

**Table 8: Effect of education on malaise score – Quantile regression**

| | Female | | | Male | | |
|---|---|---|---|---|---|---|
| | 1st Quartile | Median | 3rd Quartile | 1st Quartile | Median | 3rd Quartile |
| **Age 42** | | | | | | |
| CSE | 0.191 | -0.227 | -0.631 | -0.082 | -0.369 | -0.620 |
| | (0.325) | (0.274) | (0.341) | (0.077) | (0.160) | (0.289) |
| O-levels | -0.259 | -0.703 | -1.117 | 0.061 | -0.396 | -1.110 |
| | (0.232) | (0.234) | (0.250) | (0.100) | (0.200) | (0.245) |
| A-levels | -0.563 | -1.199 | -2.005 | -0.072 | -0.555 | -1.308 |
| | (0.362) | (0.374) | (0.438) | (0.156) | (0.340) | (0.427) |
| Higher ed. | -0.460 | -1.137 | -1.633 | 0.022 | -0.796 | -1.516 |
| | (0.237) | (0.236) | (0.366) | (0.193) | (0.348) | (0.313) |
| **O level and above** | -0.371 | -0.625 | -0.885 | 0.055 | -0.296 | -0.744 |
| | (0.153) | (0.234) | (0.298) | (0.111) | (0.156) | (0.335) |
| **IV Quantile** | -1.105 | -0.805 | -2.810 | -0.450 | -1.630 | -0.940 |
| | (0.624) | (0.715) | (1.193) | (0.440) | (0.793) | (0.909) |

Note: see note under Table 3 for specification. All three quantile regressions are estimated simultaneously when education is assumed to be exogenous.
The endogenous models are estimated using the Chernozhukov and Hansen (2005) algorithm.

**Table 9: Effect education on alternative measures of mental health**

| | Women | | | Men | | |
|---|---|---|---|---|---|---|
| Normalised non somatic malaise score | OLS | OLS | IV | OLS | OLS | IV |
| CSE | -0.072 | | | -0.124 | | |
| | (0.070) | | | (0.067) | | |
| O levels | -0.193 | | | -0.199 | | |
| | (0.061) | | | (0.059) | | |
| A levels | -0.341 | | | -0.246 | | |
| | (0.076) | | | (0.077) | | |
| Higher education | -0.285 | | | -0.244 | | |
| | (0.072) | | | (0.070) | | |
| O-level or above | | -0.179 | -0.568 | | -0.147 | -0.054 |
| | | (0.046) | (0.164) | | (0.044) | (0.168) |
| Exogeneity test: $\chi 2(1)$ | | | 5.75 | | | 0.325 |
| Probability | | | P=0.017 | | | P=0.568 |
| Over identification test | | | 3.30 | | | 0.606 |
| $\chi^2(1)$ and Probability | | | P=0.069 | | | P=0.436 |

Note: See note under Table 3 for specification details.

**References:**


Beck A. (1987), "Cognitive models of depression", *Journal of Cognitive Psychotherapy*, 1, 5-37

Berger M. and P. Leigh (1989), "Schooling, self-selection and health", *Journal of Human Resources*, 24, 433-455

Bound J., D. Jaeger and R. Baker (1995), "Problems with instrumental variables estimation when the correlation between the instruments and the endogeneous explanatory variable is weak", *Journal of the American Statistical Association*, 90, 443-450

Chernozhukov V. and C. Hansen (2005), "Instrumental variable quantile regression", Econometrica

Currie J. and E. Moretti (2003), "Mother's education and the intergenerational transmission of human capital. Evidence from college openings and longitudinal data", *Quarterly Journal of Economics*, 118, 1495-1532

Currie J. and M. Stabile (2004), "Child mental health and human capital accumulation: the case of ADHD", *National Bureau of Economic Research*, WP 10435

Dickey W. and S. Blumberg (2002), "Prevalence of Mental Disorders and Contacts with Mental Health Professionals Among Adults in the United States, National Health Interview Survey, 1999" in Mental Health, United States, 2002, R. Manderscheid and M. Henderson, Department of Health and Human Services

Duncan B. and D. Rees (2004), "Effect of smoking on depressive symptomatology: A re-examination of data from the National Longitudinal Study of Adolescent Health", *American Journal of Epidemiology*, 162, 461-470

Evans W. and E. Montgomery (1994), "Education and health: Where there's smoke there's an instrument", *National Bureau of Economic Research*, WP 4949

Feinstein, L. and Bynner, J. (2005), "Continuity and discontinuity in mid-childhood: implications for adult outcomes in the UK 1970 Birth Cohort," in Huston, A. and Ripke, M. (Eds) *Middle Childhood: Contexts of Development*, Harvard University Press, Cambridge

Fuchs V. (1982), "Time preference and health: An exploratory study" in *Economic Aspects of Health*, V. Fuchs (Ed), University of Chicago Press

Goldman D. and D. Lakdawalla (2001), "Understanding health disparities across education groups", *National Bureau of Economic Research*, WP 8326

Goldman D. and J. Smith (2002), "Can patient self-management help explain the SES health gradient?", *Proceedings, National Academy of Sciences*, 99, 10929-34

Grossman M. (2000), "The human capital model", in the *Handbook of health economics*, A. Cuyler and P. Newhouse (eds), North Holland, Amsterdam

Grossman M. (2005), "Education and Non Market Outcomes", in the *Handbook of the Economics of Education*, E. Hanushek and Finish Welch (eds), North Holland, Amsterdam.

Heckman J., H. Ichimura, P. Todd (1997), "Matching as an econometric evaluation estimator: Evidence from evaluating a job training program", *Review of Economic Studies*, 64,605-654

Heckman J., J. Stixrud and S. Urzua (2006), "The effects of cognitive and non-cognitive skills on labor and behavioural outcomes", *Journal of Labor Economics*, 24, 411-482

Jalovaara M. (2004), "Socioeconomic differentials in divorse risk by duration of marriage", *Demographic Research*, 7, 538-562

Jorm A. (2000), "Does old age reduce the risk of anxiety and depression? A review of epidemiological studies across the adult life span", *Psychological Medicine*, 30, 11-22

Kendler, K., Kessler, R., Neale,M., Heath, A. & Eaves, L., (1993), "The prediction of major depression in women: toward and integrated etiologic model", *American Journal of Psychiatry*, 150, pp1139-1148

Kenkel D. (1991), "Health behaviour, health knowledge and schooling", *Journal of Political Economy*, 99, 287-305

Kenkel D., D. Lillard and A. Mathios (2006), "The role of high school completion and GED receipt in smoking and obesity", *Journal of Labor Economics*, 24, 635-660

Koenker R. and G. Bassett (1978), "Regression quantiles", *Econometrica*, 46, 33-50

Kubzansky, L.D.; Kawachi, I.; Sparrow, D. (1999) 'SocioEconomic status, hostility, and risk factor clustering in the normative aging study: Any help from the concept of allostatic load?', *The Society of Behavioural Medicine*, 21, 4, 330-338

Llena-Nozal A., M. Lindeboom and F. Portrait (2004), "The effect of work on mental health: does occupation matter?" *Health Economics*, 13, 1045-62

Lleras-Muney A. (2005), "The relationship between education and adult mortality in the US", *Review of Economic Studies*, 72, 189-221

Lewinsohn, P., Hoberman, H. & Rosenbaum, M. (1988), "A prospective study of risk factors for unipolar depression", *Journal of Abnormal Psychiatry*. 97, pp251-264

Lindelow M., R. Hardy and B. Rodger (1997), "Development of a scale to measure symptoms of anxiety and depression in the general UK population: The psychiatric symptom frequency scale", *Journal of Epidemiology Community Health*, 51, 549-557

Marmot M., G. Smith, S. Stansfeld, C. Patel, F. North, J. Head, I. White, E. Brunner and A. Feeney (1991), "Health inequalities among British civil servants: the Whitehall II study", *The Lancet*, 337: 1387–1393

Mazure C., M. Bruce, P. Maciejewski and S. Jacobs (2000), "Adverse life events and cognitive-personality characteristics in the prediction of major depression and antidepressant response" *American Journal of Psychiatry*. 157, pp 896-903

Meltzer, H., B. Gill, M. Petticrew and K. Hinds (1995), "The prevalence of psychiatric morbidity among adults living in private households", *Office of Population Censuses & Surveys*: HMSO

Munasinghe L. and N. Sicherman (2000), "Why do dancers smoke: Time preference, occupational choice and wage growth", *National Bureau of Economic Research*, WP 7542

Propper C. , K. Jones, A. Bolster, S. Burgess, R. Johnston and R. Sarker (2004), "Local neighbourhood and mental health: evidence from the UK", *University of Bristol*, CMPO, WP 04/099

Rodgers B., A. Pickles, C. Power, S. Collishaw and B. Maughan, (1999), "Validity of the malaise inventory in general population samples", *Soc. Psychiatry Psychiatr. Edidemiol*, 34, 333-41

Rutter M. (1967), "A Children's Behaviour Questionnaire for Completion by Teachers: Preliminary Findings." *Journal of Child Psychology and Psychiatry*, Vol **8**, pp1-11.

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

Silberg J., M. Rutter, M. Neale and L. Eaves (2001), "Genetic moderation of environmental risk for depression and anxiety in adolescent girls" *British Journal of Psychiatry*, 179, pp 116-121

Singleton N., R. Bumpstead, M. O'Brien, A. Lee and H. Meltzer (2001), "Psychiatric morbidity among adults living in privage households, 2000", HMSO

Stansfeld S. and M. Marmot (1992), "Social class and minor psychiatric disorder in British civil servants: A validated screening survey using the General Health Questionnaire", *Psychol Med*, 22, 739-49

Thomas C. and S. Morris (2003), "Cost of depression among adults in England in 2000", *British Journal of Psychiatry*, 183, 514-519

Westergaard-Nielsen, N., E. Agerbo, T. Eriksson and P.B. Mortensen (2005), "Mental illness: Gender differences with respect to marital status and labour market outcomes", in Marcotte D. and Wilcox-Gok V. (eds) "*The economics of gender and mental health*"

WHO (1999), World health Report, Making a difference, WHO, Geneva

Wilson C. and A. Oswald (2005), "How does marriage affect physical and psychological health? A survey of the longitudinal evidence", *IZA*, DP 1619

Wooldbridge J. (2002), "Econometric Analysis of Cross section and panel data", MIT Press, Cambridge

Zubenko G., B. Maher, H. Hughes, W. Zubenko, J. Stiffler, B. Kaplan and M. Marazita, (2003) "Genome-wide linkage survey for genetic loci that influence the development of depressive disorders in families with recurrent, early-onset, major depression". *American Journal of Medical Genetics*

**Annex 1:**

**Table A1: Malaise score – Proportion of individuals with a given ailment over time**

| | Malaise inventory | Age 23 | Age 33 | Age 42 |
|---|---|---|---|---|
| 1 | I often have back-ache * | 0.190 | 0.243 | 0.302 |
| 2 | I feel tired most of the time | 0.173 | 0.190 | 0.339 |
| 3 | I often feel miserable or depressed | 0.144 | 0.115 | 0.202 |
| 4 | I often have bad head-aches * | 0.129 | 0.146 | 0.164 |
| 5 | I often get worried about things | 0.423 | 0.327 | 0.466 |
| 6 | I usually have great difficulty sleeping | 0.105 | 0.121 | 0.219 |
| 7 | I usually wake unnecessarily early | 0.158 | 0.167 | 0.318 |
| 8 | I wear myself out worrying about my health | 0.025 | 0.030 | 0.071 |
| 9 | I often get into a violent rage | 0.058 | 0.046 | 0.048 |
| 10 | People often annoy and irritate me | 0.267 | 0.218 | 0.305 |
| 11 | At time I have twitching of face/shoulders * | 0.079 | 0.072 | 0.092 |
| 12 | I often suddenly become scared for no good reason | 0.084 | 0.049 | 0.073 |
| 13 | I'm scared to be alone | 0.095 | 0.032 | 0.040 |
| 14 | I'm easily upset or irritated | 0.232 | 0.163 | 0.205 |
| 15 | I'm frightened of going out alone | 0.077 | 0.048 | 0.071 |
| 16 | I'm constantly keyed up and jittery | 0.038 | 0.040 | 0.058 |
| 17 | I suffer from indigestion * | 0.113 | 0.130 | 0.178 |
| 18 | I suffer from an upset stomach * | 0.101 | 0.097 | 0.126 |
| 19 | I have poor appetite | 0.044 | 0.038 | 0.052 |
| 20 | Every little things gets on my nerves | 0.024 | 0.027 | 0.046 |
| 21 | My heart often race like mad * | 0.071 | 0.057 | 0.083 |
| 22 | I often have bad pains in my eyes * | 0.050 | 0.040 | 0.030 |
| 23 | I have trouble with rheumatism or fibrosis * | 0.039 | 0.041 | 0.053 |
| 24 | I had a nervous breakdown | 0.016 | 0.026 | 0.050 |
| | | | | |
| | Malaise score | 2.53 | 2.31 | 3.45 |
| | Depression indicator | 0.061 | 0.061 | 0.138 |

Note: * denotes somatic ailment, as defined in Rodgers et al. (1999).

## Annex 2: Malaise score and self-reported measure of mental health

### Table A2-1: Mean malaise score by self-reported mental health

| Are you currently Feeling low, depressed, sad | Rutter Scale Frequency | Malaise score | Depressed |
|---|---|---|---|
| Yes, most days | 0.029 | 10.989 * | 0.751 * |
| Yes, occasionally | 0.086 | 7.168 * | 0.416 * |
| In the past, but not currently | 0.098 | 4.072 * | 0.168 * |
| Never have | 0.787 | 2.811 | 0.081 |

Note: Individuals with a malaise score strictly greater than 7 are defined as depressed. * denotes statistical difference in the means between the category defined on the row and individuals who have never been depressed.

### Table A2-2: Probability of feeling currently depressed by malaise score items

| Malaise inventory age 42 | Currently depressed Malaise=0 | Malaise=1 |
|---|---|---|
| I often have back-ache * | 0.098 | 0.165 |
| I feel tired most of the time | 0.065 | 0.221 |
| I often feel miserable or depressed | 0.042 | 0.419 |
| I often have bad head-aches * | 0.094 | 0.241 |
| I often get worried about things | 0.039 | 0.209 |
| I usually have great difficulty sleeping | 0.074 | 0.274 |
| I usually wake unnecessarily early | 0.072 | 0.217 |
| I wear myself out worrying about my health | 0.093 | 0.441 |
| I often get into a violent rage | 0.107 | 0.335 |
| People often annoy and irritate me | 0.084 | 0.196 |
| At time I have twitching of face/shoulders * | 0.105 | 0.245 |
| I often suddenly become scared for no good reason | 0.094 | 0.429 |
| I'm scared to be alone | 0.104 | 0.450 |
| I'm easily upset or irritated | 0.077 | 0.278 |
| I'm frightened of going out alone | 0.099 | 0.370 |
| I'm constantly keyed up and jittery | 0.098 | 0.444 |
| I suffer from indigestion * | 0.103 | 0.190 |
| I suffer from an upset stomach * | 0.101 | 0.239 |
| I have poor appetite | 0.103 | 0.392 |
| Every little things gets on my nerves | 0.100 | 0.491 |
| My heart often race like mad * | 0.094 | 0.387 |
| I often have bad pains in my eyes * | 0.111 | 0.346 |
| I have trouble with rheumatism or fibrosis * | 0.109 | 0.277 |
| I had a nervous breakdown | 0.099 | 0.488 |

Note: Currently depressed equals 1 if feeling low, sad or depressed most days or occasionally, zero otherwise. All differences are statistically significant at the 5% level. * denotes somatic ailment, as defined in Rodgers et al. (1999).