BERJ | British Educational Research Journal

BEI

I quit! Is there an association between leaving teaching and improvements in mental health?

Journal:	*British Educational Research Journal*
Manuscript ID	CBER-2019-0455.R1
Manuscript Type:	Original Paper
Keywords:	Biobank, well-being, mental health, Teachers < Practitioners Teachers and Trainers

Abstract:	The mental health and well-being of teachers is an issue of great policy concern. This is particularly true in England, where high workload and the associated stress is thought to be leading to a recruitment and retention crisis within the education profession. But do individuals who decide to leave teaching for another career actually see their well-being and mental health improve? We provide new evidence on this matter for individuals aged between 40 and 65, using the rich information gathered as part of the UK Biobank study. Our analysis shows that individuals who choose to leave teaching are somewhat happier in their work, but do not generally experience any improvement in their general well-being or mental health. We hence caution those middle-aged teachers who are thinking of leaving teaching that the grass may not necessarily be greener on the other side.

SCHOLARONE™
Manuscripts

The mental health and well-being of teachers is an issue of great policy concern. This is particularly true in England, where high workload and the associated stress is thought to be leading to a recruitment and retention crisis within the education profession. But do individuals who decide to leave teaching for another career actually see their well-being and mental health improve? We provide new evidence on this matter for individuals aged between 40 and 65, using the rich information gathered as part of the UK Biobank study. Our analysis shows that individuals who choose to leave teaching are somewhat happier in their work, but do not generally experience any improvement in their general well-being or mental health. We hence caution those middle-aged teachers who are thinking of leaving teaching that the grass may not necessarily be greener on the other side.

1. Introduction

Almost thirty years ago it was observed that "…teaching is an unsettled and unhappy profession at present and there is a problem to be addressed" (Smithers, 1990). A similar statement would not be out of place when describing the teaching profession today. As a job, teaching requires staff to work long hours, to keep up with changing government requirements and to manage often disruptive classrooms (*author cite*). This, of course, all occurs under the watchful eye of the accountability system, with schools (and teachers) judged by how young people perform in high-stakes national examinations.

Many teachers enter the profession for altruistic reasons, yet key issues which motivate many to leave are workload, pupil behaviour and salary (Dolton & Klaauw, 1995; Barmby, 2006). Many teachers experience dissatisfaction due to feeling a lack of control with respect to their working conditions, accompanied by the absence of a platform from which to voice their concerns (Mercer & Evans, 1991). Researchers have observed similar trends for teachers outside of England, including the USA, Australia and the Netherlands, especially in newer teachers (Tye & O'Brien, 2002; Howes and Goodman-Delahunty, 2015; den Brok et al., 2017; *author cite*). However, strikingly, many more teachers in England expressed regret about their career choice than in almost any other industrialised country (*author cite*).

Today's retention rate in the UK is low especially amongst newly qualified teachers (NQTs), and schools located in more socially deprived backgrounds experience a higher turnover of teachers (*author cite*). There is also some suggestive evidence that it is the more able

For Peer Review Only

teachers who are the most likely to leave teaching in search of alternative employment, again impacting on teacher quality, as well as there being strong monetary implications, given the cost of training teachers (Culver et al., 1990; Borman & Dowling, 2008).

Long-serving teachers in urban UK schools appear to remain in their jobs due to the deep connections and emotional ties they forge within their workplaces and communities (McIntyre, 2010). By implication then, teachers who leave the profession may do so either because they cannot forge these connections which provide an adequate level of job satisfaction, or because the challenges of workload, classroom management and salary outweigh the benefits they receive.

Such pressures may lead to work-related stress and, in turn, be detrimental to teachers' mental and physical health (Travers & Cooper, 1993; Berryhill, Linney, & Fromewick 2009; Scheuch & Seibt, 2015; Kidger et al., 2016; Merrida-Lopez & Extremara, 2017; Education Support Partnership, 2018). A number of studies have also suggested that teachers have worse mental health and wellbeing than those who work in other jobs. For instance, Johnson et al. (2003) found teachers to have one of the lowest levels of psychological wellbeing out of the 26 occupational groups they considered. Similarly, Stansfeld et al. (2011) found that teachers were at above average risk of suffering from mental ill-health, with Kidger et al. (2016) indicating that wellbeing was low and depressive symptoms high amongst teachers. Indeed, the stresses and strains of teaching are widely cited for the ongoing teacher retention and recruitment crisis that continues to create a major challenge for England's schools (e.g. Tapper, 2018). It is therefore little wonder that, in a recent nationally representative survey of teachers, around half of those working in secondary schools in England said that they wondered whether they would have been better off working in a different job (*author cite*).

This leads to an important question for teachers who are contemplating a change of career: do those who choose to leave teaching have higher levels of well-being, and a lower likelihood of developing mental health problems, than those who choose to remain? In general, occupation is known to be linked to health and wellbeing (Clark, 2010; Johnson et al., 2005; Ravesteijn, Kippersluis, & Doorslaer, 2013; Ravesteijn, Kippersluis, & Doorslaer, 2018). This is perhaps unsurprising, given the influence that work has on our social interaction with others, our ability to develop and employ new skills, our sense of personal achievement and the

For Peer Review Only

restrictions it imposes on other areas of life e.g. time with family (Creek & Hughes, 2008; Gallagher, Muldoon, & Pettigrew, 2015). Consequently, individuals who are unhappy may switch jobs in an attempt to improve their wellbeing. Indeed, empirical research suggests that voluntary job changes are associated with increased job satisfaction (Chadi & Hetschko, 2018; Gielen, 2013) and improved mental health (Longhi et al., 2019), at least in the short run.

The evidence in relation to teachers is, however, more limited. The study most comparable to ours is the interesting work of Bamford and Worth (2017). Using longitudinal data, these authors tracked job satisfaction, overall life satisfaction and the subjective well-being of a small sample of teachers within the UK. They found that teachers who left the profession experience a large increase in job satisfaction, and a small increase in subjective well-being, compared with those who decided to stay. Yet, as with all studies, the authors noted some key limitations. First, the sample size of teachers was small (e.g. 231 former teachers had left their job within the last year, 107 who had left teaching between one and two years ago and 74 who had left three years previously), with estimates surrounded by quite wide confidence intervals. Second, the focus was upon teachers who quit their job recently (mainly within the last two years) meaning it was not possible to consider the association between leaving teaching and longer-term outcomes. Finally, the Understanding Society dataset analysed collected only limited information about respondents' well-being and mental health. The combination of these factors made it challenging to compare the outcomes of individuals who chose to join, leave and stay in the teaching profession.

Our paper seeks to contribute to this understudied area, with a particular focus upon the well-being and mental health outcomes of current and former teachers aged between 40 and 65 within the UK. It is, to our knowledge, the first study to utilise the UK Biobank dataset to investigate this issue. Through this data, we have access to a wide array of information about respondents' well-being and mental health. This not only includes responses to standardised questionnaires (as have previously been used in this literature) but also prescription of common medicines used for conditions such as anxiety, depression and insomnia. These data can therefore be used to explore the mental health outcomes of current and former teachers across a wide range of important measures. Moreover, having such a rich array of data collected at baseline means that we can more credibly control for potential confounding factors within the longitudinal component of our analysis. Together, this enables us to provide

For Peer Review Only

important new evidence as to whether joining/leaving teaching has a sizeable impact upon a person's mental health, at a time when many teachers are thinking about changing career in search of a better life. It is important to recognise, however, that the data we use are based upon a sample of 40 to 65-year-olds and may not generalise to younger age groups.

To trail our key results, we find little evidence that the grass really is greener for those who quit the teaching profession. Although there is some evidence of an increase in job satisfaction for those who left teaching relatively recently (consistent with the findings of Bamford & Worth, 2017), there is little to suggest that leaving teaching reduces the risk of suffering from mental health problems, such as depression and anxiety. Similarly, there is little evidence to suggest that individuals who choose to quit teaching are happier in their life in general (i.e. outside of work). This leads us to conclude that teaching as an occupational choice is unlikely to be a major cause of poor mental health outcomes per se.

The paper now proceeds as follows. The UK Biobank dataset is described in section 2, with an overview of our empirical methodology following in section 3. Our results are detailed in section 4, with conclusions and policy discussion provided in section 5.

2. UK Biobank data

The initial UK Biobank data collection took place between 2006 and 2010. A total of around half a million volunteers between the ages of 40 and 69 participated in the study. These data therefore form a convenience sample rather than being a random sample that is representative of the wider population. Participants attended an initial assessment centre when they were first recruited, where they completed questionnaires, were interviewed by a trained health professional (in order to collect accurate information about medical conditions and currently prescribed drugs) and underwent some basic health checks (e.g. participants' blood pressure was taken; an electrocardiogram/ECG was conducted). It contains uniquely rich information about health from a very large number of individuals – many of whom were employed as teachers. A number of follow-up questionnaires have been gathered from Biobank participants since the initial assessment centre. In 2016, 117,500 participants completed an online 'occupational career' questionnaire. Respondents were first asked to type into an open text field their job title, start date and end date for each job they held. The respondent was then asked to work through a set of 'drop-down' job lists, with a list of

URL: http://mc.manuscriptcentral.com/cber

For Peer Review Only

possible job titles presented on the final screen (from which they were then asked to pick the most suitable). The relevant SOC code was then assigned to the individual, based upon their selection from the drop-down menus (information entered as free-text at the start was used for validation processes only). The data went through a validation process by an expert occupational coder, with reasonably good agreement found (Cohen's Kappa = 0.45). Further details about how the occupational career data has been captured is available within de Matteis et al (2017).

As part of the occupational career questionnaire, some further basic information was asked about each job, such as typical number of hours worked per week, whether it involved shift work and exposure to potentially hazardous substances (e.g. asbestos). For our purposes, this information allows us to identify the year respondents entered and exited the teaching profession, including the occupation that they joined when they left. Critically, this means we can identify both current and former teachers, facilitating comparisons between these two groups.

For each job recorded at the assessment centre or in the occupational history questionnaire, four-digit Standard Occupational Classification (SOC2000) codes are provided within the Biobank database, along with some further information about specific job role. Throughout this paper, we begin by identifying teachers using the following broad set of SOC codes:

- 2312 = Further education teaching professionals
- 2314 = Secondary education teaching professionals
- 2315 = Primary and nursery education teaching professionals
- 2316 = Special needs education teaching professionals

We excluded individuals from analysis if they were identified as working in a non-teaching role. For instance, for respondents with a SOC code of 2315, extra information was used to exclude nursery workers from our definition of 'teachers'[1]. Likewise, this extra information was used to remove further education lecturers, whilst retaining further education teachers.

[1] The Biobank data provides additional information about the job of each respondent, over and above the SOC code. For instance, for SOC code 2315 it provides information on whether the individual is a headteacher or not,

For Peer Review Only

Key measures collected during the 2006-2010 assessment centre

Within our analysis, we make particular use of the following measures collected during the assessment centre:

- Neuroticism score. Neuroticism was measured with the 12-item neuroticism subscale from the short form of the revised Eysenck Personality Questionnaire (EPQ-N). This encompassed 12 questions such as 'do you often feel fed-up', 'do you suffer from nerves', and 'do you often feel lonely', with participants asked to respond either yes, no, don't know or prefer not to say to each item. A total neuroticism score from these 12 items is provided as part of the Biobank dataset, which has been reported to have good levels of internal validity (Cronbach's alpha of 0.84 – Peters et al 2018). Within our analysis, we standardise this scale to mean zero and standard deviation one.
- Depression. Within the self-completion questionnaire, respondents were asked four questions about how they felt over the last two weeks. This included (a) frequency of depressed mood; (b) tiredness / lethargy; (c) unenthusiastic/ disinterest and (d) tenseness/restlessness. Respondents were asked to indicate one of following response options for each (1. Not at all; 2. Several days; 3. More than half the days; 4. Nearly every day; 5. Don't know; 6. Prefer not to say). Following McCormack et al (2015), we combine responses to these questions into an overall depression scale. This is done using a two-parameter item-response theory (IRT) model, which is then standardised to mean zero and standard deviation one.
- Sleep. A wide body of research has shown that mental health problems, such as anxiety and depression, are linked to insomnia and a lack of sleep (Freeman et al., 2017; Fujishiro et al., 2017). As part of the self-completion questionnaire, respondents were asked (a) the number of hours of sleep they typically get within a 24-hour period and (b) whether they either have trouble falling asleep or whether they wake up during the night (1. Never/rarely; 2. Sometimes; 3. Usually).
- Alcohol intake. Previous research has found that up to 40 percent of adults use alcohol as a mechanism to cope with stress (Appleton and James 2018). Within the

and whether they work in a primary or nursery setting. The variable in question is available from here: https://biobank.ctsu.ox.ac.uk/crystal/field.cgi?id=132.

For Peer Review Only

assessment centre, respondents were first asked how frequently they drink alcohol (daily/almost daily; three/four times a week; once/twice a week; one to three times a month; special occasions only; never). For those who said that they drink alcohol more than once or twice per week, they were then asked about average weekly intake of (a) pints of beer/cider; (b) white wine/Champagne; (c) fortified wine; (d) red wine; (e) spirits. For those who said that they drunk alcohol monthly, or only on special occasions, monthly figures were provided. Following Taylor et al (2018) we convert these responses into an approximate number of units of alcohol consumed per week. Both frequency of drinking per week and weekly units of alcohol consumed are considered within our analysis.

- Medications prescribed. As part of an interview with a trained nurse, respondents were asked about prescribed medications. Using this information, we create a binary variable, coded as one if they take frequently prescribed medications for anxiety, depression or insomnia[2], and zero otherwise.
- Medical conditions. In the assessment centre, participants were asked to indicate any medical conditions that they had. If the participant was uncertain of the type of illness they had, then they were asked to describe it to a trained nurse who placed it within a category. They were also asked the date or age when they were first diagnosed with the condition. Our focus is upon reports of depression, anxiety, self-harm, stress and insomnia, coded as one if they reported having one of these conditions, and zero otherwise.
- Happiness. Respondents were asked 'in general, how happy are you' with responses provided using a six-point scale (extremely happy to extremely unhappy)[3].
- Happiness with different aspects of life. Respondents were asked the same question as above, but with the focus being upon a certain aspect of their life. This included how happy they were with their (a) work; (b) family; (c) finances; (d) friends; (e) health.

[2] If respondents indicated that they were currently prescribed one of the following medications, this dummy variable was coded as one: Citalopram, Escitalopram, Fluoxetine, Fluvoxamine, Amitripyline, Paroxetine, Sertraline, Venlafaxine, Duloxteine, Pregabalin, Cymbalta, Yentreve, Mirtazapine, Anafranil, Prozac, Diazepam, Zopiclone, Temazepam, Nitrazepam.

[3] The questions about happiness and work/job satisfaction were only introduced into the assessment centre questionnaire in 2009. These data are therefore only available for a subset of respondents.

For Peer Review Only

Responses were again provided using a six-point scale (extremely happy to extremely unhappy).

A wide range of other information was also collected within the assessment centre, including demographic background, basic health assessments (blood pressure measurements), illnesses of mother/father/siblings (including depression) and whether selected life events had occurred within the last two years (e.g. death of a family member, divorce, a serious illness). Although not the focus of this paper, a selection of such variables will be used as controls within our analysis.

Follow-up assessment centres

A sub-set of Biobank participants have completed return visits to the assessment centre, providing a longitudinal element to the dataset. These longitudinal follow-ups have been concentrated within certain parts of the country; most notably the assessment centres located at Cheadle (Stockport), Reading and Newcastle. Most of the same data were collected as in the initial assessment centre, including current occupation, currently prescribed medications, current medical conditions, sleep, depression, alcohol consumption and happiness with different aspects of life. Critically, this means we can identify individuals who have changed job between the two assessment centres (e.g. individuals who joined or left the teaching profession) and measure change in the aforementioned measures of mental health and well-being.

The 2016 mental health questionnaire

In 2016, a subset of 137,000 biobank participants completed an additional questionnaire about their mental health. Most of these questionnaires were completed between the 19th of August and the 26th of September 2016. Importantly, this provides a second longitudinal component to the Biobank data, with the 2006-2010 assessment centre acting as the baseline and the 2016 mental health questionnaire as the follow-up. Within our analysis, we make particular use of the following information gathered within the mental health questionnaire:

- Current depression. Respondents were asked "*over the last 2 weeks, how often have you been bothered by any of the following problems*", with nine separate questions then following (e.g. Little interest or pleasure in doing things; feeling down, depressed,

For Peer Review Only

or hopeless; feeling tired or having little energy). They were asked to select one of four responses for each (not at all; several days; more than half the days; nearly every day). A two-parameter IRT model is estimated using these nine items in order to construct a depression scale. We standardise this scale to mean zero and standard deviation one.

- Current anxiety. Respondents answered the same question as presented above for current depression, with seven separate items (e.g. feeling nervous, anxious or on edge; worrying too much about different things; trouble relaxing). A two-parameter IRT model is again used to derive an overall anxiety scale.
- Prolonged periods of depression. First, respondents were asked the following two yes/no questions: "*have you ever had a time in your life when you felt sad, blue, or depressed for two weeks or more in a row*" and "*have you ever had a time in your life when you lost interest in most things like hobbies, work, or activities that usually give you please*". To respondents who answered yes to either of these questions, they were asked to recall the age that they first had such a spell and the age they were when they last had such a spell. For the purposes of this paper, this information can be compared to spells when respondents were and were not working as teachers.
- Alcohol consumption. Respondents were asked: "*In the next two questions, a "drink" is defined as one unit of alcohol. How often do you have six or more drinks on one occasion*?" We use responses to this question to get an indication of the extent that respondents engage in heavy drinking.
- Self-harm. A series of questions were asked about self-harm, including "*many people have thoughts that life is not worth living. Have you felt that way*" and "*have you contemplated harming yourself*?". They were also asked if they had felt this way in the last 12 months or if they had harmed themselves during the last 12 months.
- Happiness. Respondents were asked the same question as in the 2006-2010 assessment centre.
- Happiness with health. Respondents were asked "*In general how happy are you with your HEALTH*?", using the same six response options presented for the happiness scale.

For Peer Review Only

- Feel life is meaningful. Participants were asked "*to what extent do you feel your life to be meaningful*?", with responses provided on a five-point scale (not at all; a little; a moderate amount; very much; an extreme amount).

These are the main outcome measures we consider when analysing the 2016 mental health data.

3. Methodology

Outcomes measured at the initial assessment centre

To begin, we focus upon the 2006-2010 assessment centre measures described above as our outcomes of interest. These outcomes will be compared across the following groups:

1. Current teachers (reference group). Individuals who were teachers when the initial assessment centre took place (n = 16,622).
2. Former teachers who left within the last five years. Individuals who were teachers, but left the teaching profession for another career within the five years prior to the Biobank assessment centre (n = 1,271)[4].
3. Former teachers who left six to ten years ago. Individuals who were teachers, but left the teaching profession for another career between six and ten years prior to the Biobank assessment centre (n = 661).
4. Former teachers who left more than 10 years ago. Individuals who were teachers, but left the teaching profession for another career more than ten years prior to the Biobank assessment centre (n = 2,214).

Within this part of our analysis, the sample is restricted to individuals below age 65 at the time of the assessment centre, and who were still employed[5]. Table 1 provides some descriptive information about how the background characteristics of the Biobank sample compares to the estimates of the population of age 40-65-year-old teachers (based upon

[4] Around one-third of those who left teaching within the last five years had moved into another job in education, such as becoming a school inspector, private tutor or teaching assistant, while around two-thirds were employed in a job outside of education.

[5] The occupational questionnaire was typically completed in the summer of 2015; one year before the mental heath questionnaire. We assume that anyone who was recorded as a teacher when they completed the occupational history questionnaire was also a teacher when they completed the mental health questionnaire.

For Peer Review Only

nationally representative sample surveys). On the whole, the Biobank sample is reasonably similar to these population estimates, at least in terms of the observable characteristics considered.

<< Table 1 >>

The first set of results presented in the main body of the paper are based upon the following regression model:

$$link\ (O) = \alpha + \beta.T + \tau.D + \theta.P + \varphi.M + \rho.F + \sigma.L + \gamma.H + \varepsilon \quad (1)$$

Where:

Link () = The appropriate GLM link function for the outcome variable of interest. This will either be identity link for continuous variables (estimated by Ordinary Least Squares) or the logit link for binary/ordered categorical variables (estimated by logistic / ordinal logistic regression).

O = One of the outcomes collected during the assessment centre (as described above).

T = A vector of dummy variables capturing whether the individual was a teacher or a former teacher when they participated in the assessment centre.

D = A vector of demographic background variables such as age, gender, socio-economic status, household income, age they left education, a measure of fluid intelligence, whether they hold a degree and whether born outside of the UK.

P = Reported spells of depression before the individual entered teaching.

M = Month that they completed the Biobank assessment centre.

F = Family history of mental illness (reported that their mother, father or sibling suffered from depression or anxiety).

L = An indicator of whether a major life event (e.g. divorce, severe financial problems, had a relative die) occurred within the last two years.

For Peer Review Only

H = A vector of variables capturing different aspects of household structure, including whether the respondent has a partner in the household, whether there are children in the household and household size.

Multiple imputation using chained equations is used to account for missing data within the controls. The parameter of interest is β; this captures whether individuals who use to be teachers but left for another job (i.e. former teachers) have better or worse mental health outcomes than those individuals who have chosen to remain within the teaching profession (conditional upon the factors controlled for within the model).

In Appendix A, B and C we test the robustness of these results. First, we estimate alternative specifications of these models, variously including and excluding different control variables. Second, Inverse Probability Weighting (IPW) is used as an alternative estimation approach. Third, we alter the estimation approach for continuous outcomes from OLS to zero-inflated count models.

Outcomes measured during the follow-up assessment centre

The main limitation of the analytical approach outlined above is that we can only control for a limited number of potentially confounding background characteristics. Ideally, if one wishes to make causal statements about the effect teaching has upon mental health outcomes (rather than interpreting results are purely correlational), one would need to control for all factors that are both (a) associated with the decision to leave teaching and (b) are also associated with future mental health outcomes.

The second part of our analysis takes a step closer towards reaching this goal. Specifically, we now use responses provided during the follow-up assessment centre as our outcome variables, with data from the 2006-2010 assessment centre acting as a rich set of additional controls. In other words, in this longitudinal analysis, we are interested in *change* in mental health outcomes between the two timepoints.

When using this subset of the Biobank data, our primary interest is individuals who were recorded as working as a teacher within either of the assessment centres, were still employed at the time of the second assessment centre and who were below retirement age (younger than 65). We then focus upon the following groups:

For Peer Review Only

- Always teachers. Individuals who were teachers at both the baseline and follow-up assessment centres (n = 925).
- Leavers. Individuals who were teachers at the baseline assessment centre, but not the follow-up (n = 167)[6].
- Joiners. Individuals who were not teachers at the baseline assessment centre, but were at follow-up (n = 176).

A comparison of the background characteristics of these groups to all teachers who participated in the initial Biobank assessment centre can be found in Appendix D. The average amount of time elapsed between the two assessment centre time points within the sample we use within our analysis is around seven years, with an average age of 58 at follow-up.

Our empirical approach is otherwise similar to that presented above, with the main difference being that we can now control for the extremely rich health data collected from participants within the initial assessment centre. Specifically, the model becomes:

$$link\ (O) = \alpha + \beta.T + \tau.D + \varphi.M + \rho.F + \delta.BC + \varepsilon \quad (2)$$

Where:

O = One of the mental health outcomes collected during the assessment centre (as described above).

T = A vector of dummy variables capturing whether the individual was working as a teacher at both assessment centres (reference group), had left the teaching profession between the two time points or who had joined the profession.

D = A vector of demographic background variables such as age, gender, socio-economic status, age they left education, a measure of fluid intelligence, whether they hold a degree and whether a partner lives in the same household.

M = Month that they completed the Biobank assessment centre.

[6] Around one-third of those who left teaching moved into another job in education, such as becoming a school inspector, private tutor or teaching assistant, while around two-thirds were employed in a job outside of education.

For Peer Review Only

F = Family history of mental illness (reported that their mother, father or sibling suffered from depression or anxiety).

BC = Baseline (assessment centre) controls. This includes depression, anxiety, medications, medical conditions, general health, happiness, job satisfaction, alcohol consumption and sleep as reported during the 2006-2010 assessment centre.

Imputation is again used to account for missing data[7], while robustness tests using alternative model specifications and inverse probability weighting estimates are provided in Appendix A and B. The β estimates from this model now reveal whether individuals who have recently left teaching (and those who have recently joined) have better or worse mental health outcomes than their peers who have worked as teachers throughout this period.

Outcomes measured within the 2016 mental health questionnaire

We follow a similar approach in our analysis of the 2016 mental health questionnaire data; information collected during the initial assessment centre act as a rich set of controls, while responses to the 2016 mental health questionnaire are the outcome measures. Our variable of interest is based upon the information provided in the occupational history questionnaire and is defined as follows:

1. Always teachers (reference group). Individuals who were employed as teachers at both baseline (initial assessment centre) and at follow-up (2016 mental health questionnaire). N= 1,715.
2. Leavers. Those who were employed as teachers at baseline, but employed in another job at follow-up. N= 360[8].
3. Joiners. Those who not teachers at baseline, but were employed as a teacher at follow-up. N= 368.

[7] The amount of missing data is small for most covariates. The small number of covariates with large amounts of missing data are due to those questions only being included in assessment centres taken at later dates (e.g. questions about happiness were only included in later iterations of the Biobank assessment centre questionnaire) and not because of selective non-response. see Appendix E for further details on missing data by covariate.

[8] Around half of those who left teaching entered another job in education (e.g. become an inspector, teaching assistant), while the other half were working outside of education.

For Peer Review Only

Within this part of the analysis, the sample is restricted to those who were still employed and below retirement age (65) when they completed the mental health questionnaire. After making these restrictions, the average age of this analysis sample is 56. A comparison of the background characteristics of these groups to all teachers who participated in the initial Biobank assessment centre can be found in Appendix D. Imputation is again used to account for missing covariate data, while the substantive regression models are specified as outlined in the sub-sections above.

Effect sizes

Results are presented in terms of effect sizes; Cohen's d for continuous outcomes and odds-ratios for binary or ordinal outcomes. With respect to Cohen's d, we interpret values below 0.1 as evidence of essentially no effect. This is based upon two observation. First, meta-analyses have reported much larger effect sizes with respect to the impact of mental health interventions delivered in the workplace. For instance, Carolan, Harris, & Cavanagh (2017) reported an effect size of 0.37 for the impact of occupational mental health interventions upon psychological wellbeing. Second, we argue that an effect size of 0.1 would be substantively very small in this context. For instance, say that a person who left teaching scored an effect size 0.1 lower on the depression scale than those who remained in the profession. This would mean that there is only around a 52.8% chance that a person picked at random from those who have continued to be teachers will have a higher score on the depression scale than a person picked at random from the group who quit teacher for another job. In other words, the probability of suffering depression amongst members of these two groups would be little more than equal. We therefore consider any effect size below 0.1 as trivially small.

4. Results

Left teaching before the initial 2006-2010 assessment centre

The results focusing upon measures collected at the initial assessment centre can be found in Table 2. Starting with anxiety/depression, there is some limited evidence that former teachers have better outcomes on these measures than current teachers. Those individuals who left teaching within the last five years did score slightly lower on the self-reported depression

For Peer Review Only

scale than their peers who were still working as teachers (effect size difference of 0.10), though they were no less likely to report taking prescription medicines for common mental health problems (odds ratio = 0.95). Those who had left teaching within the last five years also scored slightly lower, on average, on the neuroticism scale (effect size 0.13). These differences are, however, quite modest in terms of magnitude. Differences are similar (or slightly smaller) when comparing current teachers to those who left the profession more than five years ago. Evidence of a link between teaching and these outcomes is hence mixed.

<< Table 2 >>

The next set of estimates turns to the issue of sleep. There is little evidence of a difference between current and former teachers in terms of the amount of sleep they get over a typical 24-hour period; differences when expressed as an effect size are all below 0.10. Similarly, individuals who left teaching were slightly less likely to say that they had trouble falling asleep (or woke up during the night) than the reference group (current teachers). However, the effect size is again small (odds ratio ≈ 0.90). Any benefits from quitting teaching for one's quality and quantity of sleep are hence likely to be small (if at all).

The third set of outcomes presented in Table 2 refer to the consumption of alcohol. Former teachers are found to drink slightly more regularly than current teachers, though the difference is again relatively modest in magnitude (the estimated odds ratio is, at most, around 1.2). Furthermore, no difference is found between groups in terms of number of units of alcohol consumed each week. This suggests that former teachers drink roughly the same amount as individuals who have remained in teaching,

Finally, the last set of estimates presented in Table 2 refer to self-reported happiness with different aspects of life. A similar pattern again emerges. The estimated odds ratios mostly hover around one, fluctuating between 0.9 and 1.1, with no clear or obvious pattern. This holds true across most of the five specific areas of life teachers were asked about (e.g. finances, friends, family and health) as well as happiness in life overall. The one exception, where there is a sizeable difference, is happiness at work. Those who quit teaching recently – within the last five years – are happier in their jobs than those who have remained in the profession (odds-ratio 1.42). Yet our overall interpretation of the results presented in Table 2

For Peer Review Only

is that individuals who choose to leave teaching are not generally happier in their life than those continue working in this career.

Joined / left teaching between the initial assessment centre and follow-up assessment centre

Table 3 replicates the analysis presented in the sub-section above, but is now based upon the longitudinal Biobank data, focusing upon those who attended the assessment centre twice. Recall that this allows us to measure change in outcomes over time, and how these relate to whether individuals moved out, into or remained in the teaching profession. There are perhaps two key points of note (over and above our discussion of the results presented in Table 2).

<< Table 3 >>

First, there is some suggestion that that individuals who leave teaching report slightly lower levels of anxiety and depression than those who have remained in the teaching profession. Former teachers score 0.07 standard deviations lower on the self-reported depression scale, while also being somewhat less likely to report taking prescription medicines for common mental health problems (odds ratio 0.45) or reported suffering from depression/anxiety as a medical condition (odds ratio 0.78). Interestingly, those who entered the teaching profession since the baseline assessment centre scored slightly higher on the self-reported depression scale than individuals who were working as teachers at both time points (effect size 0.13) and those who had left teaching for another career (effect size 0.20). Nevertheless, the key message from Table 3 is that leaving teaching to pursue another career may only bring small benefits for one's mental health (if any at all).

Second, the final set of estimates in Table 3 (capturing self-reported happiness) potentially help strengthen the evidence that those individuals who leave teaching have higher levels of job satisfaction than those who continue to work as a teacher as their career. In particular, those who left teaching for other employment reported higher levels of satisfaction with their work than those who stayed in teaching (odds ratio 1.41). Yet this result is very much specific to the work domain; there is no evidence that those who quit teaching were happier with their health, friendships, family, health or, indeed, with life in general.

For Peer Review Only

The final point to note from Table 3 is that the results with respect to sleep and alcohol consumption are largely the same as Table 2. In other words, there is little evidence that working as a teacher has an impact upon difficulties with sleeping and alcohol consumption.

Joined / left teaching between the initial assessment centre and the mental health follow-up

Finally, Table 4 presents results from our second longitudinal analysis, where the initial 2006-2010 assessment centre is the baseline, while the 2016 mental health questionnaire captures the outcomes.

<< Table 4 >>

Table 4 again suggests that, if there are any differences in depression or anxiety between current and former teachers, they are trivially small. Differences in terms of effect sizes are consistently below 0.1 on the anxiety and depression scales (e.g. there is just a 0.06 standard deviation difference between current and former teachers on the self-reported depression scale). Similarly, there is little difference between current and former teachers in terms of their happiness (either overall or with their health), alcohol consumption and whether they believe that their life is meaningful.

Additional analyses in appendices

A series of additional analyses are presented within the appendices, investigating the sensitivity of our results to the different methods used. In summary:

- Appendix A investigates how results change when using a different set of controls. In this we continue to find some evidence that those individuals who left teaching are somewhat happier in their work. Otherwise, the results remained mixed, with no clear evidence that those who left teaching have clearly better mental health and wellbeing outcomes than other groups.
- Appendix B uses Inverse Probability Weighting (IPW) to adjust for covariates, rather than regression modelling. The clearest result is again that those individuals who leave teaching are somewhat happier with their work, though differences for the other outcomes tend to be small and often differ across the different approaches.

For Peer Review Only

- Appendix C alters the regression modelling approach used the analysis of the anxiety, depression and Neuroticism outcome scales, accounting for the fact that many people are clustered at a single score (zero). There continues to be little evidence of sizeable differences in these outcomes between individuals who decide to join, leave and remain in the teaching profession.
- Appendix F provides alternative estimates for our models using the follow-up assessment centre and mental health questionnaire data. In particular, all prior outcome measures are removed from the model, to investigate the extent that these factors may confound the results. Interestingly, the change in the coefficients between model specifications is relatively small. This may either suggest that (a) the selection mechanism for leaving teaching for another job is relatively weak (or at least not driven by concerns about wellbeing / mental health) or (b) that selection into/out of teaching is being driven by factors that the Biobank dataset does not measure.

Summary

In summary, the evidence for whether leaving teaching leads to lower levels of depression and anxiety is mixed. Although small improvements in these outcomes were observed within a subset of our analyses, at other times no impact was found. Our conclusion is therefore that any reduction in anxiety / depression from leaving teaching is likely to be (on average) very small, at best.

Consistent with Bamford and Worth (2017), there is some suggestion that those who decide to quit teaching end up being somewhat happier in their work than those who choose to remain. Yet, critically, this does not seem to translate into greater levels of happiness in other areas of life, including satisfaction with health or happiness overall. Consequently, the benefits of leaving teaching for one's happiness seems to be relatively minor, and concentrated in satisfaction with work. Finally, there is little evidence that the decision to leave teaching has any meaningful impact upon alcohol consumption (a mechanism many adults use to cope with stress) or quality of sleep (a marker of anxiety).

This leads us to reach an overall conclusion that leaving teaching for another job is unlikely to bring significant benefits to well-being or mental health.

For Peer Review Only

5. Conclusions

It is widely thought that teaching is a demanding job, due to the long working hours, the stress imposed by the accountability system and the challenge of having to manage often disruptive classrooms (De Carlo et al., 2019). Across the UK, and England in particular, this has led to many teachers considering a change of career (Perryman & Calvert, 2019). Indeed, in a recent international study, half of secondary teachers in England said that they wondered whether they should have chosen another profession – higher than in almost any other country across the world (*author cite*). Critically, in a recent survey of teachers in England who quit the profession, half said that their job was making them ill (Perryman & Calvert, 2019) and a high number of teachers interviewed in the Teacher Wellbeing Index (2018) experienced anxiety, depression or acute stress (Education Support Partnership, 2018). But is life really that much better for those individuals who decide to quit teaching to pursue a different job? This paper has presented important new evidence on this issue, with a particular focus upon how leaving teaching is linked to well-being and mental health.

Looking across an array of outcomes, and using several different empirical approaches, we have found little evidence of a link between leaving teaching, lower prevalence of mental health problems and higher levels of general well-being. Throughout our analysis, effect sizes have been small and often changed in both magnitude and direction depending upon the empirical approach taken. The one exception is happiness with work (job satisfaction), where we find a fairly consistent improvement for those who have recently left the teaching profession. Two observations are important in interpreting this finding. First, this is consistent with empirical research a range of occupations, which finds that job satisfaction tends to fall in the period prior to an individual quitting a job, before rising during the early stages of their new employment (Chadi & Hetschko, 2018; Gielen, 2013; Longhi et al., 2019). Second, we do not observe reductions in job satisfaction for those joining the teaching profession. Taken together, this suggests that the increase in job satisfaction for those leaving teaching is not indicative of any particular problem with teaching. Rather, it is more likely to reflect a natural process by which those who are less suited to the job move into alternative occupations (Gielen, 2013).

For Peer Review Only

There are, of course, limitations with this paper and the need for future research. Five important issues stand out. First, some of the occupational career data has been based upon respondents recalling and accurately reporting such information. Although our focus upon relatively recent timepoints should limit the impact this has upon our analysis, collection of prospective longitudinal data from a cohort of teachers (tracking their entry and exits from the profession) would represent an important advance in the literature. Second, although a number of potential confounding factors were controlled within our analysis, our estimates continue to rely upon a (untestable) selection-upon-observables assumption if they are to be interpreted as causal effects. Third, while we have considered some "harder" evidence of poor mental health (e.g. prescription of antidepressant medicines), many of the outcome measures we investigated are based upon self-completion questionnaires. Although this is standard within this literature (Lenderink & Zoer, 2012), further work using a wider array of outcome data (e.g. primary care records) would represent a further step forward in this field. Fourth, the number of current and former teachers included within our analysis is relatively modest (in the hundreds rather than the thousands). Larger samples in the future would likely yield more precise results. Finally, the Biobank data is a convenience sample focused upon middle-aged participants, which is not representative of the wider teacher population. In particular, all participants were aged between 40 and 65, meaning it is not possible to generalise our findings to younger age groups. Future work using nationally representative data would undoubtedly enhance the external validity of our findings.

Despite these limitations, we believe that this paper has helped to advance our knowledge of teacher well-being and their mental health. At a time when many teachers are thinking about leaving for another career, it is vital that they are fully informed about the likely consequences. For those teachers who are not satisfied with their work, changing jobs may lead to an increase in job satisfaction. However, our results suggest that quitting teaching for alternative employment is unlikely to lead to improvements in general well-being or mental health. With respect to the latter outcomes, teaching does not stand out relative to other occupations.

For Peer Review Only

Data availability

The data that support the findings of this study are available from the UK Biobank. Restrictions apply to the availability of these data, which were used under license for this study. Data are available from https://www.ukbiobank.ac.uk/with the permission of the UK Biobank.

Ethical guidelines

The paper was produced following the BERA ethical guidelines (https://www.bera.ac.uk/). The project was approved by the UCL Institute of Education ethics committee (REC 1158).

Conflicts of Interest

The authors have no conflict of interest.

Funding

The paper has been funded by the Nuffield Foundation. The Nuffield Foundation is an independent charitable trust with a mission to advance social well-being. It funds research that informs social policy, primarily in Education, Welfare, and Justice. It also funds student programmes that provide opportunities for young people to develop skills in quantitative and scientific methods. The Nuffield Foundation is the founder and co-funder of the Nuffield Council on Bioethics and the Ada Lovelace Institute. The Foundation has funded this project, but the views expressed are those of the authors and not necessarily the Foundation. Visit www.nuffieldfoundation.org.

For Peer Review Only

References

Appleton, A., & James, R. (2018). Adults (18-75) in the UK who drink alcohol for coping reasons. Drinkaware research report. Accessed 04/07/2019 from https://www.drinkaware.co.uk/media/1852/drinking-to-cope-jan-2018.pdf

Bamford, S., & Worth, J. (2017). Is the grass greener beyond teaching? Slough: NFER. Accessed 04/07/2019 from https://www.nfer.ac.uk/teacher-retention-and-turnover-research-research-update-3-is-the-grass-greener-beyond-teaching

Berryhill, J., Linney, J., & Fromewick, J. (2009). The effects of education accountability on teachers: are policies too stress provoking for their own good? *International Journal of Education Policy & Leadership, 4*(5), 1-14.

Borman, G. D., & Dowling, N. M. (2008). Teacher attrition and retention: A meta-analytic and narrative review of the research. *Review of Educational Research, 78*(3), 367-409.

Culver, S., Wolfle, L., & Cross, L.H. (1990). Testing a model of teacher satisfaction for blacks and whites, *American Educational Research Journal, 27*(2), 323-349.

Carolan, S., Harris, P., & Cavanagh, K. (2017). Improving employee well-being and effectiveness: systematic review and meta-analysis of web-based psychological interventions delivered in the workplace. *Journal of Medical Internet Research, 19*(7).

Chadi, A., & Hetschko, C. (2018). The magic of the new: How job changes affect job satisfaction. *Journal of Economics & Management Strategy*, *27*(1), 23-39.

Clark, A. (2010). Work, jobs, and well-being across the millennium. *International differences in well-being*, 436-468.

Creek, J., & Hughes, A. (2008). Occupation and health: a review of selected literature. *British Journal of Occupational Therapy*, *71*(11), 456-468.

De Carlo, A., Girardi, D., Falco, A., Dal Corso, L., Di Sipio, A. (2019). When does work interfere with teachers' private life? An application of the job demands-resources model. *Frontiers in Psychology, 10*(1).

De Matteis, S., Jarvis, D., Young, H., Young, A., Allen, N., Potts, J., Darnton, A., Rushton, L., & Cullinan, P. (2017). Occupational self-coding and automatic recording (OSCAR): a novel web-based tool to collect and code lifetime job histories in large population-based studies. *Scandinavian Journal of Work, Environment and Health, 43*(2), 181-186.

den Brok, P., Wubbels, T., & Van Tartwijk, J. (2017). Exploring beginning teachers' attrition in the Netherlands. *Teachers and Teaching*, 23(8), 881-895.

Dolton, P., & Klaauw, W. V. D. (1995). Leaving teaching in the UK: A duration analysis. *The Economic Journal*, *105*(429), 431-444.

Education Support Partnership, 2018. *Teacher Wellbeing Index 2018*. Accessed 04/11/2019 from

https://www.educationsupportpartnership.org.uk/sites/default/files/teacher_wellbeing_index_2018.pdf

Freeman, D., Sheaves, B., Goodwin, G. M., Yu, L. M., Nickless, A., Harrison, P. J., ... & Hinds, C. (2017). The effects of improving sleep on mental health (OASIS): a randomised controlled trial with mediation analysis. *The Lancet Psychiatry, 4*(10), 749-758.

Fujishiro, K., Farley, A., Kellemen, M., & Swoboda, C. (2017). Exploring associations between state education initiatives and teachers' sleep: A social-ecological approach. *Social Science and Medicine, 191*, 151-159.

Gallagher, M., Muldoon, O. T., & Pettigrew, J. (2015). An integrative review of social and occupational factors influencing health and wellbeing. *Frontiers in Psychology*, *6*.

Gielen, A. C. (2013). Repeated job quits: stepping stones or learning about quality?. *IZA Journal of European Labor Studies*, *2*(7).

Howes, L. M., & Goodman-Delahunty, J. (2015). Teachers' career decisions: Perspectives on choosing teaching careers, and on staying or leaving. *Issues in Educational Research, 25*(1), 18.

Johnson, S., Cooper, C., Cartwright, S., Doland, I., Taylor, P., & Millet, C. (2005). The experience of work-related stress across occupations. *Journal of Managerial Psychology,* 20(2), 178-187.

Kidger, J., Brockman, R., Tilling, K., Campbell, R., Ford, T., Araya, R., King, M., & Gunnell, D. (2016). Teachers' wellbeing and depressive symptoms, and associated risk factors: A large cross-sectional study in English secondary schools. *Journal of Affective Disorders, 192*, 76-82.

Longhi, S., Nandi, A., Bryan, M., Connolly, S., & Gedikli, C. (2019). *Do all job changes increase wellbeing?* What Works Wellbeing.

Mérida-López, S., & Extremera, N. (2017). Emotional intelligence and teacher burnout: A systematic review. *International Journal of Educational Research, 85*, 121-130.

Lenderink, A., & Zoer, I. (2012). *Review on the validity and reliability of self-reported work-related illness*. Health and Safety Executive Research Report RR903. Accessed 04/07/2019 from http://www.hse.gov.uk/research/rrpdf/rr903.pdf

Mercer, D., & Evans, B. (1991). Professional myopia: Job satisfaction and the management of teachers. *School Organization, 11*(3), 291-301.

Mérida-López, S. & Extremera, N. (2017). Emotional intelligence and teacher burnout: A systematic review. *International Journal of Educational Research, 85*, 121-130.

McCormack, A., Edmondson-Jones, M., Fortnum, H., Dawes, P., Middleton, H., Munro, K., & Moore, D. (2015). Investigating the association between tinnitus severity and symptoms of depression and anxiety, while controlling for neuroticism, in a large middle-aged UK population. *International Journal of Audiology, 54*(9): 599-604.

McIntyre, J. (2010). Why they sat still: the ideas and values of long-serving teachers in challenging inner-city schools in England. *Teachers and Teaching: theory and practice, 16*(5), 595-614.

Perryman, J., & Calvert, G. (2019). What motivates people to teach, and why do they leave? Accountability, performativity and teacher retention. *British Journal of Educational Studies*, *68*(1), 3-23.

Peters, E., John, A., Bowen, R., Baetz, M., & Balbuena, L. (2018). Neuroticism and suicide in a general population cohort: results from the UK Biobank Project. *BJPsych Open*, *4*(2), 62–68.

For Peer Review Only

Ravesteijn, B., van Kippersluis, H., & van Doorslaer, E. (2013). The Contribution of Occupation to Health Inequality. *Research on Economic Inequality*, *21*, 1-20.

Ravesteijn, B., Kippersluis, H. V., & Doorslaer, E. V. (2018). The wear and tear on health: What is the role of occupation?. *Health Economics*, *27*(2), e69-e86.

Scheuch, K., Haufe, E., & Seibt, R. (2015). Teachers' health. *Deutsches Ärzteblatt International*, *112*(20), 347.

Smithers, A. (1990). *Teacher Loss*. Manchester: University of Manchester.

Stansfeld, S., Rasul, F, Head, J., & Singleton, N. (2011). Occupation and mental health in a national UK survey. *Social Psychiatry and Psychiatric Epidemiology 46*(2), 101–110.

Tapper, J. (2018, 13 May). Burned out: why are so many teachers quitting or off sick with stress? *The Guardian.*

Taylor, M., Rode, L., Bjørngaard, J., Taylor, A., Bojesen, S., Åsvold, B., Gabrielsen, M., Lewis, G., Nordestgaard, B., Romundstad, P., Hickman, M., & Munafò, M. (2018). Is smoking heaviness causally associated with alcohol use? A Mendelian randomization study in four European cohorts, *International Journal of Epidemiology, 47*(4), 1098–1105.

Tye, B., & O'brien, L. (2002). Why are experienced teachers leaving the profession?. *Phi Delta Kappan, 84*(1), 24-32.

For Peer Review Only

Table 1. Characteristics of the Biobank sample compared to population estimates

	Biobank	Population estimate	
		Estimate	Source
Average age	53	51	LFS
% male	27%	27%	LFS
% children in household	53%	44%	LFS
% Partner in household	76%	74%	LFS
% hold a degree	84%	75%	LFS
Average age left school	20	21	LFS
Born outside UK	7%	7%	LFS
Homeowner	95%	94%	LFS
Smoker	6%	4%	APS 2010
% poor general health	1%	1%	NCDS 2008
% fair general health	14%	7%	NCDS 2008
% good general health	63%	65%	NCDS 2008
% excellent general health	22%	26%	NCDS 2008
Total teachers	**16,622**		

Notes: LFS = Labour Force Survey data for 40-65-year-old teachers from January-March sweeps 2007,2008 and 2009. APS = Annual Population Survey data from 2010 for 40-65-year-old teachers. NCDS = National Child Development Survey from 2008 (when respondents were 50-years-old). The NCDS data for 'good' and 'very good' general health has been combined. APS 2010 data based upon information across all education and teaching professionals.

For Peer Review Only

Table 2. The association between leaving/remaining in the teaching profession and mental health outcomes measured in the 2006-2010 Biobank assessment centre.

	Left teaching last 5 years		Left teaching 6 - 10 years ago		Left teaching > 10 years ago	
	Beta	**T-Stat**	**Beta**	**T-Stat**	**Beta**	**T-Stat**
Depression / anxiety						
Self-reported depression (ES)	-0.10	-3.71	-0.08	-2.20	0.00	-0.21
Prescribed medicines (OR)	0.95	-0.31	0.85	-0.81	0.82	-1.57
Self-reported medical condition (OR)	1.07	0.54	0.98	-0.14	1.08	0.77
Neuroticism (ES)	-0.13	-4.19	-0.11	-2.52	-0.08	-3.65
Sleep						
Hours sleep per night (ES)	0.08	3.54	0.01	0.17	0.04	2.22
Trouble falling sleep (OR)	0.90	-1.97	0.87	-1.90	0.89	-2.56
Alcohol						
Frequent drinking (OR)	1.19	3.25	1.16	2.10	1.19	4.26
Number alcohol units per week	0.06	0.27	-0.22	-0.70	0.09	0.50
Happiness						
Happy with work (OR)	1.42	2.91	1.13	0.98	1.09	1.32
Happy with finances (OR)	0.95	-0.50	0.92	-0.63	0.97	-0.35
Happy with friends (OR)	0.85	-1.75	1.00	-0.01	0.91	-1.25
Happy with family (OR)	0.86	-1.60	0.88	-1.01	0.94	-0.79
Happy with health (OR)	0.98	-0.24	0.91	-0.42	1.22	3.67
Overall happiness (OR)	1.01	0.11	0.94	-0.48	1.05	0.66

Notes: Those individuals who were currently teachers at the time of the assessment centre are the reference group. 'ES' refers to estimated effect size for continuous outcomes and 'OR' to the odds ratio for binary/categorical outcomes. Estimates based upon regression models controlling for age, month visited the assessment centre, gender, fluid intelligence score, Townsend socio-economic status index, family history of depression, whether a major negative life event occurred within last two years, whether has a partner living in same household, whether children live in the same household, household size, household income, age left education, whether hold a degree, whether born outside of the UK and whether had ever suffered depression before working as a teacher. Multiple imputation has been used to account for missing covariate data.

For Peer Review Only

Table 3. The association between leaving/remaining in the teaching profession with outcomes measured in the follow-up Biobank assessment centre

	Left teaching		Joined teaching	
	Beta	**T-Stat**	**Beta**	**T-Stat**
Depression / anxiety				
Self-reported depression (ES)	-0.07	-1.01	0.13	1.78
Prescribed medicines (OR)	0.45	-1.47	0.82	-0.41
Self-reported medical condition (OR)	0.78	-0.75	1.20	0.61
Sleep				
Hours sleep per night (ES)	0.04	0.63	-0.04	-0.62
Trouble falling sleep (OR)	0.90	-0.60	0.73	-1.82
Alcohol				
Frequent drinking (OR)	1.00	-0.02	0.77	-1.59
Number units of alcohol per week	-1.60	-1.79	-2.30	-2.57
Happiness				
Happy with work (OR)	1.41	2.09	1.03	0.20
Happy with finances (OR)	0.84	-1.07	0.72	-1.99
Happy with friends (OR)	0.80	-1.31	0.86	-0.88
Happy with family (OR)	1.18	1.03	1.03	0.15
Happy with health (OR)	0.89	-0.70	0.83	-1.11
Overall happiness (OR)	1.02	0.12	0.84	-0.97

Notes: Those individuals who were teachers at both baseline (initial assessment centre) and follow-up (follow-up assessment centre) are the reference group. 'ES' refers to estimated effect size for continuous outcomes and 'OR' to the risk ratio for binary/categorical outcomes. Estimates based upon regression models controlling for age, month visited the assessment centre, gender, whether has a partner living in same household, whether hold a degree, fluid intelligence score, Townsend socio-economic status index, age left education, family history of depression, self-reported depression at baseline, medical conditions and prescriptions reported at baseline, general health reported at baseline, happiness with work/friends/family/finances/health at baseline, alcohol intake at baseline, sleep quality and quantity at baseline, whether suffered prolonged spell of depression before baseline. Multiple imputation has been used to account for missing covariate data.

For Peer Review Only

Table 4. The association between leaving/remaining in the teaching profession and outcomes measured in the 2016 mental health questionnaire.

	Left teaching since baseline		Joined teaching since baseline	
	Beta	**T-Stat**	**Beta**	**T-Stat**
Depression / anxiety				
Self-reported depression (ES)	0.06	0.96	0.05	0.80
Self-reported anxiety (ES)	0.09	1.28	0.06	0.85
Prolonged depress spell since baseline (OR)	1.11	0.51	0.89	-0.61
Self-harm				
Considered self-harm in last 12 months (OR)	1.28	0.72	1.82	1.82
Alcohol consumption				
Frequently drink >6 drinks (OR)	0.89	-0.82	0.73	-2.15
Other				
Happy with health (OR)	0.82	-1.30	1.06	0.35
Feel life is meaningful (OR)	1.03	0.17	0.79	-1.43
Overall happiness (OR)	0.82	-0.95	0.78	-1.68

Notes: Those individuals who were teachers at both baseline (assessment centre) and follow-up (mental health questionnaire) are the reference group. 'ES' refers to estimated effect size for continuous outcomes and 'OR' to the odds ratio for binary/categorical outcomes. Estimates based upon regression models controlling for age, month visited the assessment centre, gender, whether has a partner living in same household, whether hold a degree, fluid intelligence score, Townsend socio-economic status index, age left education, family history of depression, self-reported depression at baseline, medical conditions and prescriptions reported at baseline, general health reported at baseline, happiness with work/friends/family/finances/health at baseline, alcohol intake at baseline, sleep quality and quantity at baseline, whether suffered prolonged spell of depression before baseline. Multiple imputation has been used to account for missing covariate data.

For Peer Review Only

Appendix A. Alternative regression model estimates including different sets of controls

Table A1. The association between leaving/remaining in the teaching profession and mental health outcomes measured in the 2006-2010 Biobank assessment centre. Alternative set of controls

	Left teaching last 5 years		Left teaching 6 - 10 years ago		Left teaching > 10 years ago	
	Beta	T-Stat	Beta	T-Stat	Beta	T-Stat
Depression / anxiety						
Self-reported depression (ES)	-0.18	-3.79	-0.14	-2.05	-0.05	-1.23
Prescribed medicines (OR)	0.89	-0.81	0.78	-1.25	0.77	-2.12
Self-reported medical condition (OR)	1.05	0.41	0.98	-0.13	1.02	0.27
Neuroticism (ES)	-0.12	-4.29	-0.09	-2.34	-0.10	-4.67
Sleep						
Hours sleep per night (ES)	0.09	3.94	0.01	0.39	0.04	2.31
Trouble falling sleep (OR)	0.90	-1.94	0.85	-2.33	0.88	-3.10
Alcohol						
Frequent drinking (OR)	1.15	2.85	1.17	2.37	1.26	5.82
Units of alcohol per week	-0.02	-0.08	-0.13	-0.44	0.22	1.27
Happiness						
Happy with work (OR)	1.49	4.94	1.17	1.25	1.09	1.02
Happy with finances (OR)	0.98	-0.22	0.90	-0.90	1.11	1.50
Happy with friends (OR)	0.85	-1.85	0.96	-0.33	0.92	-1.24
Happy with family (OR)	0.85	-1.84	0.86	-1.26	0.94	-0.95
Happy with health (OR)	1.04	0.48	0.95	-0.45	1.23	2.69
Overall happiness (OR)	1.06	0.74	0.96	-0.32	1.11	1.15

Notes: Estimates based upon regression models controlling for age, month completed assessment centre, gender, mother/father/sibling ever had depression, immigrant status, partner in the household, children in the household, age finished education, whether respondent holds a degree and whether first instance of depression occurred before they became a teacher. Multiple imputation has been used to account for missing covariate data. See notes to Table 2 for further details.

For Peer Review Only

Table A2. The association between leaving/remaining in the teaching profession with outcomes measured in the follow-up Biobank assessment centre. Alternative set of controls.

	Left teaching		Joined teaching	
	Beta	**T-Stat**	**Beta**	**T-Stat**
Depression / anxiety				
Self-reported depression (ES)	-0.05	-0.76	0.11	1.67
Prescribed medicines (OR)	0.58	-1.16	0.75	-0.69
Self-reported medical condition (OR)	0.75	-0.96	1.01	0.03
Sleep				
Hours sleep per night (ES)	0.01	0.19	-0.06	-0.95
Trouble falling sleep (OR)	0.90	-0.62	0.80	-1.47
Alcohol				
Frequent drinking (OR)	1.00	-0.01	0.70	-2.38
Number of units per week	-1.13	-1.30	-1.95	-2.41
Happiness				
Happy with work (OR)	1.26	1.47	1.03	0.22
Happy with finances (OR)	0.77	-1.66	0.81	-1.46
Happy with friends (OR)	0.82	-1.21	0.97	-0.23
Happy with family (OR)	1.13	0.81	1.03	0.18
Happy with health (OR)	0.78	-1.57	0.81	-1.41
Overall happiness (OR)	0.96	-0.25	0.98	-0.15

Notes: Estimates based upon regression models controlling for age, month completed assessment centre, gender, mother/father ever had depression, partner in the household, age finished education and the following measures recorded at the baseline assessment centre: depression scale, prescription of anti-depressants, self-reported medical condition of depression/anxiety/insomnia, happiness (overall, with work and with health), general health, difficulty sleeping and alcohol consumption. Multiple imputation has been used to account for missing covariate data. See notes to Table 2 for further details.

For Peer Review Only

Table A3. The association between leaving/remaining in the teaching profession and outcomes measured in the 2016 mental health questionnaire. Alternative set of controls.

	Left teaching since baseline		Joined teaching since baseline	
	Beta	T-Stat	Beta	T-Stat
Depression / anxiety				
Self-reported depression (ES)	0.05	0.72	0.02	0.37
Self-reported anxiety (ES)	0.07	1.09	0.04	0.68
Prolonged depress spell since baseline (OR)	1.09	0.44	0.87	-0.82
Self-harm				
Considered self-harm in last 12 months (OR)	1.24	0.63	1.66	1.51
Alcohol consumption				
Frequently drink >6 drinks (OR)	0.96	-0.31	0.84	-1.44
Other				
Happy with health (OR)	0.81	-1.63	1.06	0.38
Feel life is meaningful (OR)	0.99	-0.05	0.86	-1.05
Overall happiness (OR)	0.80	-1.22	0.84	-1.17

Notes: Estimates based upon regression models controlling for age, gender, age finished education, partner in household, and the following measures recorded at the baseline assessment centre: depression scale, happiness (overall and with different areas), general health, difficulty sleeping and alcohol consumption. Imputation has been used to account for missing covariate data. See notes to Table 4 for further details.

For Peer Review Only

Appendix B. Inverse probability weighting estimates

Table B1. The association between leaving/remaining in the teaching profession and mental health outcomes measured in the initial Biobank assessment centre. IPW estimates converted to effect sizes for all variables.

	Left teaching last 5 years		Left teaching 6 - 10 years ago		Left teaching > 10 years ago	
	Beta	T-Stat	Beta	T-Stat	Beta	T-Stat
Depression / anxiety						
Self-reported depression	-0.10	-3.65	-0.21	-1.69	-0.01	-0.22
Prescribed medicines	-0.01	-0.20	0.00	-0.76	0.00	-1.51
Self-reported medical condition	0.03	0.87	0.00	0.11	0.00	0.96
Neuroticism	-0.13	-4.11	-0.10	-2.55	-0.08	-3.42
Sleep						
Hours sleep per night	0.10	3.40	0.01	0.25	0.03	1.77
Trouble falling sleep	-0.06	-1.96	-0.04	-1.93	-0.03	-2.44
Alcohol						
Frequent drinking	0.09	3.21	0.20	2.60	0.18	3.87
Units alcohol per week	0.01	0.35	-1.88	-0.70	0.74	0.44
Happiness						
Happy with work	0.19	4.12	0.09	1.94	0.02	0.90
Happy with finances	-0.01	-0.34	-0.01	-0.32	0.00	0.18
Happy with friends	-0.08	-1.74	0.00	-0.04	-0.02	-0.96
Happy with family	-0.07	-1.48	-0.05	-0.92	-0.01	-0.47
Happy with health	0.00	0.06	-0.02	-0.45	0.07	2.94
Overall happiness	-0.02	-0.34	-0.01	-0.36	0.01	0.39

Notes: Those individuals who were currently teachers at the time of the assessment centre are the reference group. Figures have been converted into an approximate effect size by dividing the estimate by the sample standard deviation for the variable. Estimates based upon IPW, including the following variables in the matching model: age, month of assessment centre, gender, fluid intelligence, socio-economic status, family history of depression, major negative life event occurred in last two years, whether born in UK, household income, household structure, age left school, whether hold a degree and whether experienced depression before becoming a teacher. Single imputation used to account for missing covariate data.

For Peer Review Only

Table B2. The association between leaving/remaining in the teaching profession with outcomes measured in the follow-up Biobank assessment centre. IPW estimates converted to effect sizes for all variables.

	Left teaching		Joined teaching	
	Beta	**T-Stat**	**Beta**	**T-Stat**
Depression / anxiety				
Self-reported depression	-0.08	-1.08	0.14	1.80
Prescribed medicines	-0.11	-1.58	-0.01	-0.19
Self-reported medical condition	-0.06	-0.83	0.04	0.47
Sleep				
Hours sleep per night	0.03	0.32	-0.06	-0.73
Trouble falling sleep	-0.05	-0.59	-0.14	-1.87
Alcohol				
Frequent drinking	0.00	-0.04	-0.13	-2.29
Number units alcohol per week	-0.13	-1.67	-0.18	-2.44
Happiness				
Happy with work	0.20	2.27	0.02	0.28
Happy with finances	-0.10	-1.21	-0.14	-1.78
Happy with friends	-0.03	-0.40	-0.03	-0.34
Happy with family	0.11	1.38	-0.01	-0.10
Happy with health	-0.06	-0.75	-0.07	-0.95
Overall happiness	0.02	0.22	0.00	-0.02

Notes: Figures have been converted into an approximate effect size by dividing the estimate by the sample standard deviation for the variable. Estimates based upon IPW, including the following variables in the matching model: age, month of assessment centre, gender, whether partner is in the household, fluid intelligence, socio-economic status, family history of depression, age left school and baseline measures of depression, prescriptions for mental health issues, self-reported mental health issue, happiness, general health, quality and quantity of sleep, alcohol intake. Single imputation used to account for missing covariate data.

For Peer Review Only

Table B3. The association between leaving/remaining in the teaching profession and outcomes measured in the 2016 mental health questionnaire. IPW estimates converted to effect sizes for all variables.

	Left teaching since baseline		Joined teaching since baseline	
	Beta	**T-Stat**	**Beta**	**T-Stat**
Depression / anxiety				
Self-reported depression	0.03	0.42	-0.01	-0.08
Self-reported anxiety	0.06	0.88	0.05	0.50
Prolonged depress spell since baseline	0.02	0.43	-0.02	-0.48
Self-harm				
Considered self-harm in last 12 months	0.00	0.02	0.05	0.55
Alcohol consumption				
Frequently drink >6 drinks	-0.08	-1.24	-0.17	-1.96
Other				
Happy with health	-0.13	-1.77	0.00	-0.02
Feel life is meaningful	-0.01	-0.20	-0.19	-1.87
Overall happiness	-0.05	-0.82	-0.14	-1.55

Notes: Figures have been converted into an approximate effect size by dividing the estimate by the sample standard deviation for the variable. Estimates based upon IPW, including the following variables in the matching model: age, gender, whether partner is in the household, socio-economic status, family history of depression, whether hold a degree, age left school and baseline measures of depression, prescriptions for mental health issues, self-reported mental health issue, happiness, general health, sleep and alcohol intake. Single imputation used to account for missing covariate data.

For Peer Review Only

Appendix C. Zero-inflated models

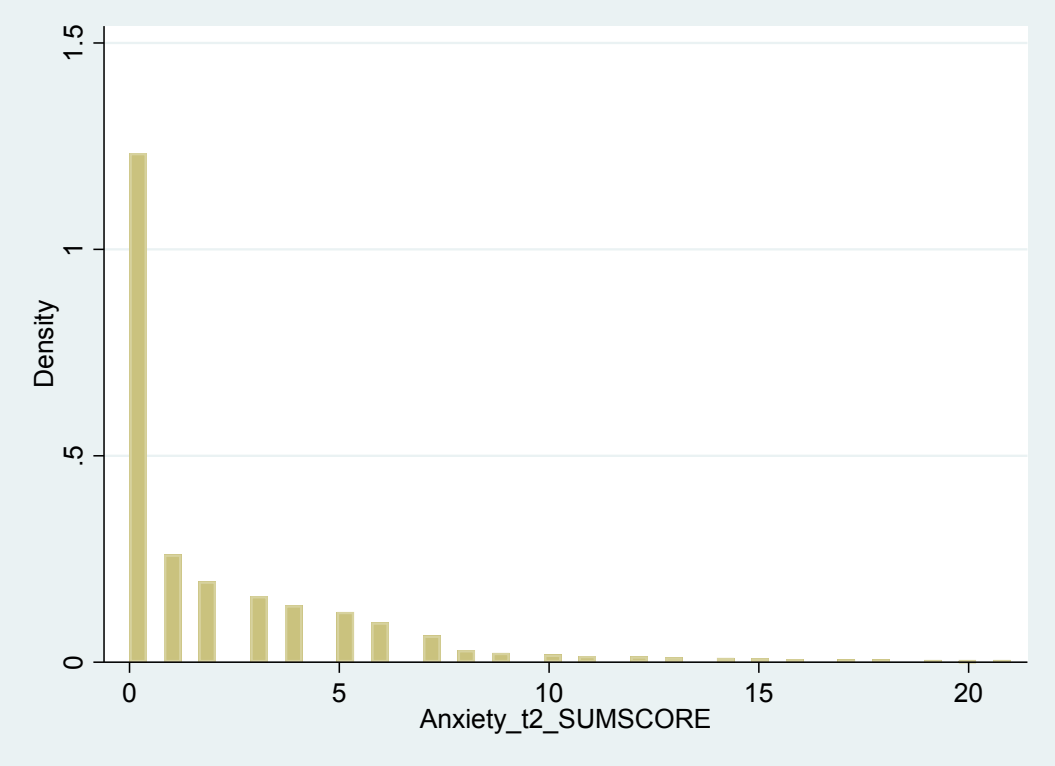
Within the main body of the paper, we treat the following variables as continuous and hence use ordinary least squares regression:

- Anxiety score
- Depression score
- Neuroticism score

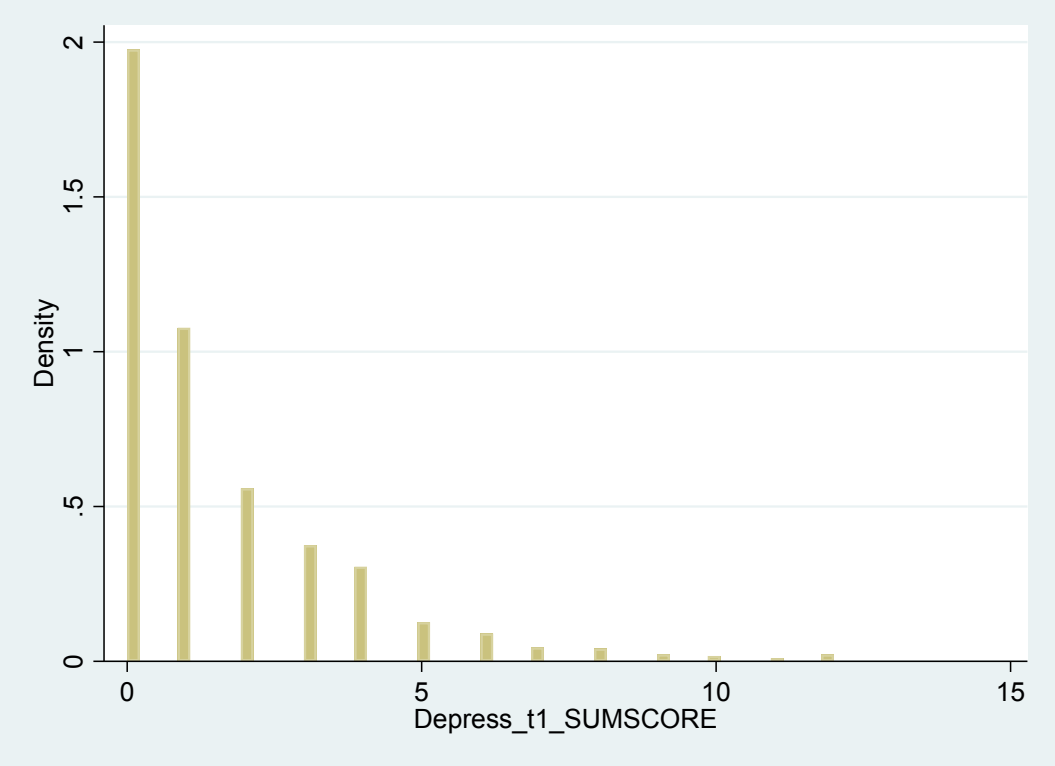
The reason for using OLS regression is that allows for straightforward estimation and interpretation as an effect size. However, as illustrated by Appendix Figures C1, C2 and C3 below, these outcomes may actually be considered to be "zero-inflated" data (meaning there is a large cluster of observations at zero). Consequently, to test the robustness of our results for these outcomes, in this appendix we estimate Zero-Inflated Poisson (ZIP) models. ZIP models essentially divide the outcome into two parts. The first is a binary response (logit) model, which estimates the probability of whether the outcome is zero or not. The second part is a count (Poisson) model, which models the values greater than zero. We implement these models using the same predictors as discussed in the main body of the paper for both parts of the ZIP model. Results are presented in Appendix Tables C1 to C3. These present estimates as odds-ratios for the 'inflation' model (i.e. the logit model which models whether the probability of the outcome being zero or not) and incidence-rate-ratios for the 'outcome' model (i.e. the Poisson count regression estimating the score on the scale).

Appendix Figure C1. Distribution of the Anxiety sum score variable

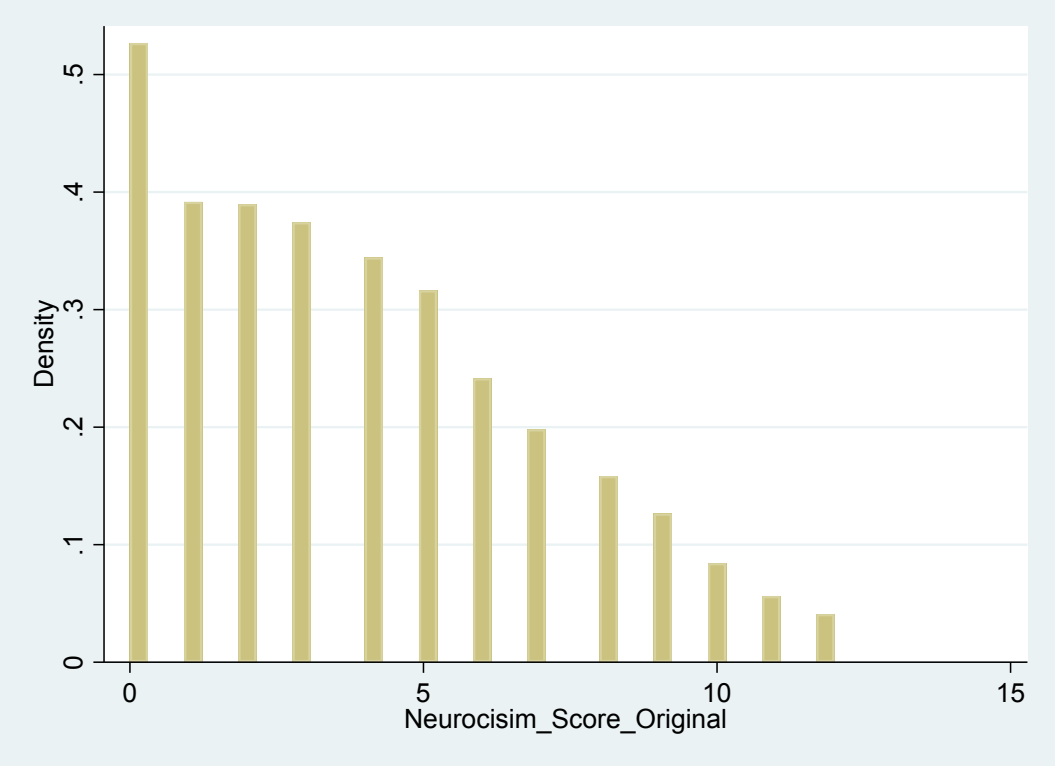
For Peer Review Only



Appendix Figure C2. Distribution of the Depression sum score variable



Appendix Figure C3. Distribution of the Neuroticism sum score variable



Appendix Table C1. Zero-Inflated Poisson estimates for selected outcome variables measures in the initial assessment centre

	Depression scale		Neuroticism scale	
	Effect	T-Stat	Effect	T-Stat
Outcome model (IRR)				
Left in last 5 years	0.90	3.43	0.94	3.92
Left 6-10 years ago	0.90	2.53	0.94	2.87
Left 11+ years ago	0.94	2.49	0.94	4.89
Inflation model (OR)				
Left in last 5 years	1.08	0.87	1.22	2.28
Left 6-10 years ago	0.90	0.83	1.02	0.14
Left 11+ years ago	0.85	2.19	1.06	0.84

Appendix Table C2. Zero-Inflated Poisson estimates for selected outcome variables measures in the follow-up assessment centre

	Depression scale	
	Effect	T-Stat
Outcome model (IRR)		
Left teaching since baseline	0.96	0.93
Joined teaching since baseline	1.04	0.94
Inflation model (OR)		
Left teaching since baseline	0.75	1.35
Joined teaching since baseline	0.62	2.38

For Peer Review Only

Appendix Table C3. Zero-Inflated Poisson estimates for selected outcome variables measures in the mental health questionnaire follow-up

	Depression scale		Anxiety scale	
	Effect	T-Stat	Effect	T-Stat
Outcome model (IRR)				
Left teaching since baseline	1.03	1.36	1.04	0.91
Joined teaching since baseline	1.02	1.08	0.94	1.47
Inflation model (OR)				
Left teaching since baseline	0.88	0.86	0.84	1.15
Joined teaching since baseline	0.74	1.87	0.69	2.46

For Peer Review Only

Appendix D. Characteristics of the assessment centre and mental health questionnaire longitudinal follow-up samples

(a) Assessment centre sample

	Original sample at assessment centre 1	Still teaching assessment centre 2	Left teaching by assessment centre 2	Joined teaching by assessment centre 2
Average age at first contact	53	51	50	51
% male	27%	29%	34%	36%
% children in household at first contact	53%	64%	66%	52%
% Partner in household at first contact	76%	77%	74%	76%
% hold a degree	84%	86%	77%	64%
Born outside UK	7%	6%	6%	5%
Homeowner at first contact	95%	98%	93%	96%
Smoker at first contact	6%	4%	4%	7%
General health at first contact				
% poor	1%	1%	1%	0%
% fair	14%	13%	11%	12%
% good	63%	62%	63%	62%
% excellent	22%	24%	25%	26%
Household income at first contact				
£18,000<	14%	11%	20%	17%
£18,000 - £30,999	38%	37%	36%	45%
£31,000 - £51,999	43%	47%	38%	36%
£52,000+	6%	6%	6%	4%
Total teachers	**16,622**	**925**	**167**	**176**

(b) Mental health questionnaire

	Original sample at assessment centre 1	Teacher at assessment centre and mental health questionnaire	Left teaching	Joined teaching
Average age at first contact	53	48	50	49
% male	27%	21%	31%	28%
% children in household at first contact	53%	71%	66%	69%
% Partner in household at first contact	76%	75%	73%	74%
% hold a degree	84%	92%	88%	78%
Born outside UK	7%	7%	10%	8%
Homeowner at first contact	95%	95%	93%	91%
Smoker at first contact	6%	5%	7%	5%
General health at first contact				
% poor	1%	1%	1%	1%
% fair	14%	12%	15%	15%
% good	63%	61%	58%	63%
% excellent	22%	26%	26%	21%
Household income at first contact				
£18,000<	14%	8%	16%	15%
£18,000 - £30,999	38%	35%	40%	36%
£31,000 - £51,999	43%	50%	37%	40%
£52,000+	6%	7%	6%	9%
Total teachers	**16,622**	**1,715**	**360**	**368**

Appendix E. The percentage of observations missing information in each covariate

	Assessment centre 1 analysis	Assessment centre 2 (longitudinal) analysis	Mental health questionnaire analysis
Age first contact	0%	0%	0%
Gender	0%	0%	0%
Fluid intelligence score	64%	67%	60%
Townsend index	0%	0%	0%
Maternal depression	1%	2%	1%
Paternal depression	3%	3%	3%
Sibling depression	3%	3%	3%
Relative die soon before first contact	0%	0%	0%
Had serious illness soon before fist contact	0%	0%	0%
Divorced soon before first contact	0%	0%	0%
Financial problems soon before first contact	0%	0%	0%
Born outside UK	0%	0%	0%
Partner in household at first contact	0%	0%	0%
Children in household at first contact	0%	0%	0%
Household size	0%	0%	0%
Household income	6%	5%	4%
Whether hold a degree	0%	0%	0%
Experienced depression before teaching	49%	29%	21%
Depression scale at 1st contact	-	3%	3%
Prescribed antidepressants at 1st contact	-	0%	0%
Depression as medical condition 1st contact	-	0%	0%
Happy at first contact	-	67%	60%

For Peer Review Only

Happy at work at first contact	-	66%	60%
General health at first contact	-	0%	0%
Happy with health at first contact	-	66%	60%
Happy with family at first contact	-	67%	60%
Happy with friends at first contact	-	67%	60%
Happy with finance at first contact	-	66%	60%
Hours sleep per night at first contact	-	0%	0%
Insomnia at first contact	-	0%	0%
Alcohol intake at first contact	-	0%	0%
Number of units at first contact	-	0%	0%
Full sample	**20768**	**1268**	**2443**

For Peer Review Only

Appendix F. Alternative estimates for the longitudinal analysis with different sets of controls

In this appendix we repeat our analysis of the follow-up assessment centre and mental health questionnaire data, but now using a different set of control variables. Specifically, we remove from the model the prior outcome measures that were measured in the initial Biobank assessment centre. The motivation behind this exploration is to investigate that the inclusion or exclusion of these controls changes the results. Appendix Table F1 lists the covariates across the two sets of models being compared. Results for the follow-up Biobank Assessment Centre (analogous to those presented in Table 3) can be found in Appendix Table F2. Those for the mental health questionnaire (analogous to those presented in Table 4) can be found in Appendix Table F3. Interestingly, the inclusion or exclusions of the prior outcome measures collected in the initial assessment centre do not seem to substantively alter the results or the key conclusions reached.

Appendix Table F1. Covariates included across the two model specifications

	Without additional controls	With additional controls
Age	Y	Y
Month of assessment centre visit	Y	Y
Gender	Y	Y
Partner in household	Y	Y
Hold a degree	Y	Y
Fluid intelligence score	Y	Y

Townsend socio-economic status index	Y	Y
Age left education	Y	Y
Family history of depression	Y	Y
Medical conditions and prescriptions reported at baseline	-	Y
General health reported at baseline	-	Y
Happiness with work at baseline	-	Y
Happiness with friends at baseline	-	Y
Happiness with family at baseline	-	Y
Happiness with finances at baseline	-	Y
Happiness with health at baseline	-	Y
Alcohol intake at baseline	-	Y
Sleep quality and quantity at baseline	-	Y
Suffered spell of depression before baseline	-	Y

Appendix Table F2. Alternative estimates of the association between leaving/remaining in the teaching profession with outcomes measured in the follow-up Biobank assessment centre (analogous to Table 3)

	Left teaching		Joined teaching	
	Without additional controls	With additional controls	Without additional controls	With additional controls
Depression / anxiety				
Self-reported depression (ES)	-0.05	-0.07	0.10	0.13
Prescribed medicines (OR)	0.56	0.45	0.84	0.82
Self-reported medical condition (OR)	0.82	0.78	1.19	1.20
Sleep				
Hours sleep per night (ES)	0.01	0.04	-0.01	-0.04
Trouble falling sleep (OR)	0.94	0.90	0.69	0.73
Alcohol				
Frequent drinking (OR)	1.03	1.00	0.87	0.77
Number units of alcohol per week	-1.74	-1.60	-2.30	-2.30
Happiness				
Happy with work (OR)	1.40	1.41	1.11	1.03
Happy with finances (OR)	0.87	0.84	0.76	0.72
Happy with friends (OR)	0.80	0.80	0.91	0.86
Happy with family (OR)	1.17	1.18	1.10	1.03
Happy with health (OR)	0.95	0.89	1.02	0.83

For Peer Review Only

Overall happiness (OR)	1.01	1.02	0.94	0.84

Appendix Table F3. Alternative estimates of the association between leaving/remaining in the teaching profession and outcomes measured in the 2016 mental health questionnaire (analogous to Table 4)

	Left teaching since baseline		Joined teaching since baseline	
	Without additional controls	With additional controls	Without additional controls	With additional controls
Depression / anxiety				
Self-reported depression (ES)	0.06	0.06	0.05	0.05
Self-reported anxiety (ES)	0.10	0.09	0.06	0.06
Prolonged depress spell since baseline (OR)	1.01	1.11	0.78	0.89
Self-harm				
Considered self-harm in last 12 months (OR)	1.22	1.28	1.62	1.82
Alcohol consumption				
Frequently drink >6 drinks (OR)	1.00	0.89	1.01	0.73
Other				
Happy with health (OR)	0.95	0.82	0.96	1.06
Feel life is meaningful (OR)	0.96	1.03	0.84	0.79
Overall happiness (OR)	0.8	0.82	0.84	0.78

For Peer Review Only

For Peer Review Only

The mental health and well-being of teachers is an issue of great policy concern. This is particularly true in England, where high workload and the associated stress is thought to be leading to a recruitment and retention crisis within the education profession. But do individuals who decide to leave teaching for another career actually see their well-being and mental health improve? We provide new evidence on this matter for individuals aged between 40 and 65, using the rich information gathered as part of the UK Biobank study. Our analysis shows that individuals who choose to leave teaching are somewhat happier in their work, but do not generally experience any improvement in their general well-being or mental health. We hence caution those middle-aged teachers who are thinking of leaving teaching that the grass may not necessarily be greener on the other side.

1. Introduction

Almost thirty years ago it was observed that "…teaching is an unsettled and unhappy profession at present and there is a problem to be addressed" (Smithers, 1990). A similar statement would not be out of place when describing the teaching profession today. As a job, teaching requires staff to work long hours, to keep up with changing government requirements and to manage often disruptive classrooms (*author cite*). This, of course, all occurs under the watchful eye of the accountability system, with schools (and teachers) judged by how young people perform in high-stakes national examinations.

Many teachers enter the profession for altruistic reasons, yet key issues which motivate many to leave are workload, pupil behaviour and salary (Dolton & Klaauw, 1995; Barmby, 2006). Many teachers experience dissatisfaction due to feeling a lack of control with respect to their working conditions, accompanied by the absence of a platform from which to voice their concerns (Mercer & Evans, 1991). Researchers have observed similar trends for teachers outside of England, including the USA, Australia and the Netherlands, especially in newer teachers (Tye & O'Brien, 2002; Howes and Goodman-Delahunty, 2015; den Brok et al., 2017; *author cite*). However, strikingly, many more teachers in England expressed regret about their career choice than in almost any other industrialised country (*author cite*).

Today's retention rate in the UK is low especially amongst newly qualified teachers (NQTs), and schools located in more socially deprived backgrounds experience a higher turnover of teachers (*author cite*). There is also some suggestive evidence that it is the more able

For Peer Review Only

teachers who are the most likely to leave teaching in search of alternative employment, again impacting on teacher quality, as well as there being strong monetary implications, given the cost of training teachers (Culver et al., 1990; Borman & Dowling, 2008).

Long-serving teachers in urban UK schools appear to remain in their jobs due to the deep connections and emotional ties they forge within their workplaces and communities (McIntyre, 2010). By implication then, teachers who leave the profession may do so either because they cannot forge these connections which provide an adequate level of job satisfaction, or because the challenges of workload, classroom management and salary outweigh the benefits they receive.

Such pressures may lead to work-related stress and, in turn, be detrimental to teachers' mental and physical health (Travers & Cooper, 1993; Berryhill, Linney, & Fromewick 2009; Scheuch & Seibt, 2015; Kidger et al., 2016; Merrida-Lopez & Extremara, 2017; Education Support Partnership, 2018). A number of studies have also suggested that teachers have worse mental health and wellbeing than those who work in other jobs. For instance, Johnson et al. (2003) found teachers to have one of the lowest levels of psychological wellbeing out of the 26 occupational groups they considered. Similarly, Stansfeld et al. (2011) found that teachers were at above average risk of suffering from mental ill-health, with Kidger et al. (2016) indicating that wellbeing was low and depressive symptoms high amongst teachers. Indeed, the stresses and strains of teaching are widely cited for the ongoing teacher retention and recruitment crisis that continues to create a major challenge for England's schools (e.g. Tapper, 2018). It is therefore little wonder that, in a recent nationally representative survey of teachers, around half of those working in secondary schools in England said that they wondered whether they would have been better off working in a different job (*author cite*).

This leads to an important question for teachers who are contemplating a change of career: do those who choose to leave teaching have higher levels of well-being, and a lower likelihood of developing mental health problems, than those who choose to remain? In general, occupation is known to be linked to health and wellbeing (Clark, 2010; Johnson et al., 2005; Ravesteijn, Kippersluis, & Doorslaer, 2013; Ravesteijn, Kippersluis, & Doorslaer, 2018). This is perhaps unsurprising, given the influence that work has on our social interaction with others, our ability to develop and employ new skills, our sense of personal achievement and the

For Peer Review Only

restrictions it imposes on other areas of life e.g. time with family (Creek & Hughes, 2008; Gallagher, Muldoon, & Pettigrew, 2015). Consequently, individuals who are unhappy may switch jobs in an attempt to improve their wellbeing. Indeed, empirical research suggests that voluntary job changes are associated with increased job satisfaction (Chadi & Hetschko, 2018; Gielen, 2013) and improved mental health (Longhi et al., 2019), at least in the short run.

The evidence in relation to teachers is, however, more limited. The study most comparable to ours is the interesting work of Bamford and Worth (2017). Using longitudinal data, these authors tracked job satisfaction, overall life satisfaction and the subjective well-being of a small sample of teachers within the UK. They found that teachers who left the profession experience a large increase in job satisfaction, and a small increase in subjective well-being, compared with those who decided to stay. Yet, as with all studies, the authors noted some key limitations. First, the sample size of teachers was small (e.g. 231 former teachers had left their job within the last year, 107 who had left teaching between one and two years ago and 74 who had left three years previously), with estimates surrounded by quite wide confidence intervals. Second, the focus was upon teachers who quit their job recently (mainly within the last two years) meaning it was not possible to consider the association between leaving teaching and longer-term outcomes. Finally, the Understanding Society dataset analysed collected only limited information about respondents' well-being and mental health. The combination of these factors made it challenging to compare the outcomes of individuals who chose to join, leave and stay in the teaching profession.

Our paper seeks to contribute to this understudied area, with a particular focus upon the well-being and mental health outcomes of current and former teachers aged between 40 and 65 within the UK. It is, to our knowledge, the first study to utilise the UK Biobank dataset to investigate this issue. Through this data, we have access to a wide array of information about respondents' well-being and mental health. This not only includes responses to standardised questionnaires (as have previously been used in this literature) but also prescription of common medicines used for conditions such as anxiety, depression and insomnia. These data can therefore be used to explore the mental health outcomes of current and former teachers across a wide range of important measures. Moreover, having such a rich array of data collected at baseline means that we can more credibly control for potential confounding factors within the longitudinal component of our analysis. Together, this enables us to provide

For Peer Review Only

important new evidence as to whether joining/leaving teaching has a sizeable impact upon a person's mental health, at a time when many teachers are thinking about changing career in search of a better life. It is important to recognise, however, that the data we use are based upon a sample of 40 to 65-year-olds and may not generalise to younger age groups.

To trail our key results, we find little evidence that the grass really is greener for those who quit the teaching profession. Although there is some evidence of an increase in job satisfaction for those who left teaching relatively recently (consistent with the findings of Bamford & Worth, 2017), there is little to suggest that leaving teaching reduces the risk of suffering from mental health problems, such as depression and anxiety. Similarly, there is little evidence to suggest that individuals who choose to quit teaching are happier in their life in general (i.e. outside of work). This leads us to conclude that teaching as an occupational choice is unlikely to be a major cause of poor mental health outcomes per se.

The paper now proceeds as follows. The UK Biobank dataset is described in section 2, with an overview of our empirical methodology following in section 3. Our results are detailed in section 4, with conclusions and policy discussion provided in section 5.

2. UK Biobank data

The initial UK Biobank data collection took place between 2006 and 2010. A total of around half a million volunteers between the ages of 40 and 69 participated in the study. These data therefore form a convenience sample rather than being a random sample that is representative of the wider population. Participants attended an initial assessment centre when they were first recruited, where they completed questionnaires, were interviewed by a trained health professional (in order to collect accurate information about medical conditions and currently prescribed drugs) and underwent some basic health checks (e.g. participants' blood pressure was taken; an electrocardiogram/ECG was conducted). It contains uniquely rich information about health from a very large number of individuals – many of whom were employed as teachers. A number of follow-up questionnaires have been gathered from Biobank participants since the initial assessment centre. In 2016, 117,500 participants completed an online 'occupational career' questionnaire. Respondents were first asked to type into an open text field their job title, start date and end date for each job they held. The respondent was then asked to work through a set of 'drop-down' job lists, with a list of

For Peer Review Only

possible job titles presented on the final screen (from which they were then asked to pick the most suitable). The relevant SOC code was then assigned to the individual, based upon their selection from the drop-down menus (information entered as free-text at the start was used for validation processes only). The data went through a validation process by an expert occupational coder, with reasonably good agreement found (Cohen's Kappa = 0.45). Further details about how the occupational career data has been captured is available within de Matteis et al (2017).

As part of the occupational career questionnaire, some further basic information was asked about each job, such as typical number of hours worked per week, whether it involved shift work and exposure to potentially hazardous substances (e.g. asbestos). For our purposes, this information allows us to identify the year respondents entered and exited the teaching profession, including the occupation that they joined when they left. Critically, this means we can identify both current and former teachers, facilitating comparisons between these two groups.

For each job recorded at the assessment centre or in the occupational history questionnaire, four-digit Standard Occupational Classification (SOC2000) codes are provided within the Biobank database, along with some further information about specific job role. Throughout this paper, we begin by identifying teachers using the following broad set of SOC codes:

- 2312 = Further education teaching professionals
- 2314 = Secondary education teaching professionals
- 2315 = Primary and nursery education teaching professionals
- 2316 = Special needs education teaching professionals

We excluded individuals from analysis if they were identified as working in a non-teaching role. For instance, for respondents with a SOC code of 2315, extra information was used to exclude nursery workers from our definition of 'teachers'[1]. Likewise, this extra information was used to remove further education lecturers, whilst retaining further education teachers.

[1] The Biobank data provides additional information about the job of each respondent, over and above the SOC code. For instance, for SOC code 2315 it provides information on whether the individual is a headteacher or not,

For Peer Review Only

Key measures collected during the 2006-2010 assessment centre

Within our analysis, we make particular use of the following measures collected during the assessment centre:

- Neuroticism score. Neuroticism was measured with the 12-item neuroticism subscale from the short form of the revised Eysenck Personality Questionnaire (EPQ-N). This encompassed 12 questions such as 'do you often feel fed-up', 'do you suffer from nerves', and 'do you often feel lonely', with participants asked to respond either yes, no, don't know or prefer not to say to each item. A total neuroticism score from these 12 items is provided as part of the Biobank dataset, which has been reported to have good levels of internal validity (Cronbach's alpha of 0.84 – Peters et al 2018). Within our analysis, we standardise this scale to mean zero and standard deviation one.
- Depression. Within the self-completion questionnaire, respondents were asked four questions about how they felt over the last two weeks. This included (a) frequency of depressed mood; (b) tiredness / lethargy; (c) unenthusiastic/ disinterest and (d) tenseness/restlessness. Respondents were asked to indicate one of following response options for each (1. Not at all; 2. Several days; 3. More than half the days; 4. Nearly every day; 5. Don't know; 6. Prefer not to say). Following McCormack et al (2015), we combine responses to these questions into an overall depression scale. This is done using a two-parameter item-response theory (IRT) model, which is then standardised to mean zero and standard deviation one.
- Sleep. A wide body of research has shown that mental health problems, such as anxiety and depression, are linked to insomnia and a lack of sleep (Freeman et al., 2017; Fujishiro et al., 2017). As part of the self-completion questionnaire, respondents were asked (a) the number of hours of sleep they typically get within a 24-hour period and (b) whether they either have trouble falling asleep or whether they wake up during the night (1. Never/rarely; 2. Sometimes; 3. Usually).
- Alcohol intake. Previous research has found that up to 40 percent of adults use alcohol as a mechanism to cope with stress (Appleton and James 2018). Within the

and whether they work in a primary or nursery setting. The variable in question is available from here: https://biobank.ctsu.ox.ac.uk/crystal/field.cgi?id=132.

For Peer Review Only

assessment centre, respondents were first asked how frequently they drink alcohol (daily/almost daily; three/four times a week; once/twice a week; one to three times a month; special occasions only; never). For those who said that they drink alcohol more than once or twice per week, they were then asked about average weekly intake of (a) pints of beer/cider; (b) white wine/Champagne; (c) fortified wine; (d) red wine; (e) spirits. For those who said that they drunk alcohol monthly, or only on special occasions, monthly figures were provided. Following Taylor et al (2018) we convert these responses into an approximate number of units of alcohol consumed per week. Both frequency of drinking per week and weekly units of alcohol consumed are considered within our analysis.

- Medications prescribed. As part of an interview with a trained nurse, respondents were asked about prescribed medications. Using this information, we create a binary variable, coded as one if they take frequently prescribed medications for anxiety, depression or insomnia[2], and zero otherwise.
- Medical conditions. In the assessment centre, participants were asked to indicate any medical conditions that they had. If the participant was uncertain of the type of illness they had, then they were asked to describe it to a trained nurse who placed it within a category. They were also asked the date or age when they were first diagnosed with the condition. Our focus is upon reports of depression, anxiety, self-harm, stress and insomnia, coded as one if they reported having one of these conditions, and zero otherwise.
- Happiness. Respondents were asked 'in general, how happy are you' with responses provided using a six-point scale (extremely happy to extremely unhappy)[3].
- Happiness with different aspects of life. Respondents were asked the same question as above, but with the focus being upon a certain aspect of their life. This included how happy they were with their (a) work; (b) family; (c) finances; (d) friends; (e) health.

[2] If respondents indicated that they were currently prescribed one of the following medications, this dummy variable was coded as one: Citalopram, Escitalopram, Fluoxetine, Fluvoxamine, Amitripyline, Paroxetine, Sertraline, Venlafaxine, Duloxteine, Pregabalin, Cymbalta, Yentreve, Mirtazapine, Anafranil, Prozac, Diazepam, Zopiclone, Temazepam, Nitrazepam.

[3] The questions about happiness and work/job satisfaction were only introduced into the assessment centre questionnaire in 2009. These data are therefore only available for a subset of respondents.

For Peer Review Only

Responses were again provided using a six-point scale (extremely happy to extremely unhappy).

A wide range of other information was also collected within the assessment centre, including demographic background, basic health assessments (blood pressure measurements), illnesses of mother/father/siblings (including depression) and whether selected life events had occurred within the last two years (e.g. death of a family member, divorce, a serious illness). Although not the focus of this paper, a selection of such variables will be used as controls within our analysis.

Follow-up assessment centres

A sub-set of Biobank participants have completed return visits to the assessment centre, providing a longitudinal element to the dataset. These longitudinal follow-ups have been concentrated within certain parts of the country; most notably the assessment centres located at Cheadle (Stockport), Reading and Newcastle. Most of the same data were collected as in the initial assessment centre, including current occupation, currently prescribed medications, current medical conditions, sleep, depression, alcohol consumption and happiness with different aspects of life. Critically, this means we can identify individuals who have changed job between the two assessment centres (e.g. individuals who joined or left the teaching profession) and measure change in the aforementioned measures of mental health and well-being.

The 2016 mental health questionnaire

In 2016, a subset of 137,000 biobank participants completed an additional questionnaire about their mental health. Most of these questionnaires were completed between the 19th of August and the 26th of September 2016. Importantly, this provides a second longitudinal component to the Biobank data, with the 2006-2010 assessment centre acting as the baseline and the 2016 mental health questionnaire as the follow-up. Within our analysis, we make particular use of the following information gathered within the mental health questionnaire:

- Current depression. Respondents were asked "*over the last 2 weeks, how often have you been bothered by any of the following problems*", with nine separate questions then following (e.g. Little interest or pleasure in doing things; feeling down, depressed,

For Peer Review Only

or hopeless; feeling tired or having little energy). They were asked to select one of four responses for each (not at all; several days; more than half the days; nearly every day). A two-parameter IRT model is estimated using these nine items in order to construct a depression scale. We standardise this scale to mean zero and standard deviation one.

- Current anxiety. Respondents answered the same question as presented above for current depression, with seven separate items (e.g. feeling nervous, anxious or on edge; worrying too much about different things; trouble relaxing). A two-parameter IRT model is again used to derive an overall anxiety scale.
- Prolonged periods of depression. First, respondents were asked the following two yes/no questions: "*have you ever had a time in your life when you felt sad, blue, or depressed for two weeks or more in a row*" and "*have you ever had a time in your life when you lost interest in most things like hobbies, work, or activities that usually give you please*". To respondents who answered yes to either of these questions, they were asked to recall the age that they first had such a spell and the age they were when they last had such a spell. For the purposes of this paper, this information can be compared to spells when respondents were and were not working as teachers.
- Alcohol consumption. Respondents were asked: "*In the next two questions, a "drink" is defined as one unit of alcohol. How often do you have six or more drinks on one occasion*?" We use responses to this question to get an indication of the extent that respondents engage in heavy drinking.
- Self-harm. A series of questions were asked about self-harm, including "*many people have thoughts that life is not worth living. Have you felt that way*" and "*have you contemplated harming yourself*?". They were also asked if they had felt this way in the last 12 months or if they had harmed themselves during the last 12 months.
- Happiness. Respondents were asked the same question as in the 2006-2010 assessment centre.
- Happiness with health. Respondents were asked "*In general how happy are you with your HEALTH*?", using the same six response options presented for the happiness scale.

For Peer Review Only

- Feel life is meaningful. Participants were asked "*to what extent do you feel your life to be meaningful*?", with responses provided on a five-point scale (not at all; a little; a moderate amount; very much; an extreme amount).

These are the main outcome measures we consider when analysing the 2016 mental health data.

3. Methodology

Outcomes measured at the initial assessment centre

To begin, we focus upon the 2006-2010 assessment centre measures described above as our outcomes of interest. These outcomes will be compared across the following groups:

1. Current teachers (reference group). Individuals who were teachers when the initial assessment centre took place (n = 16,622).
2. Former teachers who left within the last five years. Individuals who were teachers, but left the teaching profession for another career within the five years prior to the Biobank assessment centre (n = 1,271)[4].
3. Former teachers who left six to ten years ago. Individuals who were teachers, but left the teaching profession for another career between six and ten years prior to the Biobank assessment centre (n = 661).
4. Former teachers who left more than 10 years ago. Individuals who were teachers, but left the teaching profession for another career more than ten years prior to the Biobank assessment centre (n = 2,214).

Within this part of our analysis, the sample is restricted to individuals below age 65 at the time of the assessment centre, and who were still employed[5]. Table 1 provides some descriptive information about how the background characteristics of the Biobank sample compares to the estimates of the population of age 40-65-year-old teachers (based upon

[4] Around one-third of those who left teaching within the last five years had moved into another job in education, such as becoming a school inspector, private tutor or teaching assistant, while around two-thirds were employed in a job outside of education.

[5] The occupational questionnaire was typically completed in the summer of 2015; one year before the mental heath questionnaire. We assume that anyone who was recorded as a teacher when they completed the occupational history questionnaire was also a teacher when they completed the mental health questionnaire.

For Peer Review Only

nationally representative sample surveys). On the whole, the Biobank sample is reasonably similar to these population estimates, at least in terms of the observable characteristics considered.

<< Table 1 >>

The first set of results presented in the main body of the paper are based upon the following regression model:

$$link\ (O) = \alpha + \beta.T + \tau.D + \theta.P + \varphi.M + \rho.F + \sigma.L + \gamma.H + \varepsilon \quad (1)$$

Where:

Link () = The appropriate GLM link function for the outcome variable of interest. This will either be identity link for continuous variables (estimated by Ordinary Least Squares) or the logit link for binary/ordered categorical variables (estimated by logistic / ordinal logistic regression).

O = One of the outcomes collected during the assessment centre (as described above).

T = A vector of dummy variables capturing whether the individual was a teacher or a former teacher when they participated in the assessment centre.

D = A vector of demographic background variables such as age, gender, socio-economic status, household income, age they left education, a measure of fluid intelligence, whether they hold a degree and whether born outside of the UK.

P = Reported spells of depression before the individual entered teaching.

M = Month that they completed the Biobank assessment centre.

F = Family history of mental illness (reported that their mother, father or sibling suffered from depression or anxiety).

L = An indicator of whether a major life event (e.g. divorce, severe financial problems, had a relative die) occurred within the last two years.

For Peer Review Only

H = A vector of variables capturing different aspects of household structure, including whether the respondent has a partner in the household, whether there are children in the household and household size.

Multiple imputation using chained equations is used to account for missing data within the controls. The parameter of interest is β; this captures whether individuals who use to be teachers but left for another job (i.e. former teachers) have better or worse mental health outcomes than those individuals who have chosen to remain within the teaching profession (conditional upon the factors controlled for within the model).

In Appendix A, B and C we test the robustness of these results. First, we estimate alternative specifications of these models, variously including and excluding different control variables. Second, Inverse Probability Weighting (IPW) is used as an alternative estimation approach. Third, we alter the estimation approach for continuous outcomes from OLS to zero-inflated count models.

Outcomes measured during the follow-up assessment centre

The main limitation of the analytical approach outlined above is that we can only control for a limited number of potentially confounding background characteristics. Ideally, if one wishes to make causal statements about the effect teaching has upon mental health outcomes (rather than interpreting results are purely correlational), one would need to control for all factors that are both (a) associated with the decision to leave teaching and (b) are also associated with future mental health outcomes.

The second part of our analysis takes a step closer towards reaching this goal. Specifically, we now use responses provided during the follow-up assessment centre as our outcome variables, with data from the 2006-2010 assessment centre acting as a rich set of additional controls. In other words, in this longitudinal analysis, we are interested in *change* in mental health outcomes between the two timepoints.

When using this subset of the Biobank data, our primary interest is individuals who were recorded as working as a teacher within either of the assessment centres, were still employed at the time of the second assessment centre and who were below retirement age (younger than 65). We then focus upon the following groups:

For Peer Review Only

- Always teachers. Individuals who were teachers at both the baseline and follow-up assessment centres (n = 925).
- Leavers. Individuals who were teachers at the baseline assessment centre, but not the follow-up (n = 167)[6].
- Joiners. Individuals who were not teachers at the baseline assessment centre, but were at follow-up (n = 176).

A comparison of the background characteristics of these groups to all teachers who participated in the initial Biobank assessment centre can be found in Appendix D. The average amount of time elapsed between the two assessment centre time points within the sample we use within our analysis is around seven years, with an average age of 58 at follow-up.

Our empirical approach is otherwise similar to that presented above, with the main difference being that we can now control for the extremely rich health data collected from participants within the initial assessment centre. Specifically, the model becomes:

$$link\ (O) = \alpha + \beta.T + \tau.D + \varphi.M + \rho.F + \delta.BC + \varepsilon \quad (2)$$

Where:

O = One of the mental health outcomes collected during the assessment centre (as described above).

T = A vector of dummy variables capturing whether the individual was working as a teacher at both assessment centres (reference group), had left the teaching profession between the two time points or who had joined the profession.

D = A vector of demographic background variables such as age, gender, socio-economic status, age they left education, a measure of fluid intelligence, whether they hold a degree and whether a partner lives in the same household.

M = Month that they completed the Biobank assessment centre.

[6] Around one-third of those who left teaching moved into another job in education, such as becoming a school inspector, private tutor or teaching assistant, while around two-thirds were employed in a job outside of education.

For Peer Review Only

F = Family history of mental illness (reported that their mother, father or sibling suffered from depression or anxiety).

BC = Baseline (assessment centre) controls. This includes depression, anxiety, medications, medical conditions, general health, happiness, job satisfaction, alcohol consumption and sleep as reported during the 2006-2010 assessment centre.

Imputation is again used to account for missing data[7], while robustness tests using alternative model specifications and inverse probability weighting estimates are provided in Appendix A and B. The β estimates from this model now reveal whether individuals who have recently left teaching (and those who have recently joined) have better or worse mental health outcomes than their peers who have worked as teachers throughout this period.

Outcomes measured within the 2016 mental health questionnaire

We follow a similar approach in our analysis of the 2016 mental health questionnaire data; information collected during the initial assessment centre act as a rich set of controls, while responses to the 2016 mental health questionnaire are the outcome measures. Our variable of interest is based upon the information provided in the occupational history questionnaire and is defined as follows:

1. Always teachers (reference group). Individuals who were employed as teachers at both baseline (initial assessment centre) and at follow-up (2016 mental health questionnaire). N= 1,715.
2. Leavers. Those who were employed as teachers at baseline, but employed in another job at follow-up. N= 360[8].
3. Joiners. Those who not teachers at baseline, but were employed as a teacher at follow-up. N= 368.

[7] The amount of missing data is small for most covariates. The small number of covariates with large amounts of missing data are due to those questions only being included in assessment centres taken at later dates (e.g. questions about happiness were only included in later iterations of the Biobank assessment centre questionnaire) and not because of selective non-response. see Appendix E for further details on missing data by covariate.
[8] Around half of those who left teaching entered another job in education (e.g. become an inspector, teaching assistant), while the other half were working outside of education.

For Peer Review Only

Within this part of the analysis, the sample is restricted to those who were still employed and below retirement age (65) when they completed the mental health questionnaire. After making these restrictions, the average age of this analysis sample is 56. A comparison of the background characteristics of these groups to all teachers who participated in the initial Biobank assessment centre can be found in Appendix D. Imputation is again used to account for missing covariate data, while the substantive regression models are specified as outlined in the sub-sections above.

Effect sizes

Results are presented in terms of effect sizes; Cohen's d for continuous outcomes and odds-ratios for binary or ordinal outcomes. With respect to Cohen's d, we interpret values below 0.1 as evidence of essentially no effect. This is based upon two observation. First, meta-analyses have reported much larger effect sizes with respect to the impact of mental health interventions delivered in the workplace. For instance, Carolan, Harris, & Cavanagh (2017) reported an effect size of 0.37 for the impact of occupational mental health interventions upon psychological wellbeing. Second, we argue that an effect size of 0.1 would be substantively very small in this context. For instance, say that a person who left teaching scored an effect size 0.1 lower on the depression scale than those who remained in the profession. This would mean that there is only around a 52.8% chance that a person picked at random from those who have continued to be teachers will have a higher score on the depression scale than a person picked at random from the group who quit teacher for another job. In other words, the probability of suffering depression amongst members of these two groups would be little more than equal. We therefore consider any effect size below 0.1 as trivially small.

4. Results

Left teaching before the initial 2006-2010 assessment centre

The results focusing upon measures collected at the initial assessment centre can be found in Table 2. Starting with anxiety/depression, there is some limited evidence that former teachers have better outcomes on these measures than current teachers. Those individuals who left teaching within the last five years did score slightly lower on the self-reported depression

scale than their peers who were still working as teachers (effect size difference of 0.10), though they were no less likely to report taking prescription medicines for common mental health problems (odds ratio = 0.95). Those who had left teaching within the last five years also scored slightly lower, on average, on the neuroticism scale (effect size 0.13). These differences are, however, quite modest in terms of magnitude. Differences are similar (or slightly smaller) when comparing current teachers to those who left the profession more than five years ago. Evidence of a link between teaching and these outcomes is hence mixed.

<< Table 2 >>

The next set of estimates turns to the issue of sleep. There is little evidence of a difference between current and former teachers in terms of the amount of sleep they get over a typical 24-hour period; differences when expressed as an effect size are all below 0.10. Similarly, individuals who left teaching were slightly less likely to say that they had trouble falling asleep (or woke up during the night) than the reference group (current teachers). However, the effect size is again small (odds ratio ≈ 0.90). Any benefits from quitting teaching for one's quality and quantity of sleep are hence likely to be small (if at all).

The third set of outcomes presented in Table 2 refer to the consumption of alcohol. Former teachers are found to drink slightly more regularly than current teachers, though the difference is again relatively modest in magnitude (the estimated odds ratio is, at most, around 1.2). Furthermore, no difference is found between groups in terms of number of units of alcohol consumed each week. This suggests that former teachers drink roughly the same amount as individuals who have remained in teaching,

Finally, the last set of estimates presented in Table 2 refer to self-reported happiness with different aspects of life. A similar pattern again emerges. The estimated odds ratios mostly hover around one, fluctuating between 0.9 and 1.1, with no clear or obvious pattern. This holds true across most of the five specific areas of life teachers were asked about (e.g. finances, friends, family and health) as well as happiness in life overall. The one exception, where there is a sizeable difference, is happiness at work. Those who quit teaching recently – within the last five years – are happier in their jobs than those who have remained in the profession (odds-ratio 1.42). Yet our overall interpretation of the results presented in Table 2

For Peer Review Only

is that individuals who choose to leave teaching are not generally happier in their life than those continue working in this career.

Joined / left teaching between the initial assessment centre and follow-up assessment centre

Table 3 replicates the analysis presented in the sub-section above, but is now based upon the longitudinal Biobank data, focusing upon those who attended the assessment centre twice. Recall that this allows us to measure change in outcomes over time, and how these relate to whether individuals moved out, into or remained in the teaching profession. There are perhaps two key points of note (over and above our discussion of the results presented in Table 2).

<< Table 3 >>

First, there is some suggestion that that individuals who leave teaching report slightly lower levels of anxiety and depression than those who have remained in the teaching profession. Former teachers score 0.07 standard deviations lower on the self-reported depression scale, while also being somewhat less likely to report taking prescription medicines for common mental health problems (odds ratio 0.45) or reported suffering from depression/anxiety as a medical condition (odds ratio 0.78). Interestingly, those who entered the teaching profession since the baseline assessment centre scored slightly higher on the self-reported depression scale than individuals who were working as teachers at both time points (effect size 0.13) and those who had left teaching for another career (effect size 0.20). Nevertheless, the key message from Table 3 is that leaving teaching to pursue another career may only bring small benefits for one's mental health (if any at all).

Second, the final set of estimates in Table 3 (capturing self-reported happiness) potentially help strengthen the evidence that those individuals who leave teaching have higher levels of job satisfaction than those who continue to work as a teacher as their career. In particular, those who left teaching for other employment reported higher levels of satisfaction with their work than those who stayed in teaching (odds ratio 1.41). Yet this result is very much specific to the work domain; there is no evidence that those who quit teaching were happier with their health, friendships, family, health or, indeed, with life in general.

For Peer Review Only

The final point to note from Table 3 is that the results with respect to sleep and alcohol consumption are largely the same as Table 2. In other words, there is little evidence that working as a teacher has an impact upon difficulties with sleeping and alcohol consumption.

Joined / left teaching between the initial assessment centre and the mental health follow-up

Finally, Table 4 presents results from our second longitudinal analysis, where the initial 2006-2010 assessment centre is the baseline, while the 2016 mental health questionnaire captures the outcomes.

<< Table 4 >>

Table 4 again suggests that, if there are any differences in depression or anxiety between current and former teachers, they are trivially small. Differences in terms of effect sizes are consistently below 0.1 on the anxiety and depression scales (e.g. there is just a 0.06 standard deviation difference between current and former teachers on the self-reported depression scale). Similarly, there is little difference between current and former teachers in terms of their happiness (either overall or with their health), alcohol consumption and whether they believe that their life is meaningful.

Additional analyses in appendices

A series of additional analyses are presented within the appendices, investigating the sensitivity of our results to the different methods used. In summary:

- Appendix A investigates how results change when using a different set of controls. In this we continue to find some evidence that those individuals who left teaching are somewhat happier in their work. Otherwise, the results remained mixed, with no clear evidence that those who left teaching have clearly better mental health and wellbeing outcomes than other groups.
- Appendix B uses Inverse Probability Weighting (IPW) to adjust for covariates, rather than regression modelling. The clearest result is again that those individuals who leave teaching are somewhat happier with their work, though differences for the other outcomes tend to be small and often differ across the different approaches.

For Peer Review Only

- Appendix C alters the regression modelling approach used the analysis of the anxiety, depression and Neuroticism outcome scales, accounting for the fact that many people are clustered at a single score (zero). There continues to be little evidence of sizeable differences in these outcomes between individuals who decide to join, leave and remain in the teaching profession.
- Appendix F provides alternative estimates for our models using the follow-up assessment centre and mental health questionnaire data. In particular, all prior outcome measures are removed from the model, to investigate the extent that these factors may confound the results. Interestingly, the change in the coefficients between model specifications is relatively small. This may either suggest that (a) the selection mechanism for leaving teaching for another job is relatively weak (or at least not driven by concerns about wellbeing / mental health) or (b) that selection into/out of teaching is being driven by factors that the Biobank dataset does not measure.

Summary

In summary, the evidence for whether leaving teaching leads to lower levels of depression and anxiety is mixed. Although small improvements in these outcomes were observed within a subset of our analyses, at other times no impact was found. Our conclusion is therefore that any reduction in anxiety / depression from leaving teaching is likely to be (on average) very small, at best.

Consistent with Bamford and Worth (2017), there is some suggestion that those who decide to quit teaching end up being somewhat happier in their work than those who choose to remain. Yet, critically, this does not seem to translate into greater levels of happiness in other areas of life, including satisfaction with health or happiness overall. Consequently, the benefits of leaving teaching for one's happiness seems to be relatively minor, and concentrated in satisfaction with work. Finally, there is little evidence that the decision to leave teaching has any meaningful impact upon alcohol consumption (a mechanism many adults use to cope with stress) or quality of sleep (a marker of anxiety).

This leads us to reach an overall conclusion that leaving teaching for another job is unlikely to bring significant benefits to well-being or mental health.

For Peer Review Only

5. Conclusions

It is widely thought that teaching is a demanding job, due to the long working hours, the stress imposed by the accountability system and the challenge of having to manage often disruptive classrooms (De Carlo et al., 2019). Across the UK, and England in particular, this has led to many teachers considering a change of career (Perryman & Calvert, 2019). Indeed, in a recent international study, half of secondary teachers in England said that they wondered whether they should have chosen another profession – higher than in almost any other country across the world (*author cite*). Critically, in a recent survey of teachers in England who quit the profession, half said that their job was making them ill (Perryman & Calvert, 2019) and a high number of teachers interviewed in the Teacher Wellbeing Index (2018) experienced anxiety, depression or acute stress (Education Support Partnership, 2018). But is life really that much better for those individuals who decide to quit teaching to pursue a different job? This paper has presented important new evidence on this issue, with a particular focus upon how leaving teaching is linked to well-being and mental health.

Looking across an array of outcomes, and using several different empirical approaches, we have found little evidence of a link between leaving teaching, lower prevalence of mental health problems and higher levels of general well-being. Throughout our analysis, effect sizes have been small and often changed in both magnitude and direction depending upon the empirical approach taken. The one exception is happiness with work (job satisfaction), where we find a fairly consistent improvement for those who have recently left the teaching profession. Two observations are important in interpreting this finding. First, this is consistent with empirical research a range of occupations, which finds that job satisfaction tends to fall in the period prior to an individual quitting a job, before rising during the early stages of their new employment (Chadi & Hetschko, 2018; Gielen, 2013; Longhi et al., 2019). Second, we do not observe reductions in job satisfaction for those joining the teaching profession. Taken together, this suggests that the increase in job satisfaction for those leaving teaching is not indicative of any particular problem with teaching. Rather, it is more likely to reflect a natural process by which those who are less suited to the job move into alternative occupations (Gielen, 2013).

There are, of course, limitations with this paper and the need for future research. Five important issues stand out. First, some of the occupational career data has been based upon respondents recalling and accurately reporting such information. Although our focus upon relatively recent timepoints should limit the impact this has upon our analysis, collection of prospective longitudinal data from a cohort of teachers (tracking their entry and exits from the profession) would represent an important advance in the literature. Second, although a number of potential confounding factors were controlled within our analysis, our estimates continue to rely upon a (untestable) selection-upon-observables assumption if they are to be interpreted as causal effects. Third, while we have considered some "harder" evidence of poor mental health (e.g. prescription of antidepressant medicines), many of the outcome measures we investigated are based upon self-completion questionnaires. Although this is standard within this literature (Lenderink & Zoer, 2012), further work using a wider array of outcome data (e.g. primary care records) would represent a further step forward in this field. Fourth, the number of current and former teachers included within our analysis is relatively modest (in the hundreds rather than the thousands). Larger samples in the future would likely yield more precise results. Finally, the Biobank data is a convenience sample focused upon middle-aged participants, which is not representative of the wider teacher population. In particular, all participants were aged between 40 and 65, meaning it is not possible to generalise our findings to younger age groups. Future work using nationally representative data would undoubtedly enhance the external validity of our findings.

Despite these limitations, we believe that this paper has helped to advance our knowledge of teacher well-being and their mental health. At a time when many teachers are thinking about leaving for another career, it is vital that they are fully informed about the likely consequences. For those teachers who are not satisfied with their work, changing jobs may lead to an increase in job satisfaction. However, our results suggest that quitting teaching for alternative employment is unlikely to lead to improvements in general well-being or mental health. With respect to the latter outcomes, teaching does not stand out relative to other occupations.

For Peer Review Only

Data availability

The data that support the findings of this study are available from the UK Biobank. Restrictions apply to the availability of these data, which were used under license for this study. Data are available from https://www.ukbiobank.ac.uk/with the permission of the UK Biobank.

Ethical guidelines

The paper was produced following the BERA ethical guidelines (https://www.bera.ac.uk/). The project was approved by the UCL Institute of Education ethics committee (REC 1158).

Conflicts of Interest

The authors have no conflict of interest.

Funding

The paper has been funded by the Nuffield Foundation. The Nuffield Foundation is an independent charitable trust with a mission to advance social well-being. It funds research that informs social policy, primarily in Education, Welfare, and Justice. It also funds student programmes that provide opportunities for young people to develop skills in quantitative and scientific methods. The Nuffield Foundation is the founder and co-funder of the Nuffield Council on Bioethics and the Ada Lovelace Institute. The Foundation has funded this project, but the views expressed are those of the authors and not necessarily the Foundation. Visit www.nuffieldfoundation.org.

For Peer Review Only

References

Appleton, A., & James, R. (2018). Adults (18-75) in the UK who drink alcohol for coping reasons. Drinkaware research report. Accessed 04/07/2019 from https://www.drinkaware.co.uk/media/1852/drinking-to-cope-jan-2018.pdf

Bamford, S., & Worth, J. (2017). Is the grass greener beyond teaching? Slough: NFER. Accessed 04/07/2019 from https://www.nfer.ac.uk/teacher-retention-and-turnover-research-research-update-3-is-the-grass-greener-beyond-teaching

Berryhill, J., Linney, J., & Fromewick, J. (2009). The effects of education accountability on teachers: are policies too stress provoking for their own good? *International Journal of Education Policy & Leadership, 4*(5), 1-14.

Borman, G. D., & Dowling, N. M. (2008). Teacher attrition and retention: A meta-analytic and narrative review of the research. *Review of Educational Research, 78*(3), 367-409.

Culver, S., Wolfle, L., & Cross, L.H. (1990). Testing a model of teacher satisfaction for blacks and whites, *American Educational Research Journal, 27*(2), 323-349.

Carolan, S., Harris, P., & Cavanagh, K. (2017). Improving employee well-being and effectiveness: systematic review and meta-analysis of web-based psychological interventions delivered in the workplace. *Journal of Medical Internet Research, 19*(7).

Chadi, A., & Hetschko, C. (2018). The magic of the new: How job changes affect job satisfaction. *Journal of Economics & Management Strategy*, *27*(1), 23-39.

Clark, A. (2010). Work, jobs, and well-being across the millennium. *International differences in well-being*, 436-468.

Creek, J., & Hughes, A. (2008). Occupation and health: a review of selected literature. *British Journal of Occupational Therapy*, *71*(11), 456-468.

De Carlo, A., Girardi, D., Falco, A., Dal Corso, L., Di Sipio, A. (2019). When does work interfere with teachers' private life? An application of the job demands-resources model. *Frontiers in Psychology, 10*(1).

De Matteis, S., Jarvis, D., Young, H., Young, A., Allen, N., Potts, J., Darnton, A., Rushton, L., & Cullinan, P. (2017). Occupational self-coding and automatic recording (OSCAR): a novel web-based tool to collect and code lifetime job histories in large population-based studies. *Scandinavian Journal of Work, Environment and Health, 43*(2), 181-186.

den Brok, P., Wubbels, T., & Van Tartwijk, J. (2017). Exploring beginning teachers' attrition in the Netherlands. *Teachers and Teaching*, 23(8), 881-895.

Dolton, P., & Klaauw, W. V. D. (1995). Leaving teaching in the UK: A duration analysis. *The Economic Journal*, *105*(429), 431-444.

Education Support Partnership, 2018. *Teacher Wellbeing Index 2018*. Accessed 04/11/2019 from https://www.educationsupportpartnership.org.uk/sites/default/files/teacher_wellbeing_index_2018.pdf

Freeman, D., Sheaves, B., Goodwin, G. M., Yu, L. M., Nickless, A., Harrison, P. J., ... & Hinds, C. (2017). The effects of improving sleep on mental health (OASIS): a randomised controlled trial with mediation analysis. *The Lancet Psychiatry, 4*(10), 749-758.

Fujishiro, K., Farley, A., Kellemen, M., & Swoboda, C. (2017). Exploring associations between state education initiatives and teachers' sleep: A social-ecological approach. *Social Science and Medicine, 191*, 151-159.

Gallagher, M., Muldoon, O. T., & Pettigrew, J. (2015). An integrative review of social and occupational factors influencing health and wellbeing. *Frontiers in Psychology*, *6*.

Gielen, A. C. (2013). Repeated job quits: stepping stones or learning about quality?. *IZA Journal of European Labor Studies*, *2*(7).

Howes, L. M., & Goodman-Delahunty, J. (2015). Teachers' career decisions: Perspectives on choosing teaching careers, and on staying or leaving. *Issues in Educational Research, 25*(1), 18.

Johnson, S., Cooper, C., Cartwright, S., Doland, I., Taylor, P., & Millet, C. (2005). The experience of work-related stress across occupations. *Journal of Managerial Psychology,* 20(2), 178-187.

Kidger, J., Brockman, R., Tilling, K., Campbell, R., Ford, T., Araya, R., King, M., & Gunnell, D. (2016). Teachers' wellbeing and depressive symptoms, and associated risk factors: A large cross-sectional study in English secondary schools. *Journal of Affective Disorders, 192*, 76-82.

Longhi, S., Nandi, A., Bryan, M., Connolly, S., & Gedikli, C. (2019). *Do all job changes increase wellbeing?* What Works Wellbeing.

Mérida-López, S., & Extremera, N. (2017). Emotional intelligence and teacher burnout: A systematic review. *International Journal of Educational Research, 85*, 121-130.

Lenderink, A., & Zoer, I. (2012). *Review on the validity and reliability of self-reported work-related illness*. Health and Safety Executive Research Report RR903. Accessed 04/07/2019 from http://www.hse.gov.uk/research/rrpdf/rr903.pdf

Mercer, D., & Evans, B. (1991). Professional myopia: Job satisfaction and the management of teachers. *School Organization, 11*(3), 291-301.

Mérida-López, S. & Extremera, N. (2017). Emotional intelligence and teacher burnout: A systematic review. *International Journal of Educational Research, 85*, 121-130.

McCormack, A., Edmondson-Jones, M., Fortnum, H., Dawes, P., Middleton, H., Munro, K., & Moore, D. (2015). Investigating the association between tinnitus severity and symptoms of depression and anxiety, while controlling for neuroticism, in a large middle-aged UK population. *International Journal of Audiology, 54*(9): 599-604.

McIntyre, J. (2010). Why they sat still: the ideas and values of long-serving teachers in challenging inner-city schools in England. *Teachers and Teaching: theory and practice, 16*(5), 595-614.

Perryman, J., & Calvert, G. (2019). What motivates people to teach, and why do they leave? Accountability, performativity and teacher retention. *British Journal of Educational Studies*, *68*(1), 3-23.

Peters, E., John, A., Bowen, R., Baetz, M., & Balbuena, L. (2018). Neuroticism and suicide in a general population cohort: results from the UK Biobank Project. *BJPsych Open*, *4*(2), 62–68.

Ravesteijn, B., van Kippersluis, H., & van Doorslaer, E. (2013). The Contribution of Occupation to Health Inequality. *Research on Economic Inequality*, *21*, 1-20.

Ravesteijn, B., Kippersluis, H. V., & Doorslaer, E. V. (2018). The wear and tear on health: What is the role of occupation?. *Health Economics*, *27*(2), e69-e86.

Scheuch, K., Haufe, E., & Seibt, R. (2015). Teachers' health. *Deutsches Ärzteblatt International*, *112*(20), 347.

Smithers, A. (1990). *Teacher Loss*. Manchester: University of Manchester.

Stansfeld, S., Rasul, F, Head, J., & Singleton, N. (2011). Occupation and mental health in a national UK survey. *Social Psychiatry and Psychiatric Epidemiology 46*(2), 101–110.

Tapper, J. (2018, 13 May). Burned out: why are so many teachers quitting or off sick with stress? *The Guardian.*

Taylor, M., Rode, L., Bjørngaard, J., Taylor, A., Bojesen, S., Åsvold, B., Gabrielsen, M., Lewis, G., Nordestgaard, B., Romundstad, P., Hickman, M., & Munafò, M. (2018). Is smoking heaviness causally associated with alcohol use? A Mendelian randomization study in four European cohorts, *International Journal of Epidemiology, 47*(4), 1098–1105.

Tye, B., & O'brien, L. (2002). Why are experienced teachers leaving the profession?. *Phi Delta Kappan, 84*(1), 24-32.

For Peer Review Only

Table 1. Characteristics of the Biobank sample compared to population estimates

	Biobank	Population estimate	
		Estimate	Source
Average age	53	51	LFS
% male	27%	27%	LFS
% children in household	53%	44%	LFS
% Partner in household	76%	74%	LFS
% hold a degree	84%	75%	LFS
Average age left school	20	21	LFS
Born outside UK	7%	7%	LFS
Homeowner	95%	94%	LFS
Smoker	6%	4%	APS 2010
% poor general health	1%	1%	NCDS 2008
% fair general health	14%	7%	NCDS 2008
% good general health	63%	65%	NCDS 2008
% excellent general health	22%	26%	NCDS 2008
Total teachers	**16,622**		

Notes: LFS = Labour Force Survey data for 40-65-year-old teachers from January-March sweeps 2007,2008 and 2009. APS = Annual Population Survey data from 2010 for 40-65-year-old teachers. NCDS = National Child Development Survey from 2008 (when respondents were 50-years-old). The NCDS data for 'good' and 'very good' general health has been combined. APS 2010 data based upon information across all education and teaching professionals.

For Peer Review Only

Table 2. The association between leaving/remaining in the teaching profession and mental health outcomes measured in the 2006-2010 Biobank assessment centre.

	Left teaching last 5 years		Left teaching 6 - 10 years ago		Left teaching > 10 years ago	
	Beta	T-Stat	Beta	T-Stat	Beta	T-Stat
Depression / anxiety						
Self-reported depression (ES)	-0.10	-3.71	-0.08	-2.20	0.00	-0.21
Prescribed medicines (OR)	0.95	-0.31	0.85	-0.81	0.82	-1.57
Self-reported medical condition (OR)	1.07	0.54	0.98	-0.14	1.08	0.77
Neuroticism (ES)	-0.13	-4.19	-0.11	-2.52	-0.08	-3.65
Sleep						
Hours sleep per night (ES)	0.08	3.54	0.01	0.17	0.04	2.22
Trouble falling sleep (OR)	0.90	-1.97	0.87	-1.90	0.89	-2.56
Alcohol						
Frequent drinking (OR)	1.19	3.25	1.16	2.10	1.19	4.26
Number alcohol units per week	0.06	0.27	-0.22	-0.70	0.09	0.50
Happiness						
Happy with work (OR)	1.42	2.91	1.13	0.98	1.09	1.32
Happy with finances (OR)	0.95	-0.50	0.92	-0.63	0.97	-0.35
Happy with friends (OR)	0.85	-1.75	1.00	-0.01	0.91	-1.25
Happy with family (OR)	0.86	-1.60	0.88	-1.01	0.94	-0.79
Happy with health (OR)	0.98	-0.24	0.91	-0.42	1.22	3.67
Overall happiness (OR)	1.01	0.11	0.94	-0.48	1.05	0.66

Notes: Those individuals who were currently teachers at the time of the assessment centre are the reference group. 'ES' refers to estimated effect size for continuous outcomes and 'OR' to the odds ratio for binary/categorical outcomes. Estimates based upon regression models controlling for age, month visited the assessment centre, gender, fluid intelligence score, Townsend socio-economic status index, family history of depression, whether a major negative life event occurred within last two years, whether has a partner living in same household, whether children live in the same household, household size, household income, age left education, whether hold a degree, whether born outside of the UK and whether had

For Peer Review Only

ever suffered depression before working as a teacher. Multiple imputation has been used to account for missing covariate data.

Table 3. The association between leaving/remaining in the teaching profession with outcomes measured in the follow-up Biobank assessment centre

	Left teaching		Joined teaching	
	Beta	**T-Stat**	**Beta**	**T-Stat**
Depression / anxiety				
Self-reported depression (ES)	-0.07	-1.01	0.13	1.78
Prescribed medicines (OR)	0.45	-1.47	0.82	-0.41
Self-reported medical condition (OR)	0.78	-0.75	1.20	0.61
Sleep				
Hours sleep per night (ES)	0.04	0.63	-0.04	-0.62
Trouble falling sleep (OR)	0.90	-0.60	0.73	-1.82
Alcohol				
Frequent drinking (OR)	1.00	-0.02	0.77	-1.59
Number units of alcohol per week	-1.60	-1.79	-2.30	-2.57
Happiness				
Happy with work (OR)	1.41	2.09	1.03	0.20
Happy with finances (OR)	0.84	-1.07	0.72	-1.99
Happy with friends (OR)	0.80	-1.31	0.86	-0.88
Happy with family (OR)	1.18	1.03	1.03	0.15
Happy with health (OR)	0.89	-0.70	0.83	-1.11
Overall happiness (OR)	1.02	0.12	0.84	-0.97

Notes: Those individuals who were teachers at both baseline (initial assessment centre) and follow-up (follow-up assessment centre) are the reference group. 'ES' refers to estimated effect size for continuous outcomes and 'OR' to the risk ratio for binary/categorical outcomes. Estimates based upon regression models controlling for age, month visited the assessment centre, gender, whether has a partner living in same household, whether hold a degree, fluid intelligence score, Townsend socio-economic status index, age left education, family history of depression, self-reported depression at baseline, medical conditions and prescriptions

For Peer Review Only

reported at baseline, general health reported at baseline, happiness with work/friends/family/finances/health at baseline, alcohol intake at baseline, sleep quality and quantity at baseline, whether suffered prolonged spell of depression before baseline. Multiple imputation has been used to account for missing covariate data.

Table 4. The association between leaving/remaining in the teaching profession and outcomes measured in the 2016 mental health questionnaire.

	Left teaching since baseline		Joined teaching since baseline	
	Beta	**T-Stat**	**Beta**	**T-Stat**
Depression / anxiety				
Self-reported depression (ES)	0.06	0.96	0.05	0.80
Self-reported anxiety (ES)	0.09	1.28	0.06	0.85
Prolonged depress spell since baseline (OR)	1.11	0.51	0.89	-0.61
Self-harm				
Considered self-harm in last 12 months (OR)	1.28	0.72	1.82	1.82
Alcohol consumption				
Frequently drink >6 drinks (OR)	0.89	-0.82	0.73	-2.15
Other				
Happy with health (OR)	0.82	-1.30	1.06	0.35
Feel life is meaningful (OR)	1.03	0.17	0.79	-1.43
Overall happiness (OR)	0.82	-0.95	0.78	-1.68

Notes: Those individuals who were teachers at both baseline (assessment centre) and follow-up (mental health questionnaire) are the reference group. 'ES' refers to estimated effect size for continuous outcomes and 'OR' to the odds ratio for binary/categorical outcomes. Estimates based upon regression models controlling for age, month visited the assessment centre, gender, whether has a partner living in same household, whether hold a degree, fluid intelligence score, Townsend socio-economic status index, age left education, family history of depression, self-reported depression at baseline, medical conditions and prescriptions reported at baseline, general health reported at baseline, happiness with work/friends/family/finances/health at baseline, alcohol intake at baseline, sleep quality and quantity at baseline, whether suffered prolonged spell of depression before baseline. Multiple imputation has been used to account for missing covariate data.

For Peer Review Only

For Peer Review Only

Appendix A. Alternative regression model estimates including different sets of controls

Table A1. The association between leaving/remaining in the teaching profession and mental health outcomes measured in the 2006-2010 Biobank assessment centre. Alternative set of controls

	Left teaching last 5 years		Left teaching 6 - 10 years ago		Left teaching > 10 years ago	
	Beta	T-Stat	Beta	T-Stat	Beta	T-Stat
Depression / anxiety						
Self-reported depression (ES)	-0.18	-3.79	-0.14	-2.05	-0.05	-1.23
Prescribed medicines (OR)	0.89	-0.81	0.78	-1.25	0.77	-2.12
Self-reported medical condition (OR)	1.05	0.41	0.98	-0.13	1.02	0.27
Neuroticism (ES)	-0.12	-4.29	-0.09	-2.34	-0.10	-4.67
Sleep						
Hours sleep per night (ES)	0.09	3.94	0.01	0.39	0.04	2.31
Trouble falling sleep (OR)	0.90	-1.94	0.85	-2.33	0.88	-3.10
Alcohol						
Frequent drinking (OR)	1.15	2.85	1.17	2.37	1.26	5.82
Units of alcohol per week	-0.02	-0.08	-0.13	-0.44	0.22	1.27
Happiness						
Happy with work (OR)	1.49	4.94	1.17	1.25	1.09	1.02
Happy with finances (OR)	0.98	-0.22	0.90	-0.90	1.11	1.50
Happy with friends (OR)	0.85	-1.85	0.96	-0.33	0.92	-1.24
Happy with family (OR)	0.85	-1.84	0.86	-1.26	0.94	-0.95
Happy with health (OR)	1.04	0.48	0.95	-0.45	1.23	2.69
Overall happiness (OR)	1.06	0.74	0.96	-0.32	1.11	1.15

Notes: Estimates based upon regression models controlling for age, month completed assessment centre, gender, mother/father/sibling ever had depression, immigrant status, partner in the household, children in the household, age finished education, whether respondent holds a degree and whether first instance of depression occurred before they

became a teacher. Multiple imputation has been used to account for missing covariate data. See notes to Table 2 for further details.

Table A2. The association between leaving/remaining in the teaching profession with outcomes measured in the follow-up Biobank assessment centre. Alternative set of controls.

	Left teaching		Joined teaching	
	Beta	**T-Stat**	**Beta**	**T-Stat**
Depression / anxiety				
Self-reported depression (ES)	-0.05	-0.76	0.11	1.67
Prescribed medicines (OR)	0.58	-1.16	0.75	-0.69
Self-reported medical condition (OR)	0.75	-0.96	1.01	0.03
Sleep				
Hours sleep per night (ES)	0.01	0.19	-0.06	-0.95
Trouble falling sleep (OR)	0.90	-0.62	0.80	-1.47
Alcohol				
Frequent drinking (OR)	1.00	-0.01	0.70	-2.38
Number of units per week	-1.13	-1.30	-1.95	-2.41
Happiness				
Happy with work (OR)	1.26	1.47	1.03	0.22
Happy with finances (OR)	0.77	-1.66	0.81	-1.46
Happy with friends (OR)	0.82	-1.21	0.97	-0.23
Happy with family (OR)	1.13	0.81	1.03	0.18
Happy with health (OR)	0.78	-1.57	0.81	-1.41
Overall happiness (OR)	0.96	-0.25	0.98	-0.15

Notes: Estimates based upon regression models controlling for age, month completed assessment centre, gender, mother/father ever had depression, partner in the household, age finished education and the following measures recorded at the baseline assessment centre: depression scale, prescription of anti-depressants, self-reported medical condition of depression/anxiety/insomnia, happiness (overall, with work and with health), general health, difficulty sleeping and alcohol consumption. Multiple imputation has been used to account for missing covariate data. See notes to Table 2 for further details.

For Peer Review Only

Table A3. The association between leaving/remaining in the teaching profession and outcomes measured in the 2016 mental health questionnaire. Alternative set of controls.

	Left teaching since baseline		Joined teaching since baseline	
	Beta	T-Stat	Beta	T-Stat
Depression / anxiety				
Self-reported depression (ES)	0.05	0.72	0.02	0.37
Self-reported anxiety (ES)	0.07	1.09	0.04	0.68
Prolonged depress spell since baseline (OR)	1.09	0.44	0.87	-0.82
Self-harm				
Considered self-harm in last 12 months (OR)	1.24	0.63	1.66	1.51
Alcohol consumption				
Frequently drink >6 drinks (OR)	0.96	-0.31	0.84	-1.44
Other				
Happy with health (OR)	0.81	-1.63	1.06	0.38
Feel life is meaningful (OR)	0.99	-0.05	0.86	-1.05
Overall happiness (OR)	0.80	-1.22	0.84	-1.17

Notes: Estimates based upon regression models controlling for age, gender, age finished education, partner in household, and the following measures recorded at the baseline assessment centre: depression scale, happiness (overall and with different areas), general health, difficulty sleeping and alcohol consumption. Imputation has been used to account for missing covariate data. See notes to Table 4 for further details.

For Peer Review Only

Appendix B. Inverse probability weighting estimates

Table B1. The association between leaving/remaining in the teaching profession and mental health outcomes measured in the initial Biobank assessment centre. IPW estimates converted to effect sizes for all variables.

	Left teaching last 5 years		Left teaching 6 - 10 years ago		Left teaching > 10 years ago	
	Beta	T-Stat	Beta	T-Stat	Beta	T-Stat
Depression / anxiety						
Self-reported depression	-0.10	-3.65	-0.21	-1.69	-0.01	-0.22
Prescribed medicines	-0.01	-0.20	0.00	-0.76	0.00	-1.51
Self-reported medical condition	0.03	0.87	0.00	0.11	0.00	0.96
Neuroticism	-0.13	-4.11	-0.10	-2.55	-0.08	-3.42
Sleep						
Hours sleep per night	0.10	3.40	0.01	0.25	0.03	1.77
Trouble falling sleep	-0.06	-1.96	-0.04	-1.93	-0.03	-2.44
Alcohol						
Frequent drinking	0.09	3.21	0.20	2.60	0.18	3.87
Units alcohol per week	0.01	0.35	-1.88	-0.70	0.74	0.44
Happiness						
Happy with work	0.19	4.12	0.09	1.94	0.02	0.90
Happy with finances	-0.01	-0.34	-0.01	-0.32	0.00	0.18
Happy with friends	-0.08	-1.74	0.00	-0.04	-0.02	-0.96
Happy with family	-0.07	-1.48	-0.05	-0.92	-0.01	-0.47
Happy with health	0.00	0.06	-0.02	-0.45	0.07	2.94
Overall happiness	-0.02	-0.34	-0.01	-0.36	0.01	0.39

Notes: Those individuals who were currently teachers at the time of the assessment centre are the reference group. Figures have been converted into an approximate effect size by dividing the estimate by the sample standard deviation for the variable. Estimates based upon IPW, including the following variables in the matching model: age, month of assessment centre, gender, fluid intelligence, socio-economic status, family history of depression, major negative life event occurred in last two years, whether born in UK, household income, household structure, age left school, whether hold a degree and whether experienced

For Peer Review Only

depression before becoming a teacher. Single imputation used to account for missing covariate data.

Table B2. The association between leaving/remaining in the teaching profession with outcomes measured in the follow-up Biobank assessment centre. IPW estimates converted to effect sizes for all variables.

	Left teaching		Joined teaching	
	Beta	**T-Stat**	**Beta**	**T-Stat**
Depression / anxiety				
Self-reported depression	-0.08	-1.08	0.14	1.80
Prescribed medicines	-0.11	-1.58	-0.01	-0.19
Self-reported medical condition	-0.06	-0.83	0.04	0.47
Sleep				
Hours sleep per night	0.03	0.32	-0.06	-0.73
Trouble falling sleep	-0.05	-0.59	-0.14	-1.87
Alcohol				
Frequent drinking	0.00	-0.04	-0.13	-2.29
Number units alcohol per week	-0.13	-1.67	-0.18	-2.44
Happiness				
Happy with work	0.20	2.27	0.02	0.28
Happy with finances	-0.10	-1.21	-0.14	-1.78
Happy with friends	-0.03	-0.40	-0.03	-0.34
Happy with family	0.11	1.38	-0.01	-0.10
Happy with health	-0.06	-0.75	-0.07	-0.95
Overall happiness	0.02	0.22	0.00	-0.02

Notes: Figures have been converted into an approximate effect size by dividing the estimate by the sample standard deviation for the variable. Estimates based upon IPW, including the following variables in the matching model: age, month of assessment centre, gender, whether partner is in the household, fluid intelligence, socio-economic status, family history of depression, age left school and baseline measures of depression, prescriptions for mental health issues, self-reported mental health issue, happiness, general health, quality and quantity of sleep, alcohol intake. Single imputation used to account for missing covariate data.

For Peer Review Only

Table B3. The association between leaving/remaining in the teaching profession and outcomes measured in the 2016 mental health questionnaire. IPW estimates converted to effect sizes for all variables.

	Left teaching since baseline		Joined teaching since baseline	
	Beta	T-Stat	Beta	T-Stat
Depression / anxiety				
Self-reported depression	0.03	0.42	-0.01	-0.08
Self-reported anxiety	0.06	0.88	0.05	0.50
Prolonged depress spell since baseline	0.02	0.43	-0.02	-0.48
Self-harm				
Considered self-harm in last 12 months	0.00	0.02	0.05	0.55
Alcohol consumption				
Frequently drink >6 drinks	-0.08	-1.24	-0.17	-1.96
Other				
Happy with health	-0.13	-1.77	0.00	-0.02
Feel life is meaningful	-0.01	-0.20	-0.19	-1.87
Overall happiness	-0.05	-0.82	-0.14	-1.55

Notes: Figures have been converted into an approximate effect size by dividing the estimate by the sample standard deviation for the variable. Estimates based upon IPW, including the following variables in the matching model: age, gender, whether partner is in the household, socio-economic status, family history of depression, whether hold a degree, age left school and baseline measures of depression, prescriptions for mental health issues, self-reported mental health issue, happiness, general health, sleep and alcohol intake. Single imputation used to account for missing covariate data.

For Peer Review Only

For Peer Review Only

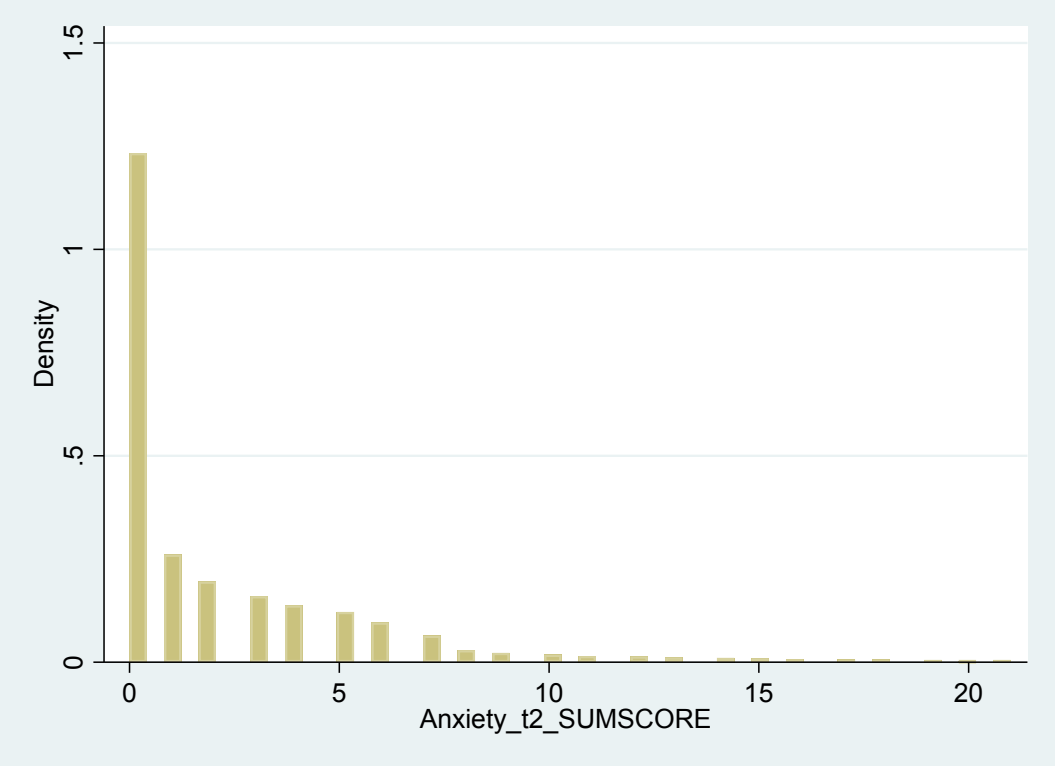
Appendix C. Zero-inflated models

Within the main body of the paper, we treat the following variables as continuous and hence use ordinary least squares regression:

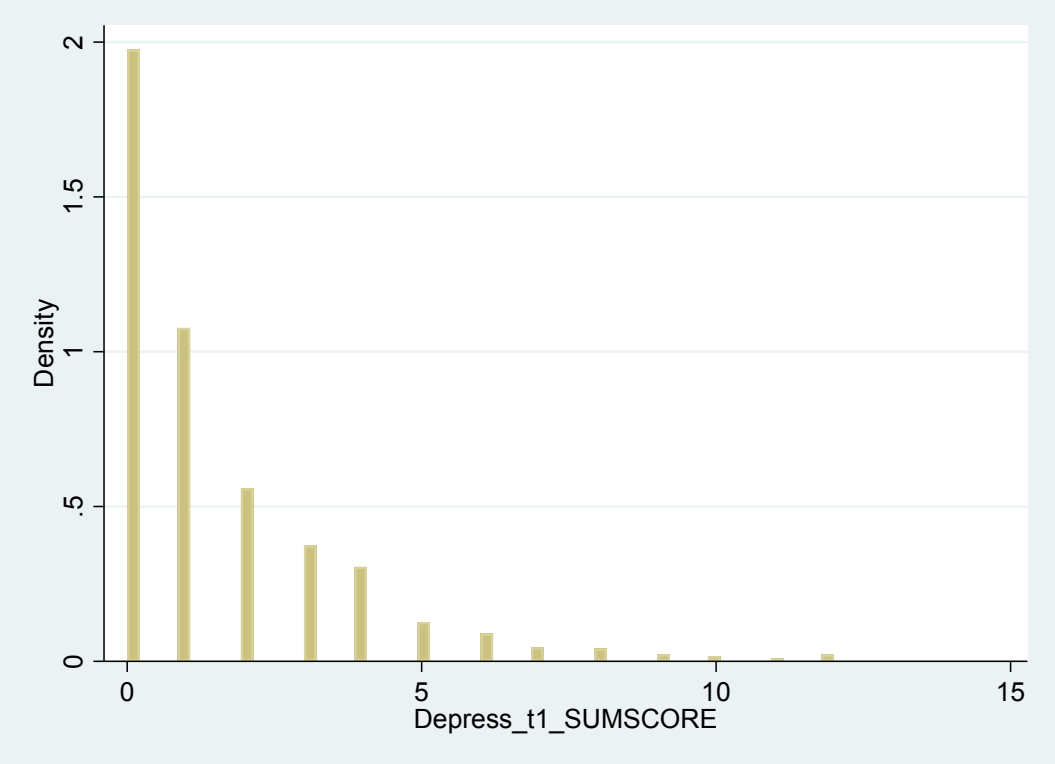
- Anxiety score
- Depression score
- Neuroticism score

The reason for using OLS regression is that allows for straightforward estimation and interpretation as an effect size. However, as illustrated by Appendix Figures C1, C2 and C3 below, these outcomes may actually be considered to be "zero-inflated" data (meaning there is a large cluster of observations at zero). Consequently, to test the robustness of our results for these outcomes, in this appendix we estimate Zero-Inflated Poisson (ZIP) models. ZIP models essentially divide the outcome into two parts. The first is a binary response (logit) model, which estimates the probability of whether the outcome is zero or not. The second part is a count (Poisson) model, which models the values greater than zero. We implement these models using the same predictors as discussed in the main body of the paper for both parts of the ZIP model. Results are presented in Appendix Tables C1 to C3. These present estimates as odds-ratios for the 'inflation' model (i.e. the logit model which models whether the probability of the outcome being zero or not) and incidence-rate-ratios for the 'outcome' model (i.e. the Poisson count regression estimating the score on the scale).

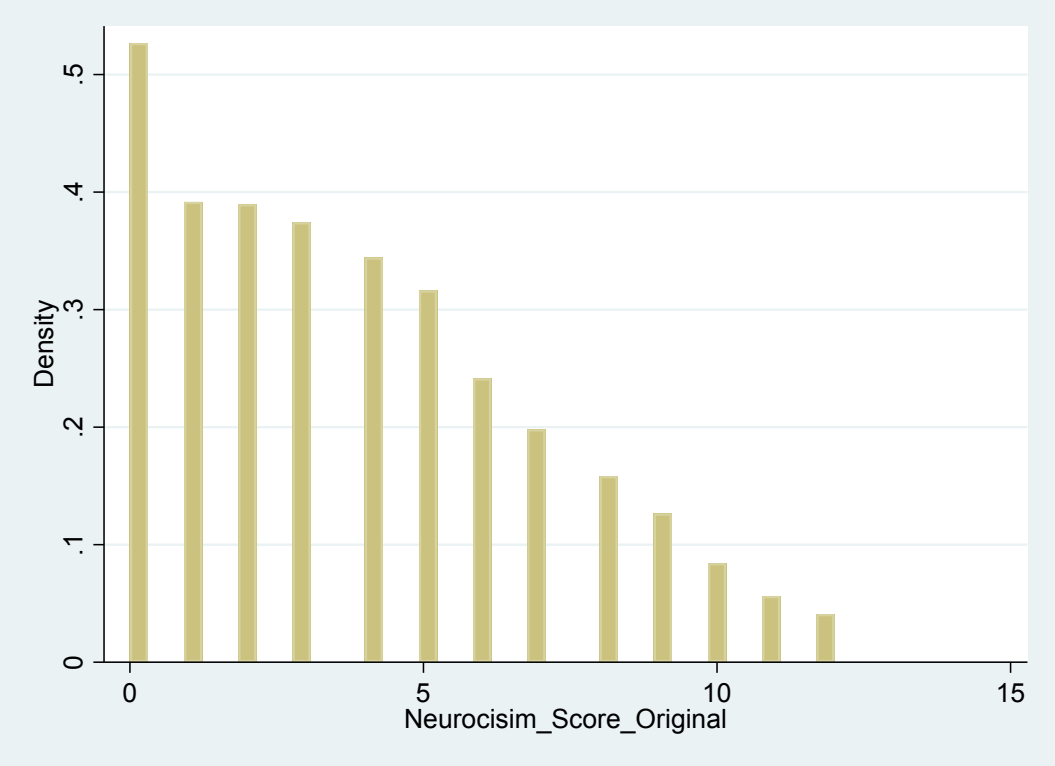
Appendix Figure C1. Distribution of the Anxiety sum score variable



Appendix Figure C2. Distribution of the Depression sum score variable



Appendix Figure C3. Distribution of the Neuroticism sum score variable



Appendix Table C1. Zero-Inflated Poisson estimates for selected outcome variables measures in the initial assessment centre

	Depression scale		Neuroticism scale	
	Effect	T-Stat	Effect	T-Stat
Outcome model (IRR)				
Left in last 5 years	0.90	3.43	0.94	3.92
Left 6-10 years ago	0.90	2.53	0.94	2.87
Left 11+ years ago	0.94	2.49	0.94	4.89
Inflation model (OR)				
Left in last 5 years	1.08	0.87	1.22	2.28
Left 6-10 years ago	0.90	0.83	1.02	0.14
Left 11+ years ago	0.85	2.19	1.06	0.84

Appendix Table C2. Zero-Inflated Poisson estimates for selected outcome variables measures in the follow-up assessment centre

	Depression scale	
	Effect	T-Stat
Outcome model (IRR)		
Left teaching since baseline	0.96	0.93
Joined teaching since baseline	1.04	0.94
Inflation model (OR)		
Left teaching since baseline	0.75	1.35
Joined teaching since baseline	0.62	2.38

For Peer Review Only

Appendix Table C3. Zero-Inflated Poisson estimates for selected outcome variables measures in the mental health questionnaire follow-up

	Depression scale		Anxiety scale	
	Effect	T-Stat	Effect	T-Stat
Outcome model (IRR)				
Left teaching since baseline	1.03	1.36	1.04	0.91
Joined teaching since baseline	1.02	1.08	0.94	1.47
Inflation model (OR)				
Left teaching since baseline	0.88	0.86	0.84	1.15
Joined teaching since baseline	0.74	1.87	0.69	2.46

For Peer Review Only

Appendix D. Characteristics of the assessment centre and mental health questionnaire longitudinal follow-up samples

(a) Assessment centre sample

	Original sample at assessment centre 1	Still teaching assessment centre 2	Left teaching by assessment centre 2	Joined teaching by assessment centre 2
Average age at first contact	53	51	50	51
% male	27%	29%	34%	36%
% children in household at first contact	53%	64%	66%	52%
% Partner in household at first contact	76%	77%	74%	76%
% hold a degree	84%	86%	77%	64%
Born outside UK	7%	6%	6%	5%
Homeowner at first contact	95%	98%	93%	96%
Smoker at first contact	6%	4%	4%	7%
General health at first contact				
% poor	1%	1%	1%	0%
% fair	14%	13%	11%	12%
% good	63%	62%	63%	62%
% excellent	22%	24%	25%	26%
Household income at first contact				
£18,000<	14%	11%	20%	17%
£18,000 - £30,999	38%	37%	36%	45%
£31,000 - £51,999	43%	47%	38%	36%
£52,000+	6%	6%	6%	4%
Total teachers	**16,622**	**925**	**167**	**176**

(b) Mental health questionnaire

	Original sample at assessment centre 1	Teacher at assessment centre and mental health questionnaire	Left teaching	Joined teaching
Average age at first contact	53	48	50	49
% male	27%	21%	31%	28%
% children in household at first contact	53%	71%	66%	69%
% Partner in household at first contact	76%	75%	73%	74%
% hold a degree	84%	92%	88%	78%
Born outside UK	7%	7%	10%	8%
Homeowner at first contact	95%	95%	93%	91%
Smoker at first contact	6%	5%	7%	5%
General health at first contact				
% poor	1%	1%	1%	1%
% fair	14%	12%	15%	15%
% good	63%	61%	58%	63%
% excellent	22%	26%	26%	21%
Household income at first contact				
£18,000<	14%	8%	16%	15%
£18,000 - £30,999	38%	35%	40%	36%
£31,000 - £51,999	43%	50%	37%	40%
£52,000+	6%	7%	6%	9%
Total teachers	**16,622**	**1,715**	**360**	**368**

Appendix E. The percentage of observations missing information in each covariate

	Assessment centre 1 analysis	Assessment centre 2 (longitudinal) analysis	Mental health questionnaire analysis
Age first contact	0%	0%	0%
Gender	0%	0%	0%
Fluid intelligence score	64%	67%	60%
Townsend index	0%	0%	0%
Maternal depression	1%	2%	1%
Paternal depression	3%	3%	3%
Sibling depression	3%	3%	3%
Relative die soon before first contact	0%	0%	0%
Had serious illness soon before fist contact	0%	0%	0%
Divorced soon before first contact	0%	0%	0%
Financial problems soon before first contact	0%	0%	0%
Born outside UK	0%	0%	0%
Partner in household at first contact	0%	0%	0%
Children in household at first contact	0%	0%	0%
Household size	0%	0%	0%
Household income	6%	5%	4%
Whether hold a degree	0%	0%	0%
Experienced depression before teaching	49%	29%	21%
Depression scale at 1st contact	-	3%	3%
Prescribed antidepressants at 1st contact	-	0%	0%
Depression as medical condition 1st contact	-	0%	0%
Happy at first contact	-	67%	60%

For Peer Review Only

Happy at work at first contact	-	66%	60%
General health at first contact	-	0%	0%
Happy with health at first contact	-	66%	60%
Happy with family at first contact	-	67%	60%
Happy with friends at first contact	-	67%	60%
Happy with finance at first contact	-	66%	60%
Hours sleep per night at first contact	-	0%	0%
Insomnia at first contact	-	0%	0%
Alcohol intake at first contact	-	0%	0%
Number of units at first contact	-	0%	0%
Full sample	**20768**	**1268**	**2443**

For Peer Review Only

Appendix F. Alternative estimates for the longitudinal analysis with different sets of controls

In this appendix we repeat our analysis of the follow-up assessment centre and mental health questionnaire data, but now using a different set of control variables. Specifically, we remove from the model the prior outcome measures that were measured in the initial Biobank assessment centre. The motivation behind this exploration is to investigate that the inclusion or exclusion of these controls changes the results. Appendix Table F1 lists the covariates across the two sets of models being compared. Results for the follow-up Biobank Assessment Centre (analogous to those presented in Table 3) can be found in Appendix Table F2. Those for the mental health questionnaire (analogous to those presented in Table 4) can be found in Appendix Table F3. Interestingly, the inclusion or exclusions of the prior outcome measures collected in the initial assessment centre do not seem to substantively alter the results or the key conclusions reached.

Appendix Table F1. Covariates included across the two model specifications

	Without additional controls	With additional controls
Age	Y	Y
Month of assessment centre visit	Y	Y
Gender	Y	Y
Partner in household	Y	Y
Hold a degree	Y	Y
Fluid intelligence score	Y	Y

Townsend socio-economic status index	Y	Y
Age left education	Y	Y
Family history of depression	Y	Y
Medical conditions and prescriptions reported at baseline	-	Y
General health reported at baseline	-	Y
Happiness with work at baseline	-	Y
Happiness with friends at baseline	-	Y
Happiness with family at baseline	-	Y
Happiness with finances at baseline	-	Y
Happiness with health at baseline	-	Y
Alcohol intake at baseline	-	Y
Sleep quality and quantity at baseline	-	Y
Suffered spell of depression before baseline	-	Y

For Peer Review Only

Appendix Table F2. Alternative estimates of the association between leaving/remaining in the teaching profession with outcomes measured in the follow-up Biobank assessment centre (analogous to Table 3)

	Left teaching		Joined teaching	
	Without additional controls	With additional controls	Without additional controls	With additional controls
Depression / anxiety				
Self-reported depression (ES)	-0.05	-0.07	0.10	0.13
Prescribed medicines (OR)	0.56	0.45	0.84	0.82
Self-reported medical condition (OR)	0.82	0.78	1.19	1.20
Sleep				
Hours sleep per night (ES)	0.01	0.04	-0.01	-0.04
Trouble falling sleep (OR)	0.94	0.90	0.69	0.73
Alcohol				
Frequent drinking (OR)	1.03	1.00	0.87	0.77
Number units of alcohol per week	-1.74	-1.60	-2.30	-2.30
Happiness				
Happy with work (OR)	1.40	1.41	1.11	1.03
Happy with finances (OR)	0.87	0.84	0.76	0.72
Happy with friends (OR)	0.80	0.80	0.91	0.86
Happy with family (OR)	1.17	1.18	1.10	1.03
Happy with health (OR)	0.95	0.89	1.02	0.83

Overall happiness (OR)	1.01	1.02	0.94	0.84

Appendix Table F3. Alternative estimates of the association between leaving/remaining in the teaching profession and outcomes measured in the 2016 mental health questionnaire (analogous to Table 4)

	Left teaching since baseline		Joined teaching since baseline	
	Without additional controls	With additional controls	Without additional controls	With additional controls
Depression / anxiety				
Self-reported depression (ES)	0.06	0.06	0.05	0.05
Self-reported anxiety (ES)	0.10	0.09	0.06	0.06
Prolonged depress spell since baseline (OR)	1.01	1.11	0.78	0.89
Self-harm				
Considered self-harm in last 12 months (OR)	1.22	1.28	1.62	1.82
Alcohol consumption				
Frequently drink >6 drinks (OR)	1.00	0.89	1.01	0.73
Other				
Happy with health (OR)	0.95	0.82	0.96	1.06
Feel life is meaningful (OR)	0.96	1.03	0.84	0.79
Overall happiness (OR)	0.8	0.82	0.84	0.78

For Peer Review Only

For Peer Review Only

I quit! Is there an association between leaving teaching and improvements in mental health?

John Jerrim
Sam Sims
Hannah Taylor
UCL Institute of Education

April 2020

The mental health and well-being of teachers is an issue of great policy concern. This is particularly true in England, where high workload and the associated stress is thought to be leading to a recruitment and retention crisis within the education profession. But do individuals who decide to leave teaching for another career actually see their well-being and mental health improve? We provide new evidence on this matter for individuals aged between 40 and 65, using the rich information gathered as part of the UK Biobank study. Our analysis shows that individuals who choose to leave teaching are somewhat happier in their work, but do not generally experience any improvement in their general well-being or mental health. We hence caution those middle-aged teachers who are thinking of leaving teaching that the grass may not necessarily be greener on the other side.

Key Words: Biobank; mental health; teachers

Contact details: John Jerrim (J.Jerrim@ucl.ac.uk) Department of Social Science, UCL Institute of Education, University College London, 20 Bedford Way London, WC1H 0AL

Acknowledgements: This research has been funded by the Nuffield Foundation. The Nuffield Foundation is an endowed charitable trust that aims to improve social well-being in the widest sense. It funds research and innovation in education and social policy and also works to build capacity in education, science and social science research. The Nuffield Foundation has funded this project, but the views expressed are those of the authors and not necessarily those of the Foundation. More information is available at www.nuffieldfoundation.org. We are grateful for their support. Helpful comments have been received on the draft from our project steering group, who we would like to thank for their input and support. This research has been conducted using the UK Biobank Resource.

For Peer Review Only

1. Introduction

Almost thirty years ago it was observed that "…teaching is an unsettled and unhappy profession at present and there is a problem to be addressed" (Smithers, 1990). A similar statement would not be out of place when describing the teaching profession today. As a job, teaching requires staff to work long hours, to keep up with changing government requirements and to manage often disruptive classrooms (Jerrim and Sims 2019). This, of course, all occurs under the watchful eye of the accountability system, with schools (and teachers) judged by how young people perform in high-stakes national examinations.

Many teachers enter the profession for altruistic reasons, yet key issues which motivate many to leave are workload, pupil behaviour and salary (Dolton & Klaauw, 1995; Barmby, 2006). Many teachers experience dissatisfaction due to feeling a lack of control with respect to their working conditions, accompanied by the absence of a platform from which to voice their concerns (Mercer & Evans, 1991). Researchers have observed similar trends for teachers outside of England, including the USA, Australia and the Netherlands, especially in newer teachers (Tye & O'Brien, 2002; Howes and Goodman-Delahunty, 2015; den Brok et al., 2017). However, strikingly, many more teachers in England expressed regret about their career choice than in almost any other industrialised country (Jerrim and Sims 2019).

Today's retention rate in the UK is low especially amongst newly qualified teachers (NQTs), and schools located in more socially deprived backgrounds experience a higher turnover of teachers (Sims and Allen, 2018). There is also some suggestive evidence that it is the more able teachers who are the most likely to leave teaching in search of alternative employment, again impacting on teacher quality, as well as there being strong monetary implications, given the cost of training teachers (Culver et al., 1990; Borman & Dowling, 2008).

Long-serving teachers in urban UK schools appear to remain in their jobs due to the deep connections and emotional ties they forge within their workplaces and communities (McIntyre, 2010). By implication then, teachers who leave the profession may do so either because they cannot forge these connections which provide an adequate level of job satisfaction, or because the challenges of workload, classroom management and salary outweigh the benefits they receive.

For Peer Review Only

Such pressures may lead to work-related stress and, in turn, be detrimental to teachers' mental and physical health (Travers & Cooper, 1993; Berryhill, Linney, & Fromewick 2009; Scheuch & Seibt, 2015; Kidger et al., 2016; Merrida-Lopez & Extremara, 2017; Education Support Partnership, 2018). A number of studies have also suggested that teachers have worse mental health and wellbeing than those who work in other jobs. For instance, Johnson et al. (2003) found teachers to have one of the lowest levels of psychological wellbeing out of the 26 occupational groups they considered. Similarly, Stansfeld et al. (2011) found that teachers were at above average risk of suffering from mental ill-health, with Kidger et al. (2016) indicating that wellbeing was low and depressive symptoms high amongst teachers. Indeed, the stresses and strains of teaching are widely cited for the ongoing teacher retention and recruitment crisis that continues to create a major challenge for England's schools (e.g. Tapper, 2018). It is therefore little wonder that, in a recent nationally representative survey of teachers, around half of those working in secondary schools in England said that they wondered whether they would have been better off working in a different job (Jerrim and Sims 2019).

This leads to an important question for teachers who are contemplating a change of career: do those who choose to leave teaching have higher levels of well-being, and a lower likelihood of developing mental health problems, than those who choose to remain? In general, occupation is known to be linked to health and wellbeing (Clark, 2010; Johnson et al., 2005; Ravesteijn, Kippersluis, & Doorslaer, 2013; Ravesteijn, Kippersluis, & Doorslaer, 2018). This is perhaps unsurprising, given the influence that work has on our social interaction with others, our ability to develop and employ new skills, our sense of personal achievement and the restrictions it imposes on other areas of life e.g. time with family (Creek & Hughes, 2008; Gallagher, Muldoon, & Pettigrew, 2015). Consequently, individuals who are unhappy may switch jobs in an attempt to improve their wellbeing. Indeed, empirical research suggests that voluntary job changes are associated with increased job satisfaction (Chadi & Hetschko, 2018; Gielen, 2013) and improved mental health (Longhi et al., 2019), at least in the short run.

The evidence in relation to teachers is, however, more limited. The study most comparable to ours is the interesting work of Bamford and Worth (2017). Using longitudinal data, these authors tracked job satisfaction, overall life satisfaction and the subjective well-being of a small sample of teachers within the UK. They found that teachers who left the profession

For Peer Review Only

experience a large increase in job satisfaction, and a small increase in subjective well-being, compared with those who decided to stay. Yet, as with all studies, the authors noted some key limitations. First, the sample size of teachers was small (e.g. 231 former teachers had left their job within the last year, 107 who had left teaching between one and two years ago and 74 who had left three years previously), with estimates surrounded by quite wide confidence intervals. Second, the focus was upon teachers who quit their job recently (mainly within the last two years) meaning it was not possible to consider the association between leaving teaching and longer-term outcomes. Finally, the Understanding Society dataset analysed collected only limited information about respondents' well-being and mental health. The combination of these factors made it challenging to compare the outcomes of individuals who chose to join, leave and stay in the teaching profession.

Our paper seeks to contribute to this understudied area, with a particular focus upon the well-being and mental health outcomes of current and former teachers aged between 40 and 65 within the UK. It is, to our knowledge, the first study to utilise the UK Biobank dataset to investigate this issue. Through this data, we have access to a wide array of information about respondents' well-being and mental health. This not only includes responses to standardised questionnaires (as have previously been used in this literature) but also prescription of common medicines used for conditions such as anxiety, depression and insomnia. These data can therefore be used to explore the mental health outcomes of current and former teachers across a wide range of important measures. Moreover, having such a rich array of data collected at baseline means that we can more credibly control for potential confounding factors within the longitudinal component of our analysis. Together, this enables us to provide important new evidence as to whether joining/leaving teaching has a sizeable impact upon a person's mental health, at a time when many teachers are thinking about changing career in search of a better life. It is important to recognise, however, that the data we use are based upon a sample of 40 to 65-year-olds and may not generalise to younger age groups.

To trail our key results, we find little evidence that the grass really is greener for those who quit the teaching profession. Although there is some evidence of an increase in job satisfaction for those who left teaching relatively recently (consistent with the findings of Bamford & Worth, 2017), there is little to suggest that leaving teaching reduces the risk of suffering from mental health problems, such as depression and anxiety. Similarly, there is

For Peer Review Only

little evidence to suggest that individuals who choose to quit teaching are happier in their life in general (i.e. outside of work). This leads us to conclude that teaching as an occupational choice is unlikely to be a major cause of poor mental health outcomes per se.

The paper now proceeds as follows. The UK Biobank dataset is described in section 2, with an overview of our empirical methodology following in section 3. Our results are detailed in section 4, with conclusions and policy discussion provided in section 5.

2. UK Biobank data

The initial UK Biobank data collection took place between 2006 and 2010. A total of around half a million volunteers between the ages of 40 and 69 participated in the study. These data therefore form a convenience sample rather than being a random sample that is representative of the wider population. Participants attended an initial assessment centre when they were first recruited, where they completed questionnaires, were interviewed by a trained health professional (in order to collect accurate information about medical conditions and currently prescribed drugs) and underwent some basic health checks (e.g. participants' blood pressure was taken; an electrocardiogram/ECG was conducted). It contains uniquely rich information about health from a very large number of individuals – many of whom were employed as teachers. A number of follow-up questionnaires have been gathered from Biobank participants since the initial assessment centre. In 2016, 117,500 participants completed an online 'occupational career' questionnaire. Respondents were first asked to type into an open text field their job title, start date and end date for each job they held. The respondent was then asked to work through a set of 'drop-down' job lists, with a list of possible job titles presented on the final screen (from which they were then asked to pick the most suitable). The relevant SOC code was then assigned to the individual, based upon their selection from the drop-down menus (information entered as free-text at the start was used for validation processes only). The data went through a validation process by an expert occupational coder, with reasonably good agreement found (Cohen's Kappa = 0.45). Further details about how the occupational career data has been captured is available within de Matteis et al (2017).

As part of the occupational career questionnaire, some further basic information was asked about each job, such as typical number of hours worked per week, whether it involved shift

For Peer Review Only

work and exposure to potentially hazardous substances (e.g. asbestos). For our purposes, this information allows us to identify the year respondents entered and exited the teaching profession, including the occupation that they joined when they left. Critically, this means we can identify both current and former teachers, facilitating comparisons between these two groups.

For each job recorded at the assessment centre or in the occupational history questionnaire, four-digit Standard Occupational Classification (SOC2000) codes are provided within the Biobank database, along with some further information about specific job role. Throughout this paper, we begin by identifying teachers using the following broad set of SOC codes:

- 2312 = Further education teaching professionals
- 2314 = Secondary education teaching professionals
- 2315 = Primary and nursery education teaching professionals
- 2316 = Special needs education teaching professionals

We excluded individuals from analysis if they were identified as working in a non-teaching role. For instance, for respondents with a SOC code of 2315, extra information was used to exclude nursery workers from our definition of 'teachers'[1]. Likewise, this extra information was used to remove further education lecturers, whilst retaining further education teachers.

Key measures collected during the 2006-2010 assessment centre

Within our analysis, we make particular use of the following measures collected during the assessment centre:

- Neuroticism score. Neuroticism was measured with the 12-item neuroticism subscale from the short form of the revised Eysenck Personality Questionnaire (EPQ-N). This encompassed 12 questions such as 'do you often feel fed-up', 'do you suffer from nerves', and 'do you often feel lonely', with participants asked to respond either yes, no, don't know or prefer not to say to each item. A total neuroticism score from these

[1] The Biobank data provides additional information about the job of each respondent, over and above the SOC code. For instance, for SOC code 2315 it provides information on whether the individual is a headteacher or not, and whether they work in a primary or nursery setting. The variable in question is available from here: https://biobank.ctsu.ox.ac.uk/crystal/field.cgi?id=132.

For Peer Review Only

12 items is provided as part of the Biobank dataset, which has been reported to have good levels of internal validity (Cronbach's alpha of 0.84 – Peters et al 2018). Within our analysis, we standardise this scale to mean zero and standard deviation one.

- Depression. Within the self-completion questionnaire, respondents were asked four questions about how they felt over the last two weeks. This included (a) frequency of depressed mood; (b) tiredness / lethargy; (c) unenthusiastic/ disinterest and (d) tenseness/restlessness. Respondents were asked to indicate one of following response options for each (1. Not at all; 2. Several days; 3. More than half the days; 4. Nearly every day; 5. Don't know; 6. Prefer not to say). Following McCormack et al (2015), we combine responses to these questions into an overall depression scale. This is done using a two-parameter item-response theory (IRT) model, which is then standardised to mean zero and standard deviation one.
- Sleep. A wide body of research has shown that mental health problems, such as anxiety and depression, are linked to insomnia and a lack of sleep (Freeman et al., 2017; Fujishiro et al., 2017). As part of the self-completion questionnaire, respondents were asked (a) the number of hours of sleep they typically get within a 24-hour period and (b) whether they either have trouble falling asleep or whether they wake up during the night (1. Never/rarely; 2. Sometimes; 3. Usually).
- Alcohol intake. Previous research has found that up to 40 percent of adults use alcohol as a mechanism to cope with stress (Appleton and James 2018). Within the assessment centre, respondents were first asked how frequently they drink alcohol (daily/almost daily; three/four times a week; once/twice a week; one to three times a month; special occasions only; never). For those who said that they drink alcohol more than once or twice per week, they were then asked about average weekly intake of (a) pints of beer/cider; (b) white wine/Champagne; (c) fortified wine; (d) red wine; (e) spirits. For those who said that they drunk alcohol monthly, or only on special occasions, monthly figures were provided. Following Taylor et al (2018) we convert these responses into an approximate number of units of alcohol consumed per week. Both frequency of drinking per week and weekly units of alcohol consumed are considered within our analysis.

For Peer Review Only

- Medications prescribed. As part of an interview with a trained nurse, respondents were asked about prescribed medications. Using this information, we create a binary variable, coded as one if they take frequently prescribed medications for anxiety, depression or insomnia[2], and zero otherwise.
- Medical conditions. In the assessment centre, participants were asked to indicate any medical conditions that they had. If the participant was uncertain of the type of illness they had, then they were asked to describe it to a trained nurse who placed it within a category. They were also asked the date or age when they were first diagnosed with the condition. Our focus is upon reports of depression, anxiety, self-harm, stress and insomnia, coded as one if they reported having one of these conditions, and zero otherwise.
- Happiness. Respondents were asked 'in general, how happy are you' with responses provided using a six-point scale (extremely happy to extremely unhappy)[3].
- Happiness with different aspects of life. Respondents were asked the same question as above, but with the focus being upon a certain aspect of their life. This included how happy they were with their (a) work; (b) family; (c) finances; (d) friends; (e) health. Responses were again provided using a six-point scale (extremely happy to extremely unhappy).

A wide range of other information was also collected within the assessment centre, including demographic background, basic health assessments (blood pressure measurements), illnesses of mother/father/siblings (including depression) and whether selected life events had occurred within the last two years (e.g. death of a family member, divorce, a serious illness). Although not the focus of this paper, a selection of such variables will be used as controls within our analysis.

Follow-up assessment centres

[2] If respondents indicated that they were currently prescribed one of the following medications, this dummy variable was coded as one: Citalopram, Escitalopram, Fluoxetine, Fluvoxamine, Amitripyline, Paroxetine, Sertraline, Venlafaxine, Duloxteine, Pregabalin, Cymbalta, Yentreve, Mirtazapine, Anafranil, Prozac, Diazepam, Zopiclone, Temazepam, Nitrazepam.

[3] The questions about happiness and work/job satisfaction were only introduced into the assessment centre questionnaire in 2009. These data are therefore only available for a subset of respondents.

For Peer Review Only

A sub-set of Biobank participants have completed return visits to the assessment centre, providing a longitudinal element to the dataset. These longitudinal follow-ups have been concentrated within certain parts of the country; most notably the assessment centres located at Cheadle (Stockport), Reading and Newcastle. Most of the same data were collected as in the initial assessment centre, including current occupation, currently prescribed medications, current medical conditions, sleep, depression, alcohol consumption and happiness with different aspects of life. Critically, this means we can identify individuals who have changed job between the two assessment centres (e.g. individuals who joined or left the teaching profession) and measure change in the aforementioned measures of mental health and well-being.

The 2016 mental health questionnaire

In 2016, a subset of 137,000 biobank participants completed an additional questionnaire about their mental health. Most of these questionnaires were completed between the 19th of August and the 26th of September 2016. Importantly, this provides a second longitudinal component to the Biobank data, with the 2006-2010 assessment centre acting as the baseline and the 2016 mental health questionnaire as the follow-up. Within our analysis, we make particular use of the following information gathered within the mental health questionnaire:

- Current depression. Respondents were asked "*over the last 2 weeks, how often have you been bothered by any of the following problems*", with nine separate questions then following (e.g. Little interest or pleasure in doing things; feeling down, depressed, or hopeless; feeling tired or having little energy). They were asked to select one of four responses for each (not at all; several days; more than half the days; nearly every day). A two-parameter IRT model is estimated using these nine items in order to construct a depression scale. We standardise this scale to mean zero and standard deviation one.
- Current anxiety. Respondents answered the same question as presented above for current depression, with seven separate items (e.g. feeling nervous, anxious or on edge; worrying too much about different things; trouble relaxing). A two-parameter IRT model is again used to derive an overall anxiety scale.

For Peer Review Only

- Prolonged periods of depression. First, respondents were asked the following two yes/no questions: "*have you ever had a time in your life when you felt sad, blue, or depressed for two weeks or more in a row*" and "*have you ever had a time in your life when you lost interest in most things like hobbies, work, or activities that usually give you please*". To respondents who answered yes to either of these questions, they were asked to recall the age that they first had such a spell and the age they were when they last had such a spell. For the purposes of this paper, this information can be compared to spells when respondents were and were not working as teachers.
- Alcohol consumption. Respondents were asked: "*In the next two questions, a "drink" is defined as one unit of alcohol. How often do you have six or more drinks on one occasion*?" We use responses to this question to get an indication of the extent that respondents engage in heavy drinking.
- Self-harm. A series of questions were asked about self-harm, including "*many people have thoughts that life is not worth living. Have you felt that way*" and "*have you contemplated harming yourself*?". They were also asked if they had felt this way in the last 12 months or if they had harmed themselves during the last 12 months.
- Happiness. Respondents were asked the same question as in the 2006-2010 assessment centre.
- Happiness with health. Respondents were asked "*In general how happy are you with your HEALTH*?", using the same six response options presented for the happiness scale.
- Feel life is meaningful. Participants were asked "*to what extent do you feel your life to be meaningful*?", with responses provided on a five-point scale (not at all; a little; a moderate amount; very much; an extreme amount).

These are the main outcome measures we consider when analysing the 2016 mental health data.

3. Methodology

Outcomes measured at the initial assessment centre

To begin, we focus upon the 2006-2010 assessment centre measures described above as our outcomes of interest. These outcomes will be compared across the following groups:

For Peer Review Only

1. Current teachers (reference group). Individuals who were teachers when the initial assessment centre took place (n = 16,622).
2. Former teachers who left within the last five years. Individuals who were teachers, but left the teaching profession for another career within the five years prior to the Biobank assessment centre (n = 1,271)[4].
3. Former teachers who left six to ten years ago. Individuals who were teachers, but left the teaching profession for another career between six and ten years prior to the Biobank assessment centre (n = 661).
4. Former teachers who left more than 10 years ago. Individuals who were teachers, but left the teaching profession for another career more than ten years prior to the Biobank assessment centre (n = 2,214).

Within this part of our analysis, the sample is restricted to individuals below age 65 at the time of the assessment centre, and who were still employed[5]. Table 1 provides some descriptive information about how the background characteristics of the Biobank sample compares to the estimates of the population of age 40-65-year-old teachers (based upon nationally representative sample surveys). On the whole, the Biobank sample is reasonably similar to these population estimates, at least in terms of the observable characteristics considered.

<< Table 1 >>

The first set of results presented in the main body of the paper are based upon the following regression model:

$$link\ (O) = \alpha + \beta.T + \tau.D + \theta.P + \varphi.M + \rho.F + \sigma.L + \gamma.H + \varepsilon \qquad (1)$$

Where:

[4] Around one-third of those who left teaching within the last five years had moved into another job in education, such as becoming a school inspector, private tutor or teaching assistant, while around two-thirds were employed in a job outside of education.

[5] The occupational questionnaire was typically completed in the summer of 2015; one year before the mental heath questionnaire. We assume that anyone who was recorded as a teacher when they completed the occupational history questionnaire was also a teacher when they completed the mental health questionnaire.

For Peer Review Only

Link () = The appropriate GLM link function for the outcome variable of interest. This will either be identity link for continuous variables (estimated by Ordinary Least Squares) or the logit link for binary/ordered categorical variables (estimated by logistic / ordinal logistic regression).

O = One of the outcomes collected during the assessment centre (as described above).

T = A vector of dummy variables capturing whether the individual was a teacher or a former teacher when they participated in the assessment centre.

D = A vector of demographic background variables such as age, gender, socio-economic status, household income, age they left education, a measure of fluid intelligence, whether they hold a degree and whether born outside of the UK.

P = Reported spells of depression before the individual entered teaching.

M = Month that they completed the Biobank assessment centre.

F = Family history of mental illness (reported that their mother, father or sibling suffered from depression or anxiety).

L = An indicator of whether a major life event (e.g. divorce, severe financial problems, had a relative die) occurred within the last two years.

H = A vector of variables capturing different aspects of household structure, including whether the respondent has a partner in the household, whether there are children in the household and household size.

Multiple imputation using chained equations is used to account for missing data within the controls. The parameter of interest is β; this captures whether individuals who use to be teachers but left for another job (i.e. former teachers) have better or worse mental health outcomes than those individuals who have chosen to remain within the teaching profession (conditional upon the factors controlled for within the model).

In Appendix A, B and C we test the robustness of these results. First, we estimate alternative specifications of these models, variously including and excluding different control variables. Second, Inverse Probability Weighting (IPW) is used as an alternative estimation approach.

URL: http://mc.manuscriptcentral.com/cber

For Peer Review Only

Third, we alter the estimation approach for continuous outcomes from OLS to zero-inflated count models.

Outcomes measured during the follow-up assessment centre

The main limitation of the analytical approach outlined above is that we can only control for a limited number of potentially confounding background characteristics. Ideally, if one wishes to make causal statements about the effect teaching has upon mental health outcomes (rather than interpreting results are purely correlational), one would need to control for all factors that are both (a) associated with the decision to leave teaching and (b) are also associated with future mental health outcomes.

The second part of our analysis takes a step closer towards reaching this goal. Specifically, we now use responses provided during the follow-up assessment centre as our outcome variables, with data from the 2006-2010 assessment centre acting as a rich set of additional controls. In other words, in this longitudinal analysis, we are interested in *change* in mental health outcomes between the two timepoints.

When using this subset of the Biobank data, our primary interest is individuals who were recorded as working as a teacher within either of the assessment centres, were still employed at the time of the second assessment centre and who were below retirement age (younger than 65). We then focus upon the following groups:

- Always teachers. Individuals who were teachers at both the baseline and follow-up assessment centres (n = 925).
- Leavers. Individuals who were teachers at the baseline assessment centre, but not the follow-up (n = 167)[6].
- Joiners. Individuals who were not teachers at the baseline assessment centre, but were at follow-up (n = 176).

A comparison of the background characteristics of these groups to all teachers who participated in the initial Biobank assessment centre can be found in Appendix D. The average

[6] Around one-third of those who left teaching moved into another job in education, such as becoming a school inspector, private tutor or teaching assistant, while around two-thirds were employed in a job outside of education.

For Peer Review Only

amount of time elapsed between the two assessment centre time points within the sample we use within our analysis is around seven years, with an average age of 58 at follow-up.

Our empirical approach is otherwise similar to that presented above, with the main difference being that we can now control for the extremely rich health data collected from participants within the initial assessment centre. Specifically, the model becomes:

$$link\ (O) = \ \alpha + \beta.T + \ \tau.D + \varphi.M + \rho.F + \delta.BC + \varepsilon \qquad (2)$$

Where:

O = One of the mental health outcomes collected during the assessment centre (as described above).

T = A vector of dummy variables capturing whether the individual was working as a teacher at both assessment centres (reference group), had left the teaching profession between the two time points or who had joined the profession.

D = A vector of demographic background variables such as age, gender, socio-economic status, age they left education, a measure of fluid intelligence, whether they hold a degree and whether a partner lives in the same household.

M = Month that they completed the Biobank assessment centre.

F = Family history of mental illness (reported that their mother, father or sibling suffered from depression or anxiety).

BC = Baseline (assessment centre) controls. This includes depression, anxiety, medications, medical conditions, general health, happiness, job satisfaction, alcohol consumption and sleep as reported during the 2006-2010 assessment centre.

Imputation is again used to account for missing data[7], while robustness tests using alternative model specifications and inverse probability weighting estimates are provided in Appendix A

[7] The amount of missing data is small for most covariates. The small number of covariates with large amounts of missing data are due to those questions only being included in assessment centres taken at later dates (e.g. questions about happiness were only included in later iterations of the Biobank assessment centre questionnaire) and not because of selective non-response. see Appendix E for further details on missing data by covariate.

For Peer Review Only

and B. The β estimates from this model now reveal whether individuals who have recently left teaching (and those who have recently joined) have better or worse mental health outcomes than their peers who have worked as teachers throughout this period.

Outcomes measured within the 2016 mental health questionnaire

We follow a similar approach in our analysis of the 2016 mental health questionnaire data; information collected during the initial assessment centre act as a rich set of controls, while responses to the 2016 mental health questionnaire are the outcome measures. Our variable of interest is based upon the information provided in the occupational history questionnaire and is defined as follows:

1. Always teachers (reference group). Individuals who were employed as teachers at both baseline (initial assessment centre) and at follow-up (2016 mental health questionnaire). N= 1,715.
2. Leavers. Those who were employed as teachers at baseline, but employed in another job at follow-up. N= 360[8].
3. Joiners. Those who not teachers at baseline, but were employed as a teacher at follow-up. N= 368.

Within this part of the analysis, the sample is restricted to those who were still employed and below retirement age (65) when they completed the mental health questionnaire. After making these restrictions, the average age of this analysis sample is 56. A comparison of the background characteristics of these groups to all teachers who participated in the initial Biobank assessment centre can be found in Appendix D. Imputation is again used to account for missing covariate data, while the substantive regression models are specified as outlined in the sub-sections above.

Effect sizes

Results are presented in terms of effect sizes; Cohen's d for continuous outcomes and odds-ratios for binary or ordinal outcomes. With respect to Cohen's d, we interpret values below

[8] Around half of those who left teaching entered another job in education (e.g. become an inspector, teaching assistant), while the other half were working outside of education.

For Peer Review Only

0.1 as evidence of essentially no effect. This is based upon two observation. First, meta-analyses have reported much larger effect sizes with respect to the impact of mental health interventions delivered in the workplace. For instance, Carolan, Harris, & Cavanagh (2017) reported an effect size of 0.37 for the impact of occupational mental health interventions upon psychological wellbeing. Second, we argue that an effect size of 0.1 would be substantively very small in this context. For instance, say that a person who left teaching scored an effect size 0.1 lower on the depression scale than those who remained in the profession. This would mean that there is only around a 52.8% chance that a person picked at random from those who have continued to be teachers will have a higher score on the depression scale than a person picked at random from the group who quit teacher for another job. In other words, the probability of suffering depression amongst members of these two groups would be little more than equal. We therefore consider any effect size below 0.1 as trivially small.

4. Results

Left teaching before the initial 2006-2010 assessment centre

The results focusing upon measures collected at the initial assessment centre can be found in Table 2. Starting with anxiety/depression, there is some limited evidence that former teachers have better outcomes on these measures than current teachers. Those individuals who left teaching within the last five years did score slightly lower on the self-reported depression scale than their peers who were still working as teachers (effect size difference of 0.10), though they were no less likely to report taking prescription medicines for common mental health problems (odds ratio = 0.95). Those who had left teaching within the last five years also scored slightly lower, on average, on the neuroticism scale (effect size 0.13). These differences are, however, quite modest in terms of magnitude. Differences are similar (or slightly smaller) when comparing current teachers to those who left the profession more than five years ago. Evidence of a link between teaching and these outcomes is hence mixed.

<< Table 2 >>

The next set of estimates turns to the issue of sleep. There is little evidence of a difference between current and former teachers in terms of the amount of sleep they get over a typical

For Peer Review Only

24-hour period; differences when expressed as an effect size are all below 0.10. Similarly, individuals who left teaching were slightly less likely to say that they had trouble falling asleep (or woke up during the night) than the reference group (current teachers). However, the effect size is again small (odds ratio ≈ 0.90). Any benefits from quitting teaching for one's quality and quantity of sleep are hence likely to be small (if at all).

The third set of outcomes presented in Table 2 refer to the consumption of alcohol. Former teachers are found to drink slightly more regularly than current teachers, though the difference is again relatively modest in magnitude (the estimated odds ratio is, at most, around 1.2). Furthermore, no difference is found between groups in terms of number of units of alcohol consumed each week. This suggests that former teachers drink roughly the same amount as individuals who have remained in teaching,

Finally, the last set of estimates presented in Table 2 refer to self-reported happiness with different aspects of life. A similar pattern again emerges. The estimated odds ratios mostly hover around one, fluctuating between 0.9 and 1.1, with no clear or obvious pattern. This holds true across most of the five specific areas of life teachers were asked about (e.g. finances, friends, family and health) as well as happiness in life overall. The one exception, where there is a sizeable difference, is happiness at work. Those who quit teaching recently – within the last five years – are happier in their jobs than those who have remained in the profession (odds-ratio 1.42). Yet our overall interpretation of the results presented in Table 2 is that individuals who choose to leave teaching are not generally happier in their life than those continue working in this career.

Joined / left teaching between the initial assessment centre and follow-up assessment centre

Table 3 replicates the analysis presented in the sub-section above, but is now based upon the longitudinal Biobank data, focusing upon those who attended the assessment centre twice. Recall that this allows us to measure change in outcomes over time, and how these relate to whether individuals moved out, into or remained in the teaching profession. There are perhaps two key points of note (over and above our discussion of the results presented in Table 2).

For Peer Review Only

<< Table 3 >>

First, there is some suggestion that that individuals who leave teaching report slightly lower levels of anxiety and depression than those who have remained in the teaching profession. Former teachers score 0.07 standard deviations lower on the self-reported depression scale, while also being somewhat less likely to report taking prescription medicines for common mental health problems (odds ratio 0.45) or reported suffering from depression/anxiety as a medical condition (odds ratio 0.78). Interestingly, those who entered the teaching profession since the baseline assessment centre scored slightly higher on the self-reported depression scale than individuals who were working as teachers at both time points (effect size 0.13) and those who had left teaching for another career (effect size 0.20). Nevertheless, the key message from Table 3 is that leaving teaching to pursue another career may only bring small benefits for one's mental health (if any at all).

Second, the final set of estimates in Table 3 (capturing self-reported happiness) potentially help strengthen the evidence that those individuals who leave teaching have higher levels of job satisfaction than those who continue to work as a teacher as their career. In particular, those who left teaching for other employment reported higher levels of satisfaction with their work than those who stayed in teaching (odds ratio 1.41). Yet this result is very much specific to the work domain; there is no evidence that those who quit teaching were happier with their health, friendships, family, health or, indeed, with life in general.

The final point to note from Table 3 is that the results with respect to sleep and alcohol consumption are largely the same as Table 2. In other words, there is little evidence that working as a teacher has an impact upon difficulties with sleeping and alcohol consumption.

Joined / left teaching between the initial assessment centre and the mental health follow-up

Finally, Table 4 presents results from our second longitudinal analysis, where the initial 2006-2010 assessment centre is the baseline, while the 2016 mental health questionnaire captures the outcomes.

<< Table 4 >>

URL: http://mc.manuscriptcentral.com/cber

For Peer Review Only

Table 4 again suggests that, if there are any differences in depression or anxiety between current and former teachers, they are trivially small. Differences in terms of effect sizes are consistently below 0.1 on the anxiety and depression scales (e.g. there is just a 0.06 standard deviation difference between current and former teachers on the self-reported depression scale). Similarly, there is little difference between current and former teachers in terms of their happiness (either overall or with their health), alcohol consumption and whether they believe that their life is meaningful.

Additional analyses in appendices

A series of additional analyses are presented within the appendices, investigating the sensitivity of our results to the different methods used. In summary:

- Appendix A investigates how results change when using a different set of controls. In this we continue to find some evidence that those individuals who left teaching are somewhat happier in their work. Otherwise, the results remained mixed, with no clear evidence that those who left teaching have clearly better mental health and wellbeing outcomes than other groups.
- Appendix B uses Inverse Probability Weighting (IPW) to adjust for covariates, rather than regression modelling. The clearest result is again that those individuals who leave teaching are somewhat happier with their work, though differences for the other outcomes tend to be small and often differ across the different approaches.
- Appendix C alters the regression modelling approach used the analysis of the anxiety, depression and Neuroticism outcome scales, accounting for the fact that many people are clustered at a single score (zero). There continues to be little evidence of sizeable differences in these outcomes between individuals who decide to join, leave and remain in the teaching profession.
- Appendix F provides alternative estimates for our models using the follow-up assessment centre and mental health questionnaire data. In particular, all prior outcome measures are removed from the model, to investigate the extent that these factors may confound the results. Interestingly, the change in the coefficients between model specifications is relatively small. This may either suggest that (a) the selection mechanism for leaving teaching for another job is relatively weak (or at least

not driven by concerns about wellbeing / mental health) or (b) that selection into/out of teaching is being driven by factors that the Biobank dataset does not measure.

Summary

In summary, the evidence for whether leaving teaching leads to lower levels of depression and anxiety is mixed. Although small improvements in these outcomes were observed within a subset of our analyses, at other times no impact was found. Our conclusion is therefore that any reduction in anxiety / depression from leaving teaching is likely to be (on average) very small, at best.

Consistent with Bamford and Worth (2017), there is some suggestion that those who decide to quit teaching end up being somewhat happier in their work than those who choose to remain. Yet, critically, this does not seem to translate into greater levels of happiness in other areas of life, including satisfaction with health or happiness overall. Consequently, the benefits of leaving teaching for one's happiness seems to be relatively minor, and concentrated in satisfaction with work. Finally, there is little evidence that the decision to leave teaching has any meaningful impact upon alcohol consumption (a mechanism many adults use to cope with stress) or quality of sleep (a marker of anxiety).

This leads us to reach an overall conclusion that leaving teaching for another job is unlikely to bring significant benefits to well-being or mental health.

5. Conclusions

It is widely thought that teaching is a demanding job, due to the long working hours, the stress imposed by the accountability system and the challenge of having to manage often disruptive classrooms (De Carlo et al., 2019). Across the UK, and England in particular, this has led to many teachers considering a change of career (Perryman & Calvert, 2019). Indeed, in a recent international study, half of secondary teachers in England said that they wondered whether they should have chosen another profession – higher than in almost any other country across the world (Jerrim and Sims 2019). Critically, in a recent survey of teachers in England who quit the profession, half said that their job was making them ill (Perryman & Calvert, 2019) and a high number of teachers interviewed in the Teacher Wellbeing Index (2018) experienced

For Peer Review Only

anxiety, depression or acute stress (Education Support Partnership, 2018). But is life really that much better for those individuals who decide to quit teaching to pursue a different job? This paper has presented important new evidence on this issue, with a particular focus upon how leaving teaching is linked to well-being and mental health.

Looking across an array of outcomes, and using several different empirical approaches, we have found little evidence of a link between leaving teaching, lower prevalence of mental health problems and higher levels of general well-being. Throughout our analysis, effect sizes have been small and often changed in both magnitude and direction depending upon the empirical approach taken. The one exception is happiness with work (job satisfaction), where we find a fairly consistent improvement for those who have recently left the teaching profession. Two observations are important in interpreting this finding. First, this is consistent with empirical research a range of occupations, which finds that job satisfaction tends to fall in the period prior to an individual quitting a job, before rising during the early stages of their new employment (Chadi & Hetschko, 2018; Gielen, 2013; Longhi et al., 2019). Second, we do not observe reductions in job satisfaction for those joining the teaching profession. Taken together, this suggests that the increase in job satisfaction for those leaving teaching is not indicative of any particular problem with teaching. Rather, it is more likely to reflect a natural process by which those who are less suited to the job move into alternative occupations (Gielen, 2013).

There are, of course, limitations with this paper and the need for future research. Five important issues stand out. First, some of the occupational career data has been based upon respondents recalling and accurately reporting such information. Although our focus upon relatively recent timepoints should limit the impact this has upon our analysis, collection of prospective longitudinal data from a cohort of teachers (tracking their entry and exits from the profession) would represent an important advance in the literature. Second, although a number of potential confounding factors were controlled within our analysis, our estimates continue to rely upon a (untestable) selection-upon-observables assumption if they are to be interpreted as causal effects. Third, while we have considered some "harder" evidence of poor mental health (e.g. prescription of antidepressant medicines), many of the outcome measures we investigated are based upon self-completion questionnaires. Although this is standard within this literature (Lenderink & Zoer, 2012), further work using a wider array of

For Peer Review Only

outcome data (e.g. primary care records) would represent a further step forward in this field. Fourth, the number of current and former teachers included within our analysis is relatively modest (in the hundreds rather than the thousands). Larger samples in the future would likely yield more precise results. Finally, the Biobank data is a convenience sample focused upon middle-aged participants, which is not representative of the wider teacher population. In particular, all participants were aged between 40 and 65, meaning it is not possible to generalise our findings to younger age groups. Future work using nationally representative data would undoubtedly enhance the external validity of our findings.

Despite these limitations, we believe that this paper has helped to advance our knowledge of teacher well-being and their mental health. At a time when many teachers are thinking about leaving for another career, it is vital that they are fully informed about the likely consequences. For those teachers who are not satisfied with their work, changing jobs may lead to an increase in job satisfaction. However, our results suggest that quitting teaching for alternative employment is unlikely to lead to improvements in general well-being or mental health. With respect to the latter outcomes, teaching does not stand out relative to other occupations.

For Peer Review Only

Data availability

The data that support the findings of this study are available from the UK Biobank. Restrictions apply to the availability of these data, which were used under license for this study. Data are available from https://www.ukbiobank.ac.uk/with the permission of the UK Biobank.

Ethical guidelines

The paper was produced following the BERA ethical guidelines (https://www.bera.ac.uk/). The project was approved by the UCL Institute of Education ethics committee (REC 1158).

Conflicts of Interest

The authors have no conflict of interest.

Funding

The paper has been funded by the Nuffield Foundation. The Nuffield Foundation is an independent charitable trust with a mission to advance social well-being. It funds research that informs social policy, primarily in Education, Welfare, and Justice. It also funds student programmes that provide opportunities for young people to develop skills in quantitative and scientific methods. The Nuffield Foundation is the founder and co-funder of the Nuffield Council on Bioethics and the Ada Lovelace Institute. The Foundation has funded this project, but the views expressed are those of the authors and not necessarily the Foundation. Visit www.nuffieldfoundation.org.

For Peer Review Only

References

Appleton, A., & James, R. (2018). Adults (18-75) in the UK who drink alcohol for coping reasons. Drinkaware research report. Accessed 04/07/2019 from https://www.drinkaware.co.uk/media/1852/drinking-to-cope-jan-2018.pdf

Bamford, S., & Worth, J. (2017). Is the grass greener beyond teaching? Slough: NFER. Accessed 04/07/2019 from https://www.nfer.ac.uk/teacher-retention-and-turnover-research-research-update-3-is-the-grass-greener-beyond-teaching

Berryhill, J., Linney, J., & Fromewick, J. (2009). The effects of education accountability on teachers: are policies too stress provoking for their own good? *International Journal of Education Policy & Leadership, 4*(5), 1-14.

Borman, G. D., & Dowling, N. M. (2008). Teacher attrition and retention: A meta-analytic and narrative review of the research. *Review of Educational Research, 78*(3), 367-409.

Culver, S., Wolfle, L., & Cross, L.H. (1990). Testing a model of teacher satisfaction for blacks and whites, *American Educational Research Journal, 27*(2), 323-349.

Carolan, S., Harris, P., & Cavanagh, K. (2017). Improving employee well-being and effectiveness: systematic review and meta-analysis of web-based psychological interventions delivered in the workplace. *Journal of Medical Internet Research, 19*(7).

Chadi, A., & Hetschko, C. (2018). The magic of the new: How job changes affect job satisfaction. *Journal of Economics & Management Strategy*, *27*(1), 23-39.

Clark, A. (2010). Work, jobs, and well-being across the millennium. *International differences in well-being*, 436-468.

Creek, J., & Hughes, A. (2008). Occupation and health: a review of selected literature. *British Journal of Occupational Therapy*, *71*(11), 456-468.

De Carlo, A., Girardi, D., Falco, A., Dal Corso, L., Di Sipio, A. (2019). When does work interfere with teachers' private life? An application of the job demands-resources model. *Frontiers in Psychology, 10*(1).

De Matteis, S., Jarvis, D., Young, H., Young, A., Allen, N., Potts, J., Darnton, A., Rushton, L., & Cullinan, P. (2017). Occupational self-coding and automatic recording (OSCAR): a novel web-based tool to collect and code lifetime job histories in large population-based studies. *Scandinavian Journal of Work, Environment and Health, 43*(2), 181-186.

den Brok, P., Wubbels, T., & Van Tartwijk, J. (2017). Exploring beginning teachers' attrition in the Netherlands. *Teachers and Teaching*, 23(8), 881-895.

Dolton, P., & Klaauw, W. V. D. (1995). Leaving teaching in the UK: A duration analysis. *The Economic Journal*, *105*(429), 431-444.

Education Support Partnership, 2018. *Teacher Wellbeing Index 2018*. Accessed 04/11/2019 from https://www.educationsupportpartnership.org.uk/sites/default/files/teacher_wellbeing_index_2018.pdf

Freeman, D., Sheaves, B., Goodwin, G. M., Yu, L. M., Nickless, A., Harrison, P. J., ... & Hinds, C. (2017). The effects of improving sleep on mental health (OASIS): a randomised controlled trial with mediation analysis. *The Lancet Psychiatry, 4*(10), 749-758.

Fujishiro, K., Farley, A., Kellemen, M., & Swoboda, C. (2017). Exploring associations between state education initiatives and teachers' sleep: A social-ecological approach. *Social Science and Medicine, 191*, 151-159.

Gallagher, M., Muldoon, O. T., & Pettigrew, J. (2015). An integrative review of social and occupational factors influencing health and wellbeing. *Frontiers in Psychology*, *6*.

Gielen, A. C. (2013). Repeated job quits: stepping stones or learning about quality?. *IZA Journal of European Labor Studies*, *2*(7).

Howes, L. M., & Goodman-Delahunty, J. (2015). Teachers' career decisions: Perspectives on choosing teaching careers, and on staying or leaving. *Issues in Educational Research, 25*(1), 18.

Jerrim, J. and Sims, S. 2019. The Teaching and Learning International Survey 2018. Department for Education Research Report. Accessed 04/07/2019 from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/809737/TALIS_2018_research.pdf

Johnson, S., Cooper, C., Cartwright, S., Doland, I., Taylor, P., & Millet, C. (2005). The experience of work-related stress across occupations. *Journal of Managerial Psychology,* 20(2), 178-187.

Kidger, J., Brockman, R., Tilling, K., Campbell, R., Ford, T., Araya, R., King, M., & Gunnell, D. (2016). Teachers' wellbeing and depressive symptoms, and associated risk factors: A large cross-sectional study in English secondary schools. *Journal of Affective Disorders, 192*, 76-82.

Longhi, S., Nandi, A., Bryan, M., Connolly, S., & Gedikli, C. (2019). *Do all job changes increase wellbeing?* What Works Wellbeing.

Mérida-López, S., & Extremera, N. (2017). Emotional intelligence and teacher burnout: A systematic review. *International Journal of Educational Research, 85*, 121-130.

Lenderink, A., & Zoer, I. (2012). *Review on the validity and reliability of self-reported work-related illness*. Health and Safety Executive Research Report RR903. Accessed 04/07/2019 from http://www.hse.gov.uk/research/rrpdf/rr903.pdf

Mercer, D., & Evans, B. (1991). Professional myopia: Job satisfaction and the management of teachers. *School Organization, 11*(3), 291-301.

Mérida-López, S. & Extremera, N. (2017). Emotional intelligence and teacher burnout: A systematic review. *International Journal of Educational Research, 85*, 121-130.

McCormack, A., Edmondson-Jones, M., Fortnum, H., Dawes, P., Middleton, H., Munro, K., & Moore, D. (2015). Investigating the association between tinnitus severity and symptoms of depression and anxiety, while controlling for neuroticism, in a large middle-aged UK population. *International Journal of Audiology, 54*(9): 599-604.

McIntyre, J. (2010). Why they sat still: the ideas and values of long-serving teachers in challenging inner-city schools in England. *Teachers and Teaching: theory and practice, 16*(5), 595-614.

Perryman, J., & Calvert, G. (2019). What motivates people to teach, and why do they leave? Accountability, performativity and teacher retention. *British Journal of Educational Studies*, *68*(1), 3-23.

Peters, E., John, A., Bowen, R., Baetz, M., & Balbuena, L. (2018). Neuroticism and suicide in a general population cohort: results from the UK Biobank Project. *BJPsych Open*, *4*(2), 62–68.

Ravesteijn, B., van Kippersluis, H., & van Doorslaer, E. (2013). The Contribution of Occupation to Health Inequality. *Research on Economic Inequality*, *21*, 1-20.

Ravesteijn, B., Kippersluis, H. V., & Doorslaer, E. V. (2018). The wear and tear on health: What is the role of occupation?. *Health Economics*, *27*(2), e69-e86.

Scheuch, K., Haufe, E., & Seibt, R. (2015). Teachers' health. *Deutsches Ärzteblatt International*, *112*(20), 347.

Sims, S., & Allen, R. (2018). Identifying schools with high usage and high loss of newly qualified teachers. *National Institute Economic Review*, 243(1), R27-R36

Smithers, A. (1990). *Teacher Loss*. Manchester: University of Manchester.

Stansfeld, S., Rasul, F, Head, J., & Singleton, N. (2011). Occupation and mental health in a national UK survey. *Social Psychiatry and Psychiatric Epidemiology 46*(2), 101–110.

Tapper, J. (2018, 13 May). Burned out: why are so many teachers quitting or off sick with stress? *The Guardian.*

Taylor, M., Rode, L., Bjørngaard, J., Taylor, A., Bojesen, S., Åsvold, B., Gabrielsen, M., Lewis, G., Nordestgaard, B., Romundstad, P., Hickman, M., & Munafò, M. (2018). Is smoking heaviness causally associated with alcohol use? A Mendelian randomization study in four European cohorts, *International Journal of Epidemiology, 47*(4), 1098–1105.

Tye, B., & O'brien, L. (2002). Why are experienced teachers leaving the profession?. *Phi Delta Kappan, 84*(1), 24-32.

For Peer Review Only

Table 1. Characteristics of the Biobank sample compared to population estimates

	Biobank	Population estimate	
		Estimate	Source
Average age	53	51	LFS
% male	27%	27%	LFS
% children in household	53%	44%	LFS
% Partner in household	76%	74%	LFS
% hold a degree	84%	75%	LFS
Average age left school	20	21	LFS
Born outside UK	7%	7%	LFS
Homeowner	95%	94%	LFS
Smoker	6%	4%	APS 2010
% poor general health	1%	1%	NCDS 2008
% fair general health	14%	7%	NCDS 2008
% good general health	63%	65%	NCDS 2008
% excellent general health	22%	26%	NCDS 2008
Total teachers	**16,622**		

Notes: LFS = Labour Force Survey data for 40-65-year-old teachers from January-March sweeps 2007,2008 and 2009. APS = Annual Population Survey data from 2010 for 40-65-year-old teachers. NCDS = National Child Development Survey from 2008 (when respondents were 50-years-old). The NCDS data for 'good' and 'very good' general health has been combined. APS 2010 data based upon information across all education and teaching professionals.

For Peer Review Only

Table 2. The association between leaving/remaining in the teaching profession and mental health outcomes measured in the 2006-2010 Biobank assessment centre.

	Left teaching last 5 years		Left teaching 6 - 10 years ago		Left teaching > 10 years ago	
	Beta	**T-Stat**	**Beta**	**T-Stat**	**Beta**	**T-Stat**
Depression / anxiety						
Self-reported depression (ES)	-0.10	-3.71	-0.08	-2.20	0.00	-0.21
Prescribed medicines (OR)	0.95	-0.31	0.85	-0.81	0.82	-1.57
Self-reported medical condition (OR)	1.07	0.54	0.98	-0.14	1.08	0.77
Neuroticism (ES)	-0.13	-4.19	-0.11	-2.52	-0.08	-3.65
Sleep						
Hours sleep per night (ES)	0.08	3.54	0.01	0.17	0.04	2.22
Trouble falling sleep (OR)	0.90	-1.97	0.87	-1.90	0.89	-2.56
Alcohol						
Frequent drinking (OR)	1.19	3.25	1.16	2.10	1.19	4.26
Number alcohol units per week	0.06	0.27	-0.22	-0.70	0.09	0.50
Happiness						
Happy with work (OR)	1.42	2.91	1.13	0.98	1.09	1.32
Happy with finances (OR)	0.95	-0.50	0.92	-0.63	0.97	-0.35
Happy with friends (OR)	0.85	-1.75	1.00	-0.01	0.91	-1.25
Happy with family (OR)	0.86	-1.60	0.88	-1.01	0.94	-0.79
Happy with health (OR)	0.98	-0.24	0.91	-0.42	1.22	3.67
Overall happiness (OR)	1.01	0.11	0.94	-0.48	1.05	0.66

Notes: Those individuals who were currently teachers at the time of the assessment centre are the reference group. 'ES' refers to estimated effect size for continuous outcomes and 'OR' to the odds ratio for binary/categorical outcomes. Estimates based upon regression models controlling for age, month visited the assessment centre, gender, fluid intelligence score, Townsend socio-economic status index, family history of depression, whether a major negative life event occurred within last two years, whether has a partner living in same household, whether children live in the same household, household size, household income, age left education, whether hold a degree, whether born outside of the UK and whether had ever suffered depression before working as a teacher. Multiple imputation has been used to account for missing covariate data.

For Peer Review Only

Table 3. The association between leaving/remaining in the teaching profession with outcomes measured in the follow-up Biobank assessment centre

	Left teaching		Joined teaching	
	Beta	**T-Stat**	**Beta**	**T-Stat**
Depression / anxiety				
Self-reported depression (ES)	-0.07	-1.01	0.13	1.78
Prescribed medicines (OR)	0.45	-1.47	0.82	-0.41
Self-reported medical condition (OR)	0.78	-0.75	1.20	0.61
Sleep				
Hours sleep per night (ES)	0.04	0.63	-0.04	-0.62
Trouble falling sleep (OR)	0.90	-0.60	0.73	-1.82
Alcohol				
Frequent drinking (OR)	1.00	-0.02	0.77	-1.59
Number units of alcohol per week	-1.60	-1.79	-2.30	-2.57
Happiness				
Happy with work (OR)	1.41	2.09	1.03	0.20
Happy with finances (OR)	0.84	-1.07	0.72	-1.99
Happy with friends (OR)	0.80	-1.31	0.86	-0.88
Happy with family (OR)	1.18	1.03	1.03	0.15
Happy with health (OR)	0.89	-0.70	0.83	-1.11
Overall happiness (OR)	1.02	0.12	0.84	-0.97

Notes: Those individuals who were teachers at both baseline (initial assessment centre) and follow-up (follow-up assessment centre) are the reference group. 'ES' refers to estimated effect size for continuous outcomes and 'OR' to the risk ratio for binary/categorical outcomes. Estimates based upon regression models controlling for age, month visited the assessment centre, gender, whether has a partner living in same household, whether hold a degree, fluid intelligence score, Townsend socio-economic status index, age left education, family history of depression, self-reported depression at baseline, medical conditions and prescriptions reported at baseline, general health reported at baseline, happiness with work/friends/family/finances/health at baseline, alcohol intake at baseline, sleep quality and quantity at baseline, whether suffered prolonged spell of depression before baseline. Multiple imputation has been used to account for missing covariate data.

For Peer Review Only

Table 4. The association between leaving/remaining in the teaching profession and outcomes measured in the 2016 mental health questionnaire.

	Left teaching since baseline		Joined teaching since baseline	
	Beta	**T-Stat**	**Beta**	**T-Stat**
Depression / anxiety				
Self-reported depression (ES)	0.06	0.96	0.05	0.80
Self-reported anxiety (ES)	0.09	1.28	0.06	0.85
Prolonged depress spell since baseline (OR)	1.11	0.51	0.89	-0.61
Self-harm				
Considered self-harm in last 12 months (OR)	1.28	0.72	1.82	1.82
Alcohol consumption				
Frequently drink >6 drinks (OR)	0.89	-0.82	0.73	-2.15
Other				
Happy with health (OR)	0.82	-1.30	1.06	0.35
Feel life is meaningful (OR)	1.03	0.17	0.79	-1.43
Overall happiness (OR)	0.82	-0.95	0.78	-1.68

Notes: Those individuals who were teachers at both baseline (assessment centre) and follow-up (mental health questionnaire) are the reference group. 'ES' refers to estimated effect size for continuous outcomes and 'OR' to the odds ratio for binary/categorical outcomes. Estimates based upon regression models controlling for age, month visited the assessment centre, gender, whether has a partner living in same household, whether hold a degree, fluid intelligence score, Townsend socio-economic status index, age left education, family history of depression, self-reported depression at baseline, medical conditions and prescriptions reported at baseline, general health reported at baseline, happiness with work/friends/family/finances/health at baseline, alcohol intake at baseline, sleep quality and quantity at baseline, whether suffered prolonged spell of depression before baseline. Multiple imputation has been used to account for missing covariate data.

For Peer Review Only

Appendix A. Alternative regression model estimates including different sets of controls

Table A1. The association between leaving/remaining in the teaching profession and mental health outcomes measured in the 2006-2010 Biobank assessment centre. Alternative set of controls

	Left teaching last 5 years		Left teaching 6 - 10 years ago		Left teaching > 10 years ago	
	Beta	T-Stat	Beta	T-Stat	Beta	T-Stat
Depression / anxiety						
Self-reported depression (ES)	-0.18	-3.79	-0.14	-2.05	-0.05	-1.23
Prescribed medicines (OR)	0.89	-0.81	0.78	-1.25	0.77	-2.12
Self-reported medical condition (OR)	1.05	0.41	0.98	-0.13	1.02	0.27
Neuroticism (ES)	-0.12	-4.29	-0.09	-2.34	-0.10	-4.67
Sleep						
Hours sleep per night (ES)	0.09	3.94	0.01	0.39	0.04	2.31
Trouble falling sleep (OR)	0.90	-1.94	0.85	-2.33	0.88	-3.10
Alcohol						
Frequent drinking (OR)	1.15	2.85	1.17	2.37	1.26	5.82
Units of alcohol per week	-0.02	-0.08	-0.13	-0.44	0.22	1.27
Happiness						
Happy with work (OR)	1.49	4.94	1.17	1.25	1.09	1.02
Happy with finances (OR)	0.98	-0.22	0.90	-0.90	1.11	1.50
Happy with friends (OR)	0.85	-1.85	0.96	-0.33	0.92	-1.24
Happy with family (OR)	0.85	-1.84	0.86	-1.26	0.94	-0.95
Happy with health (OR)	1.04	0.48	0.95	-0.45	1.23	2.69
Overall happiness (OR)	1.06	0.74	0.96	-0.32	1.11	1.15

Notes: Estimates based upon regression models controlling for age, month completed assessment centre, gender, mother/father/sibling ever had depression, immigrant status, partner in the household, children in the household, age finished education, whether respondent holds a degree and whether first instance of depression occurred before they became a teacher. Multiple imputation has been used to account for missing covariate data. See notes to Table 2 for further details.

For Peer Review Only

Table A2. The association between leaving/remaining in the teaching profession with outcomes measured in the follow-up Biobank assessment centre. Alternative set of controls.

	Left teaching		Joined teaching	
	Beta	**T-Stat**	**Beta**	**T-Stat**
Depression / anxiety				
Self-reported depression (ES)	-0.05	-0.76	0.11	1.67
Prescribed medicines (OR)	0.58	-1.16	0.75	-0.69
Self-reported medical condition (OR)	0.75	-0.96	1.01	0.03
Sleep				
Hours sleep per night (ES)	0.01	0.19	-0.06	-0.95
Trouble falling sleep (OR)	0.90	-0.62	0.80	-1.47
Alcohol				
Frequent drinking (OR)	1.00	-0.01	0.70	-2.38
Number of units per week	-1.13	-1.30	-1.95	-2.41
Happiness				
Happy with work (OR)	1.26	1.47	1.03	0.22
Happy with finances (OR)	0.77	-1.66	0.81	-1.46
Happy with friends (OR)	0.82	-1.21	0.97	-0.23
Happy with family (OR)	1.13	0.81	1.03	0.18
Happy with health (OR)	0.78	-1.57	0.81	-1.41
Overall happiness (OR)	0.96	-0.25	0.98	-0.15

Notes: Estimates based upon regression models controlling for age, month completed assessment centre, gender, mother/father ever had depression, partner in the household, age finished education and the following measures recorded at the baseline assessment centre: depression scale, prescription of anti-depressants, self-reported medical condition of depression/anxiety/insomnia, happiness (overall, with work and with health), general health, difficulty sleeping and alcohol consumption. Multiple imputation has been used to account for missing covariate data. See notes to Table 2 for further details.

For Peer Review Only

Table A3. The association between leaving/remaining in the teaching profession and outcomes measured in the 2016 mental health questionnaire. Alternative set of controls.

	Left teaching since baseline		Joined teaching since baseline	
	Beta	T-Stat	Beta	T-Stat
Depression / anxiety				
Self-reported depression (ES)	0.05	0.72	0.02	0.37
Self-reported anxiety (ES)	0.07	1.09	0.04	0.68
Prolonged depress spell since baseline (OR)	1.09	0.44	0.87	-0.82
Self-harm				
Considered self-harm in last 12 months (OR)	1.24	0.63	1.66	1.51
Alcohol consumption				
Frequently drink >6 drinks (OR)	0.96	-0.31	0.84	-1.44
Other				
Happy with health (OR)	0.81	-1.63	1.06	0.38
Feel life is meaningful (OR)	0.99	-0.05	0.86	-1.05
Overall happiness (OR)	0.80	-1.22	0.84	-1.17

Notes: Estimates based upon regression models controlling for age, gender, age finished education, partner in household, and the following measures recorded at the baseline assessment centre: depression scale, happiness (overall and with different areas), general health, difficulty sleeping and alcohol consumption. Imputation has been used to account for missing covariate data. See notes to Table 4 for further details.

For Peer Review Only

Appendix B. Inverse probability weighting estimates

Table B1. The association between leaving/remaining in the teaching profession and mental health outcomes measured in the initial Biobank assessment centre. IPW estimates converted to effect sizes for all variables.

	Left teaching last 5 years		Left teaching 6 - 10 years ago		Left teaching > 10 years ago	
	Beta	T-Stat	Beta	T-Stat	Beta	T-Stat
Depression / anxiety						
Self-reported depression	-0.10	-3.65	-0.21	-1.69	-0.01	-0.22
Prescribed medicines	-0.01	-0.20	0.00	-0.76	0.00	-1.51
Self-reported medical condition	0.03	0.87	0.00	0.11	0.00	0.96
Neuroticism	-0.13	-4.11	-0.10	-2.55	-0.08	-3.42
Sleep						
Hours sleep per night	0.10	3.40	0.01	0.25	0.03	1.77
Trouble falling sleep	-0.06	-1.96	-0.04	-1.93	-0.03	-2.44
Alcohol						
Frequent drinking	0.09	3.21	0.20	2.60	0.18	3.87
Units alcohol per week	0.01	0.35	-1.88	-0.70	0.74	0.44
Happiness						
Happy with work	0.19	4.12	0.09	1.94	0.02	0.90
Happy with finances	-0.01	-0.34	-0.01	-0.32	0.00	0.18
Happy with friends	-0.08	-1.74	0.00	-0.04	-0.02	-0.96
Happy with family	-0.07	-1.48	-0.05	-0.92	-0.01	-0.47
Happy with health	0.00	0.06	-0.02	-0.45	0.07	2.94
Overall happiness	-0.02	-0.34	-0.01	-0.36	0.01	0.39

Notes: Those individuals who were currently teachers at the time of the assessment centre are the reference group. Figures have been converted into an approximate effect size by dividing the estimate by the sample standard deviation for the variable. Estimates based upon IPW, including the following variables in the matching model: age, month of assessment centre, gender, fluid intelligence, socio-economic status, family history of depression, major negative life event occurred in last two years, whether born in UK, household income, household structure, age left school, whether hold a degree and whether experienced depression before becoming a teacher. Single imputation used to account for missing covariate data.

For Peer Review Only

Table B2. The association between leaving/remaining in the teaching profession with outcomes measured in the follow-up Biobank assessment centre. IPW estimates converted to effect sizes for all variables.

	Left teaching		Joined teaching	
	Beta	**T-Stat**	**Beta**	**T-Stat**
Depression / anxiety				
Self-reported depression	-0.08	-1.08	0.14	1.80
Prescribed medicines	-0.11	-1.58	-0.01	-0.19
Self-reported medical condition	-0.06	-0.83	0.04	0.47
Sleep				
Hours sleep per night	0.03	0.32	-0.06	-0.73
Trouble falling sleep	-0.05	-0.59	-0.14	-1.87
Alcohol				
Frequent drinking	0.00	-0.04	-0.13	-2.29
Number units alcohol per week	-0.13	-1.67	-0.18	-2.44
Happiness				
Happy with work	0.20	2.27	0.02	0.28
Happy with finances	-0.10	-1.21	-0.14	-1.78
Happy with friends	-0.03	-0.40	-0.03	-0.34
Happy with family	0.11	1.38	-0.01	-0.10
Happy with health	-0.06	-0.75	-0.07	-0.95
Overall happiness	0.02	0.22	0.00	-0.02

Notes: Figures have been converted into an approximate effect size by dividing the estimate by the sample standard deviation for the variable. Estimates based upon IPW, including the following variables in the matching model: age, month of assessment centre, gender, whether partner is in the household, fluid intelligence, socio-economic status, family history of depression, age left school and baseline measures of depression, prescriptions for mental health issues, self-reported mental health issue, happiness, general health, quality and quantity of sleep, alcohol intake. Single imputation used to account for missing covariate data.

For Peer Review Only

Table B3. The association between leaving/remaining in the teaching profession and outcomes measured in the 2016 mental health questionnaire. IPW estimates converted to effect sizes for all variables.

	Left teaching since baseline		Joined teaching since baseline	
	Beta	**T-Stat**	**Beta**	**T-Stat**
Depression / anxiety				
Self-reported depression	0.03	0.42	-0.01	-0.08
Self-reported anxiety	0.06	0.88	0.05	0.50
Prolonged depress spell since baseline	0.02	0.43	-0.02	-0.48
Self-harm				
Considered self-harm in last 12 months	0.00	0.02	0.05	0.55
Alcohol consumption				
Frequently drink >6 drinks	-0.08	-1.24	-0.17	-1.96
Other				
Happy with health	-0.13	-1.77	0.00	-0.02
Feel life is meaningful	-0.01	-0.20	-0.19	-1.87
Overall happiness	-0.05	-0.82	-0.14	-1.55

Notes: Figures have been converted into an approximate effect size by dividing the estimate by the sample standard deviation for the variable. Estimates based upon IPW, including the following variables in the matching model: age, gender, whether partner is in the household, socio-economic status, family history of depression, whether hold a degree, age left school and baseline measures of depression, prescriptions for mental health issues, self-reported mental health issue, happiness, general health, sleep and alcohol intake. Single imputation used to account for missing covariate data.

For Peer Review Only

Appendix C. Zero-inflated models

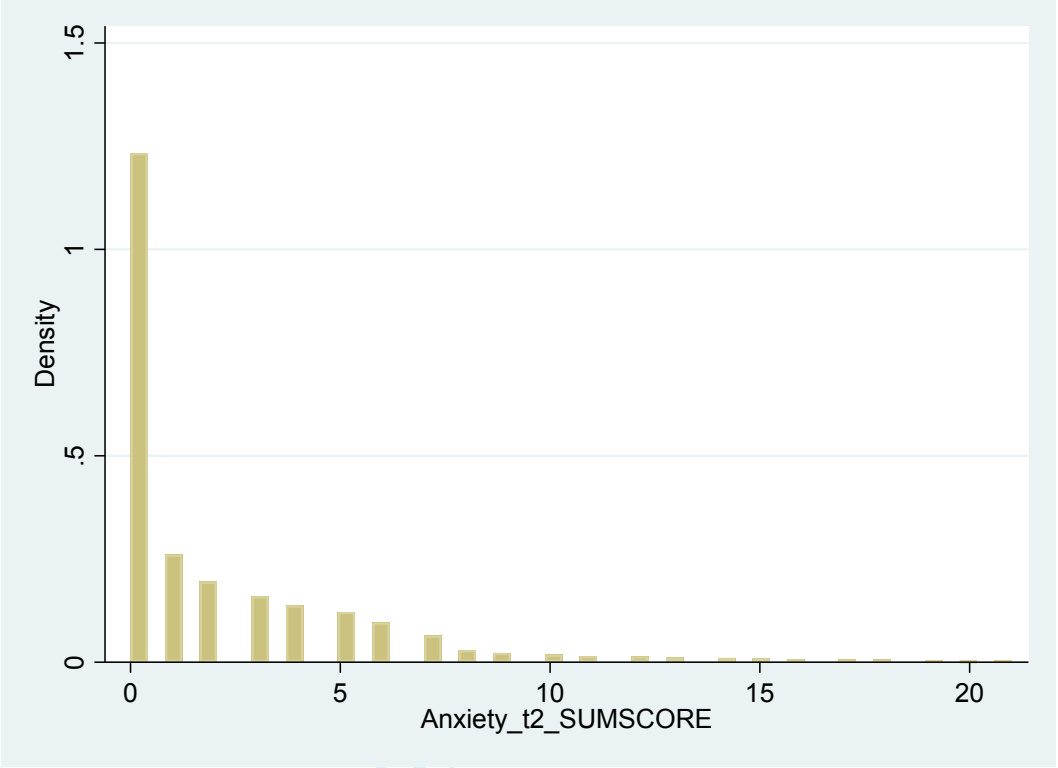
Within the main body of the paper, we treat the following variables as continuous and hence use ordinary least squares regression:

- Anxiety score
- Depression score
- Neuroticism score

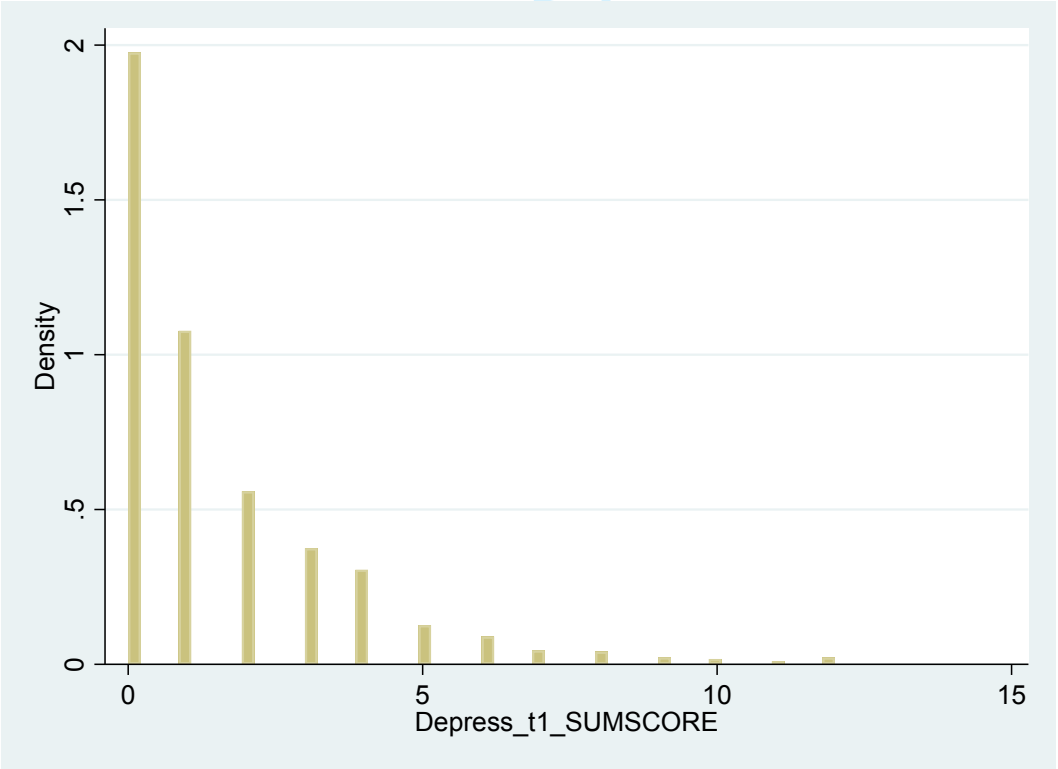
The reason for using OLS regression is that allows for straightforward estimation and interpretation as an effect size. However, as illustrated by Appendix Figures C1, C2 and C3 below, these outcomes may actually be considered to be "zero-inflated" data (meaning there is a large cluster of observations at zero). Consequently, to test the robustness of our results for these outcomes, in this appendix we estimate Zero-Inflated Poisson (ZIP) models. ZIP models essentially divide the outcome into two parts. The first is a binary response (logit) model, which estimates the probability of whether the outcome is zero or not. The second part is a count (Poisson) model, which models the values greater than zero. We implement these models using the same predictors as discussed in the main body of the paper for both parts of the ZIP model. Results are presented in Appendix Tables C1 to C3. These present estimates as odds-ratios for the 'inflation' model (i.e. the logit model which models whether the probability of the outcome being zero or not) and incidence-rate-ratios for the 'outcome' model (i.e. the Poisson count regression estimating the score on the scale).

Appendix Figure C1. Distribution of the Anxiety sum score variable

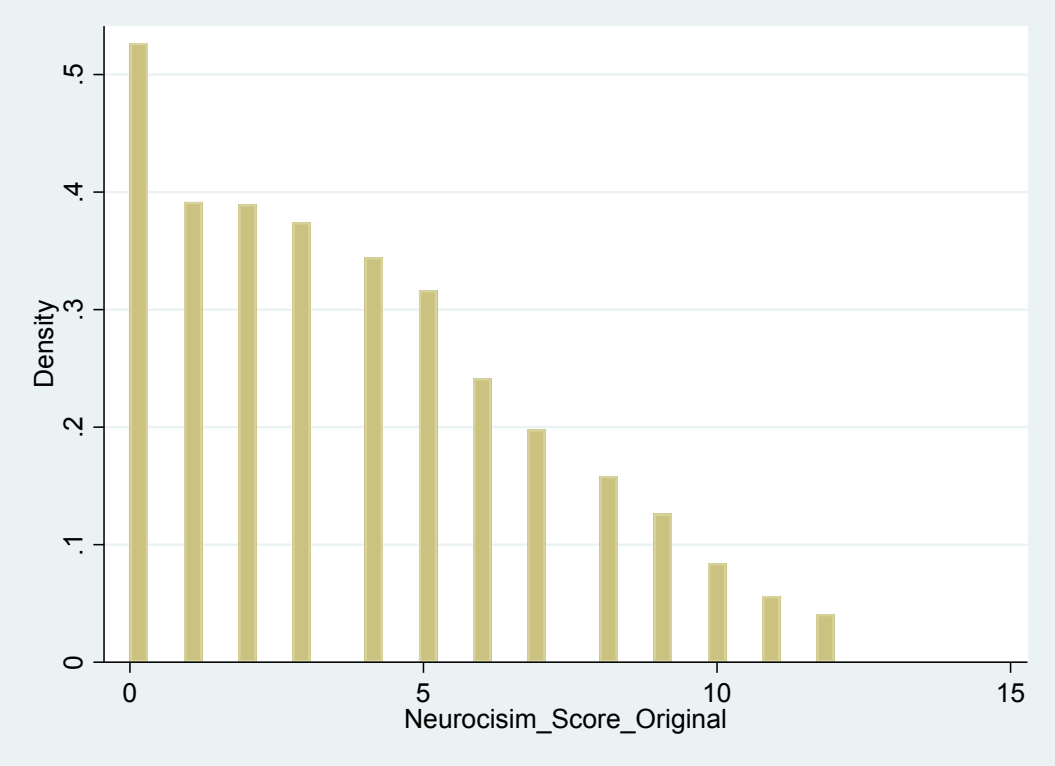
For Peer Review Only



Appendix Figure C2. Distribution of the Depression sum score variable



Appendix Figure C3. Distribution of the Neuroticism sum score variable



Appendix Table C1. Zero-Inflated Poisson estimates for selected outcome variables measures in the initial assessment centre

	Depression scale		Neuroticism scale	
	Effect	T-Stat	Effect	T-Stat
Outcome model (IRR)				
Left in last 5 years	0.90	3.43	0.94	3.92
Left 6-10 years ago	0.90	2.53	0.94	2.87
Left 11+ years ago	0.94	2.49	0.94	4.89
Inflation model (OR)				
Left in last 5 years	1.08	0.87	1.22	2.28
Left 6-10 years ago	0.90	0.83	1.02	0.14
Left 11+ years ago	0.85	2.19	1.06	0.84

Appendix Table C2. Zero-Inflated Poisson estimates for selected outcome variables measures in the follow-up assessment centre

	Depression scale	
	Effect	T-Stat
Outcome model (IRR)		
Left teaching since baseline	0.96	0.93
Joined teaching since baseline	1.04	0.94
Inflation model (OR)		
Left teaching since baseline	0.75	1.35
Joined teaching since baseline	0.62	2.38

For Peer Review Only

Appendix Table C3. Zero-Inflated Poisson estimates for selected outcome variables measures in the mental health questionnaire follow-up

	Depression scale		Anxiety scale	
	Effect	T-Stat	Effect	T-Stat
Outcome model (IRR)				
Left teaching since baseline	1.03	1.36	1.04	0.91
Joined teaching since baseline	1.02	1.08	0.94	1.47
Inflation model (OR)				
Left teaching since baseline	0.88	0.86	0.84	1.15
Joined teaching since baseline	0.74	1.87	0.69	2.46

For Peer Review Only

Appendix D. Characteristics of the assessment centre and mental health questionnaire longitudinal follow-up samples

(a) Assessment centre sample

	Original sample at assessment centre 1	Still teaching assessment centre 2	Left teaching by assessment centre 2	Joined teaching by assessment centre 2
Average age at first contact	53	51	50	51
% male	27%	29%	34%	36%
% children in household at first contact	53%	64%	66%	52%
% Partner in household at first contact	76%	77%	74%	76%
% hold a degree	84%	86%	77%	64%
Born outside UK	7%	6%	6%	5%
Homeowner at first contact	95%	98%	93%	96%
Smoker at first contact	6%	4%	4%	7%
General health at first contact				
% poor	1%	1%	1%	0%
% fair	14%	13%	11%	12%
% good	63%	62%	63%	62%
% excellent	22%	24%	25%	26%
Household income at first contact				
£18,000<	14%	11%	20%	17%
£18,000 - £30,999	38%	37%	36%	45%
£31,000 - £51,999	43%	47%	38%	36%
£52,000+	6%	6%	6%	4%
Total teachers	**16,622**	**925**	**167**	**176**

(b) Mental health questionnaire

	Original sample at assessment centre 1	Teacher at assessment centre and mental health questionnaire	Left teaching	Joined teaching
Average age at first contact	53	48	50	49
% male	27%	21%	31%	28%
% children in household at first contact	53%	71%	66%	69%
% Partner in household at first contact	76%	75%	73%	74%
% hold a degree	84%	92%	88%	78%
Born outside UK	7%	7%	10%	8%
Homeowner at first contact	95%	95%	93%	91%
Smoker at first contact	6%	5%	7%	5%
General health at first contact				
% poor	1%	1%	1%	1%
% fair	14%	12%	15%	15%
% good	63%	61%	58%	63%
% excellent	22%	26%	26%	21%
Household income at first contact				
£18,000<	14%	8%	16%	15%
£18,000 - £30,999	38%	35%	40%	36%
£31,000 - £51,999	43%	50%	37%	40%
£52,000+	6%	7%	6%	9%
Total teachers	**16,622**	**1,715**	**360**	**368**

Appendix E. The percentage of observations missing information in each covariate

	Assessment centre 1 analysis	Assessment centre 2 (longitudinal) analysis	Mental health questionnaire analysis
Age first contact	0%	0%	0%
Gender	0%	0%	0%
Fluid intelligence score	64%	67%	60%
Townsend index	0%	0%	0%
Maternal depression	1%	2%	1%
Paternal depression	3%	3%	3%
Sibling depression	3%	3%	3%
Relative die soon before first contact	0%	0%	0%
Had serious illness soon before fist contact	0%	0%	0%
Divorced soon before first contact	0%	0%	0%
Financial problems soon before first contact	0%	0%	0%
Born outside UK	0%	0%	0%
Partner in household at first contact	0%	0%	0%
Children in household at first contact	0%	0%	0%
Household size	0%	0%	0%
Household income	6%	5%	4%
Whether hold a degree	0%	0%	0%
Experienced depression before teaching	49%	29%	21%
Depression scale at 1st contact	-	3%	3%
Prescribed antidepressants at 1st contact	-	0%	0%
Depression as medical condition 1st contact	-	0%	0%
Happy at first contact	-	67%	60%

For Peer Review Only

Happy at work at first contact	-	66%	60%
General health at first contact	-	0%	0%
Happy with health at first contact	-	66%	60%
Happy with family at first contact	-	67%	60%
Happy with friends at first contact	-	67%	60%
Happy with finance at first contact	-	66%	60%
Hours sleep per night at first contact	-	0%	0%
Insomnia at first contact	-	0%	0%
Alcohol intake at first contact	-	0%	0%
Number of units at first contact	-	0%	0%
Full sample	**20768**	**1268**	**2443**

For Peer Review Only

Appendix F. Alternative estimates for the longitudinal analysis with different sets of controls

In this appendix we repeat our analysis of the follow-up assessment centre and mental health questionnaire data, but now using a different set of control variables. Specifically, we remove from the model the prior outcome measures that were measured in the initial Biobank assessment centre. The motivation behind this exploration is to investigate that the inclusion or exclusion of these controls changes the results. Appendix Table F1 lists the covariates across the two sets of models being compared. Results for the follow-up Biobank Assessment Centre (analogous to those presented in Table 3) can be found in Appendix Table F2. Those for the mental health questionnaire (analogous to those presented in Table 4) can be found in Appendix Table F3. Interestingly, the inclusion or exclusions of the prior outcome measures collected in the initial assessment centre do not seem to substantively alter the results or the key conclusions reached.

Appendix Table F1. Covariates included across the two model specifications

	Without additional controls	With additional controls
Age	Y	Y
Month of assessment centre visit	Y	Y
Gender	Y	Y
Partner in household	Y	Y
Hold a degree	Y	Y
Fluid intelligence score	Y	Y

Townsend socio-economic status index	Y	Y
Age left education	Y	Y
Family history of depression	Y	Y
Medical conditions and prescriptions reported at baseline	-	Y
General health reported at baseline	-	Y
Happiness with work at baseline	-	Y
Happiness with friends at baseline	-	Y
Happiness with family at baseline	-	Y
Happiness with finances at baseline	-	Y
Happiness with health at baseline	-	Y
Alcohol intake at baseline	-	Y
Sleep quality and quantity at baseline	-	Y
Suffered spell of depression before baseline	-	Y

Appendix Table F2. Alternative estimates of the association between leaving/remaining in the teaching profession with outcomes measured in the follow-up Biobank assessment centre (analogous to Table 3)

	Left teaching		Joined teaching	
	Without additional controls	With additional controls	Without additional controls	With additional controls
Depression / anxiety				
Self-reported depression (ES)	-0.05	-0.07	0.10	0.13
Prescribed medicines (OR)	0.56	0.45	0.84	0.82
Self-reported medical condition (OR)	0.82	0.78	1.19	1.20
Sleep				
Hours sleep per night (ES)	0.01	0.04	-0.01	-0.04
Trouble falling sleep (OR)	0.94	0.90	0.69	0.73
Alcohol				
Frequent drinking (OR)	1.03	1.00	0.87	0.77
Number units of alcohol per week	-1.74	-1.60	-2.30	-2.30
Happiness				
Happy with work (OR)	1.40	1.41	1.11	1.03
Happy with finances (OR)	0.87	0.84	0.76	0.72
Happy with friends (OR)	0.80	0.80	0.91	0.86
Happy with family (OR)	1.17	1.18	1.10	1.03
Happy with health (OR)	0.95	0.89	1.02	0.83

For Peer Review Only

Overall happiness (OR)	1.01	1.02	0.94	0.84

Appendix Table F3. Alternative estimates of the association between leaving/remaining in the teaching profession and outcomes measured in the 2016 mental health questionnaire (analogous to Table 4)

	Left teaching since baseline		Joined teaching since baseline	
	Without additional controls	With additional controls	Without additional controls	With additional controls
Depression / anxiety				
Self-reported depression (ES)	0.06	0.06	0.05	0.05
Self-reported anxiety (ES)	0.10	0.09	0.06	0.06
Prolonged depress spell since baseline (OR)	1.01	1.11	0.78	0.89
Self-harm				
Considered self-harm in last 12 months (OR)	1.22	1.28	1.62	1.82
Alcohol consumption				
Frequently drink >6 drinks (OR)	1.00	0.89	1.01	0.73
Other				
Happy with health (OR)	0.95	0.82	0.96	1.06
Feel life is meaningful (OR)	0.96	1.03	0.84	0.79
Overall happiness (OR)	0.8	0.82	0.84	0.78

For Peer Review Only

For Peer Review Only