

The *when* and the *how* of the emergence of social inequality in mental health: exploring social causation and health selection through employment transitions

Version accepted for publication in *Research in Social Stratification and Mobility*

Author

Lewis R. Anderson

Department of Sociology & Nuffield College

University of Oxford

lewis.anderson@sociology.ox.ac.uk

ORCID: 0000-0001-6295-7391

The published version of this article can be accessed at:

<https://doi.org/10.1016/j.rssm.2021.100642>

Materials to reproduce the analysis are available at:

<https://doi.org/10.17605/OSF.IO/9AE45>

This document includes the Appendix.

Abstract

How, and at what stage of the adult lifecourse, does the association between income and mental health problems arise? Research mostly tests whether mechanisms of social causation or health selection *are evident* in a given sample. I test for these mechanisms, and the contribution of unobserved heterogeneity, but additionally explore *the extent to which each contributes* to growth in social inequality in mental health, specifically symptoms of depression and anxiety and their negative correlation with household income. Using Understanding Society data, I first show that inequality in mental health among adults emerges around ages 20-30 and then persists, only weakening from around 60. Inequality is much lower controlling for employment. Next, I apply a novel fixed-effects longitudinal structural equation model to test three mechanisms generating growth of this inequality, operationalising social causation and health selection through employment transitions. While leaving employment exhibits a negative association with subsequent mental health, neither this mechanism nor the reverse – health selection out of employment – can account for growth in mental health inequality. Rather, only unobserved heterogeneity between individuals accounts for a substantial portion of this growth – around a third. This result is similar for men and women and across age groups. These findings lend support to the relatively neglected indirect health selection hypothesis, and indicate that a priority for future work should be to more clearly delineate what sorts of relatively fixed characteristics of individuals might matter, through what sorts of pathways, and how these characteristics are moulded across childhood and adolescence.

1 Introduction

The association between income (and other indicators of socioeconomic status) and mental health problems is a topic of enduring importance for social scientists and public health researchers (Brown & Harris, 1978; Costello et al., 2003; Dohrenwend et al., 1992; Reeves et al., 2017). The mental health problems examined here are symptoms of depression and anxiety, which impose enormous costs on society and exhibit a negative social gradient in their distribution (Department for Work and Pensions & Department of Health, 2009; Friedrich, 2017; Fryers et al., 2005; Kessler, 2012; Lorant et al., 2003). In this paper I draw on several strands of the health inequalities literature to attempt to answer the following question: when and how does the association between income and mental health problems arise? In past

research, hypotheses relating to the *how* – social causation, health selection, or indirect health selection – have received considerably more attention than the *when* (Kröger et al., 2015). This paper aims to attend to both elements of the question, hypothesising that different mechanisms for the generation of mental health inequality may take precedence at different lifecourse stages.

In attending to the *how*, much health inequalities research has tested whether mechanisms of social causation and health selection *are evident* in a given sample. Blane et al. (1999) and Chandola et al. (2003) are among those to have drawn the distinction between such tests and – further – examining the contribution of these mechanisms to explaining gradients in health. That is, social position may affect subsequent health, but this in itself does not imply anything about the importance of social causation in the generation of health inequalities: the effect may be detectable but small in magnitude and ultimately not a substantively important mechanism. The same of course may be said about health selection, that is, an effect of health on social position.

The two studies cited above apply this insight with respect to health inequality between social classes and employment grades respectively, but exclude the non-employed. Attending only to the continuously employed may however miss a substantial part of the true extent of health inequality across the adult population (Dahl, 1993; van de Mheen et al., 1999). Indeed some evidence indicates that health (both physical and mental) is related more strongly to transitions in and out of employment than to moves between occupational or class categories (Elstad, 2004; Ki et al., 2011; Richards & Paskov, 2016). Methodologically advanced recent work has reported negative health effects of unemployment and job loss (Krug & Eberl, 2018; Steele et al., 2013), adding support to the social causation hypothesis. However this work in turn does not examine the contribution of such employment transitions to explaining social gradients in health. Addressing this limitation may be particularly important given that in the UK context, ‘[b]etween 1973 and 2009, the relation between good health and securing and sustaining employment has strengthened for both men and women’ (Minton et al., 2012: 1).

It is not straightforward to incorporate the non-employed in a model of the generation of health inequality, since the measure of socioeconomic status used is often based on an individual’s current employment, and simply ranking the unemployed or non-employed as the lowest grade or social class has notable limitations. In this paper I include employment transitions in a model

of mental health inequality in which the extent of that inequality is measured by the contemporaneous association between household income and mental health. Existing research suggests that mental health problems may be more robustly related to income than social class (Fryers et al., 2003; Richards & Paskov, 2016).

Associations between mental health, SES, and transitions in and out of employment are likely to vary across the lifecourse, making certain life stages and turning points more plausible than others as sites for the generation of mental health inequalities. In particular, I make the case that the transition to adulthood is a period during which the association between mental health and SES is especially likely to strengthen due to health selection, while social causation may play a more important role in mid-adulthood. I conceptualise the transition to adulthood as spanning exit from compulsory education to economic independence from one's parents. Lifecourse scholars argue that this transition has expanded from being concentrated in the early 20s, to spanning the ages from 18 to the early- or even mid-30s (Cook & Furstenberg, 2002; Furstenberg et al., 2004; Swartz et al., 2011).

Consideration of the three hypotheses generally applied to the explanation of health inequalities should lead scholars to view the *when* as inextricably linked to the *how*. One potential answer to the *how* is the indirect health selection hypothesis – the idea that the health-SES relation is confounded rather than arising from the causal effects of one upon the other. This hypothesis has drawn support, as Hoffmann et al. (2018: 30) observe, from ‘several studies that begin observing the relationship between SES and health in mid-life, find little or no mutual effects, and conclude that health inequalities can be explained by indirect selection.’ Such studies are agnostic as to whether finding no evidence of social causation or health selection within a given age range means that omitted (and perhaps time-invariant) variables determine health and SES through the lifecourse, or – a very different conclusion – whether in fact social causation and health selection *do* generate this inequality, but do so at earlier ages, prior to the observation window, beyond which it merely persists as cohorts age.

Attention to timing, then, is crucial both because different causal mechanisms may take priority at different ages, and because ignoring timing can lead to confusion around the interpretation of tests of the different mechanisms. Policy implications follow from each of these points. Measures to counteract or ameliorate the impact of SES on health or vice versa will be mistargeted if focused on age groups in which those mechanisms do not in fact operate.

Moreover, perceptions of the fairness of health inequalities could be coloured by studies which attend only to restricted age ranges, find no evidence of either causal mechanism, and conclude that the SES-health relationship is confounded by fixed traits such as personality or cognitive ability (Gottfredson, 2004; Link et al., 2008), or by adolescent lifestyle (Glendinning et al., 1995).

This paper uses eight waves of nationally representative panel data from the UK Household Longitudinal Study, also known as Understanding Society. First, I describe cross-sectional inequality in mental health by age, revealing at which stages of the lifecourse inequalities are more or less apparent, and over which periods they increase, remain stable, and decline. Adding economic status to this static model accounts for a large part of the contemporaneous income-mental health association and motivates a focus on the role of employment transitions. In the next step I present a series of longitudinal structural equation models. These models test for the presence of social causation, health selection, and indirect health selection; and they measure the extent to which these mechanisms account for observed increases in mental health inequality. Finally, results are compared for different age groups, to examine the timing of the different mechanisms generating mental health inequalities. Before coming to the empirical part of the paper, I further discuss relevant literature to contextualise what follows.

2 Background and hypotheses

Three longstanding hypotheses regarding the question of *how* social inequalities in mental health arise are social causation (SES affects mental health), health selection (mental health affects SES), and indirect health selection (confounding; no causal effects between the two). A focus of this paper is the *when*: if social causation and health selection occur, does the strength of each of these effects – and their contribution to generating health inequality – vary by age, and if so how? I consider first some evidence on when the social gradient in mental health forms; this work directs our attention to the period of transition to adulthood and the processes that occur therein.

2.1 Generation of the mental health gradient

One particularly relevant cluster of evidence describes the (re-)emergence of inequalities in mental health from late adolescence to early adulthood. Both cross-sectional (Vallejo-Torres

et al., 2014) and longitudinal (Sweeting et al., 2016) evidence from the UK indicate an absence of mental health inequality in late adolescence (ages 16-19 for Vallejo-Torres et al., and age 18 for Sweeting et al.). In each of these studies, mental health inequality has emerged by the next age group or measurement point respectively (ages 20-24 and age 24). Similarly, Wickrama et al. (2009), using US data, find equalisation of depressive symptoms according to parental education from early (ages 13/14) to late (19/20) adolescence, but re-emergence of inequalities by age 23.

If the aforementioned studies are indicative of a lower bound for the lifecourse stage at which the social gradient in mental health is generated, a second relevant body of work suggests an upper bound. Though highly informative contributions, this latter group effectively ignores the transition to adulthood as a site for the generation of (mental) health inequalities. Foverskov and Holm (2016) aim to test the hypotheses of social causation, health selection, and indirect selection using the British Household Panel Survey. They restrict their analyses to men and women aged between 30 and 60, and model health as a single latent factor including a measure of psychological wellbeing. In a sensitivity analysis the GHQ-12 measure of depressive symptoms is used as the measure of health, with the same results. While cross-sectional inequalities are evident, they find no evidence of prior health influencing later SES or vice versa, and so conclude that '[t]he well-known cross-sectional correlations between health and SEP in adulthood seem not to be driven by a causal relationship, but instead by dynamics and influences in place before the respondents turn 30 years old that affect both their health and SEP onwards' (Foverskov & Holm, 2016: 180-181). Richards and Paskov (2016) do not explicitly test these causal hypotheses, but rather explore mechanisms which might account for the social class gradient in mental health, limiting their analyses to those aged 25-65. Their fixed-effects regressions show that changes in social class are unrelated to changes in GHQ score. They therefore reach a similar conclusion, suggesting the importance of unobserved heterogeneity due to factors acting earlier in life which influence both the psychological wellbeing and social class of individuals in adulthood. These studies both, then, point towards younger ages for the actual generation of mental health inequalities.

These two strands of literature taken together suggest that social causation and/or health selection occur across the transition to adulthood, generating mental health inequalities which persist thereafter. However, there is further complexity to consider. In the case of Foverskov and Holm (2016), the authors test for effects at various lag lengths of household income on

mental health and vice versa. Economic status is not included in the model.¹ This appears to reflect a tacit assumption that economic status only affects mental health through income. On the contrary, a large literature attests psychosocial and social pathways too: the loss of purpose, social networks, status, time structure, and sense of control (Brand, 2015; Krug & Eberl, 2018; Newman, 1988). A more plausible causal model includes economic status as a direct and synchronous determinant of both income and mental health; indeed this is a central assumption of the current paper.

In the case of Richards and Paskov (2016), economic status also looms large. A second major finding from their paper is that, conditional on economic status, there is no social class gradient in mental health. The implication is that lower social class is associated with poorer mental health precisely because it entails a greater risk of unemployment. This is wholly in line with the theoretical underpinnings of the NS-SEC measure used, under which lower social class position is characterised in part by the greater instability of the type of employment contract that defines it (Erikson & Goldthorpe, 1992).

The finding of no social causation and no health selection among individuals in mid-adulthood thus neglects the possibility that those processes do occur within this age range, and do so through employment transitions: job loss may represent social causation if it leads to worse mental health, and health selection if it is precipitated by poor mental health.

2.2 Health selection and social causation over the lifecourse

Clark and Oswald (1994) report a stronger association between unemployment and depressive symptoms among those aged 30-49 than older or younger. In general however, previous work on the links between labour market status and mental health appears to have paid little attention to variation by age, tending either to aggregate the whole working-age population or examine specific age groups. At what lifecourse stages are employment transitions – and other pathways of social causation and health selection – likely to have the greatest impact? Considerations

¹ Foverskov and Holm (2016) use the Cambridge Social Interaction and Stratification Scale as an alternative measure of SES, but not, as they indeed admit, in a manner that takes into account employment transitions: non-employed individuals are given a value based on their ‘last know [sic] position’ (2016: 181).

deriving from lifecourse theory highlight the transition to adulthood as a period when health selection is especially likely to play out.

This paper conceptualises the transition to adulthood primarily as a period during which an individual's SES shifts from that of their parents, to their own. This shift from parental to own SES entails the *formation* of an individual's own SES. The distal causes of success here – for instance genetic endowment, socialisation, parental time investments, and educational resources – may extend back throughout the earlier stages of life. However the proximal causes of one's early own SES, such as educational and early-career decision-making, performance in school and university examinations, and job applications, are contingent events occurring around the ages 16-25. By contingent I mean to emphasise that although individuals may be endowed with or gradually develop traits and abilities which predispose them to a high early own SES, nevertheless the translation of this latent predisposition into actual, realised SES depends on events and processes which are vulnerable to disruption by negative shocks to mental health. One's nascent own SES could for instance be impacted by an episode of poor mental health coinciding with examinations, or disrupting the job search process. The cognitive characteristics of common mental health problems may influence decision-making relevant to SES formation. Relatively pessimistic thinking among the depressed is one such characteristic (Alloy et al., 1990; Wisco, 2009), implying that individuals would make less risky – or ambitious – educational and career decisions.

There is ample evidence for health selection at this lifecourse stage. US research indicates that adolescent depressive symptoms may disrupt the status attainment process (Kessler et al., 1995; McLeod & Fettes, 2007), for instance through an association with risky behaviours such as committing crimes and excessive drinking (Wickrama & Wickrama, 2010). Studies of the 1979 US National Longitudinal Survey of Youth report associations between higher initial levels of depressive symptoms (at ages 27-35) and higher odds of both remaining (Prause & Dooley, 2001) and becoming (Dooley et al., 2000) unemployed two years later. Samples from the lower end of the transition to adulthood age range report similar conclusions: Swedish 16-year-olds with above-average levels of depressive symptoms were at greater risk both of any and of long-term unemployment over the following five years (Hammarström & Janlert, 1997); low self-esteem among high school students in the US was associated with spending a greater proportion of the following seven years unemployed (Dooley & Prause, 1997). The most robust evidence for mental health influencing employment status among young adults is the sibling fixed-

effects models reported by Egan et al. (2016): individuals with depressive symptom scores at least one standard deviation above the mean at age 16-20 were substantially more likely to be unemployed or out of the labour force 11 years later, and experienced substantially more time unemployed over the follow-up period.

Due then to its contingent and provisional nature, SES in the transition to adulthood may be particularly vulnerable to health selection effects. At later stages of working-age life, episodes of poor mental health may adversely affect SES, but individuals are more likely to have accrued resources with which to buffer this impact, such as financial savings, the support of a long-term partner, or tenure and firm-specific knowledge as barriers to job loss.

Social causation on the other hand – and particularly the impact of job loss and unemployment – may be especially strong through mid-adulthood. This is a period of increasing responsibility, not only in the workplace and for children, but also for elderly parents (Aneshensel et al., 1995; Lachman et al., 2015). While job loss may be less likely in this period, it may have a greater impact on mental health for being relatively unexpected and arriving when many have reached their peak in seniority and earnings (Lachman, 2004). Moreover, the accumulation through adulthood of social and financial resources mentioned earlier is socially stratified; individuals living in economic precariousness will be especially vulnerable to the effects of job loss.

Finally, in contrast to the above, the large literature documenting the importance of childhood factors for mental health in adulthood suggests that causal relations between mental health and SES might not be observed in the transition to adulthood, nor at any point in working-age life (Kessler et al., 2010; McLaughlin et al., 2012; Power et al., 2002, 2007). Rather, this would suggest that a conclusion similar to that of Foverskov and Holm (2016) and Richards and Paskov (2016) will be found, supporting the indirect health selection hypothesis and highlighting the importance of even earlier ages in determining individuals' SES and mental health trajectories through the lifecourse.

2.3 Unobserved heterogeneity

I consider one final topic before presenting the empirical part of the paper: unobserved heterogeneity and its relevance for the social causation and health selection hypotheses.

Recent research has increasingly used cross-lagged panel models to test these two hypotheses. Foverskov and Holm's (2016) study is one example, in which the authors apply the Anderson-Hsiao fixed-effects estimator (Anderson & Hsiao, 1982). As these authors note, in contrast to their own analysis the majority of similar previous studies harness both between- and within-individual variation rather than excising unobserved heterogeneity by focusing on within variation (Aittomäki et al., 2012; Chandola et al., 2003; Elovainio et al., 2011; Mulatu & Schooler, 2002; Steele et al., 2013; Warren, 2009). Moreover, these studies all find support for social causation, and nonsignificant or relatively weak health selection effects. For instance, Chandola et al. (2003) report consistent associations of prior employment grade with subsequent mental health in the Whitehall II study. They only find the opposite when modelling SES as financial deprivation, and that only for men, with effects only 25-40% the size of the equivalent social causation estimates. Warren (2009) explicitly addresses the threat of unobserved heterogeneity by including measures of child health and SES, but these observed measures are unlikely to capture the full range of possibly relevant time-invariant individual factors which might jointly influence SES and health across the lifecourse. Steele et al. (2013) explore the (social causation) effects of employment transitions through simultaneous equations for the set of transitions and GHQ. The authors adjust for selection through terms linking prior GHQ both with GHQ at t and with the $t-1$ to t employment transitions; they also allow nonzero correlations between individual-level random effects for each outcome, thus allowing 'for the possibility that changes in mental health and employment transitions are jointly determined' by time-invariant unobservables (2013: 706). This is a relatively strong design but nevertheless requires that the random effects assumption be made with respect to the other controls in the model: age, marital status, care responsibilities, and occupational class.

The one analysis taking a similar approach to Foverskov and Holm (2016) and utilising only within-individual variation in a cross-lagged model, and which like Chandola et al. (2003) uses the Whitehall II study, in fact finds support for health selection and not for social causation (Case & Paxson, 2011). However the lag structure proposed in Case and Paxson's model is highly questionable, implying that a health shock between measurement points 1 and 2 causes a decline in SES *between* measurement points 3 and 4. It is hard to see why negative effects might not show up at measurement point 2, in which case they are missed by this model, or at measurement point 3, in which case they lower the baseline from which deviations at point 4 are measured. Moreover, the Whitehall II study represents a selected group, as in fact Case and Paxson (2011) demonstrate by comparison with the nationally representative NCDS and BCS.

Nevertheless, it is suggestive that conclusions differ depending on whether unobserved heterogeneity is accounted for.

Foverskov and Holm (2016) thus both identify and add to an important discrepancy in the literature. The current evidence, despite its limitations, suggests a major role for unobserved heterogeneity in biasing tests of the health selection and social causation hypotheses. One of the contributions of this paper is to strengthen the evidence with a method robust to unobserved heterogeneity – by including individual fixed effects.

3 Data and measures

This paper uses eight waves of the nationally representative UK Household Longitudinal Study, also known as Understanding Society. The first wave was collected between 2009 and 2011, and the eighth between 2016 and 2018, with individuals surveyed at approximately annual intervals. From wave 2 onwards, Understanding Society has absorbed the British Household Panel Survey (BHPS), yielding a total combined sample of 39,942 households. Of these, 30,428 (76.2%) responded partially or fully at wave 2, with 50,389 (59.4%) of 84,891 adults (those aged 16 or older) giving a full interview.

Understanding Society has a complex survey design, which is accounted for by including design weights and using the *svy* suite of commands in Stata 14. The missing data approach is described in the Appendix.

3.1 Mental health problems

The General Health Questionnaire 12-item version (GHQ or GHQ-12) is used to measure mental health. The GHQ is a widely-used measure of the symptoms of common mental health problems including depression, anxiety, and associated somatic complaints (Goldberg et al., 1997; Jackson, 2007). Each of the 12 items is scored from 0 to 3; for example: ‘Have you recently lost much sleep over worry?’ (0 = *not at all*, 1 = *no more than usual*, 2 = *rather more than usual*, 3 = *much more than usual*). This ‘Likert’ scoring method gives a total score ranging from 0 to 36, which I then standardise (within sex and wave). Items are listed in full in the Appendix.

3.2 Household income

The raw variable in the Understanding Society data is monthly household income, net of taxes on earnings.² This measure sums the net incomes of all household members, including those who declined an interview or were interviewed by proxy. For individuals with missing data on any component of income, Understanding Society imputes values based on a wide range of personal and contextual information. Extensive detail is given in the *User Guide* (Knies, 2018). In the following order, values of this raw variable are equivalised, logged, top- and bottom-coded at the 99th and 1st percentiles, and within-wave standardised.

Equivalisation is applied using the modified OECD equivalence scale supplied with the data. This adjusts for household size and composition. The log transform is applied to reflect expected diminishing returns of income on mental health; in other words, linearity in the relationship is not expected to hold at extreme values (this was confirmed by a binned scatterplot – see Supplementary Figure S1 in the Appendix). Top- and bottom-coding is used for the same reason, as well as to reduce the influence of outliers. The within-wave standardisation has two purposes: to ease interpretation and to adjust for inflation and shifts in the income distribution between waves.

Understanding Society income data has been validated against the Households Below Average Income series, the data source for official UK statistics on the income distribution, which is based on a high-quality specialist income survey, the Family Resources Survey (Fisher et al., 2019).

3.3 Employment transitions

Self-reported economic status is categorised as either *employed* or *not employed*. On this basis, four binary variables are constructed indicating each of the possible (non-)transitions in employment status across adjacent waves. I refer to these as *EN* (i.e. employed at $t-1$, not employed at t), *NE*, *EE*, and *NN*. Because the models place high demands on the data, in the analysis I aggregate *NE* and *EE* for simplicity, leaving **E* as the reference category – those

² As also reported by Hounkpatin et al. (2015) from their examination of two other sources of data, *personal* income correlated no more closely with depressive symptoms than did household income.

currently in employment, regardless of their $t-1$ status. The groups represented by *NE* and *EE* report similar GHQ scores (Thomas et al., 2005).

Descriptive statistics (Table S1) and a disaggregation of *NE* and *NN* into specific *not employed* statuses (Table S2) are presented in the Appendix.

4 Analytical approach

I take a structural equation modelling (SEM) approach throughout the paper, and consistently index the (unexplained) social gradient in mental health by the residual correlation between GHQ and household income z-scores; that is, the synchronous correlation between GHQ and household income that is not accounted for by any of the other processes specified in the model. In the figures below, the idiosyncratic error terms for income and GHQ are denoted by ϵ and υ respectively.

The first stage of the analysis is descriptive. To describe cross-sectional inequality in mental health at each age, I pool all available observations and, for each age from 16 to 70, report the income-mental health correlation (a in Figure 1a). This is done separately for men and women, as for all analyses in this paper. Next, I explore to what extent any evident gradient remains, conditional on economic status. A dichotomous variable indicating whether or not the individual is employed is added to the model. This variable is included as a contemporaneous influence on both income and mental health (Figure 1b), and I again report a . Results from these preliminary steps (presented below) confirm two expectations from the literature reviewed above: that the gradient in mental health is negligible in late adolescence and grows over the following decades; and second, economic status explains a large portion of this cross-sectional inequality and should be central to a proposed causal model.

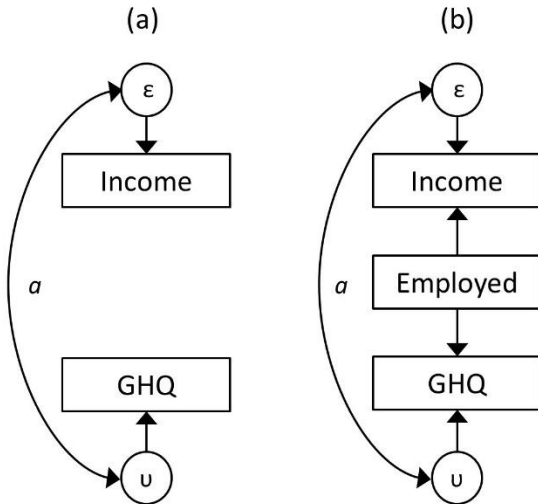


Figure 1 Cross-sectional structural equation models describing the social gradient in mental health, unconditionally (a) and conditional on economic status (b)

In the second stage of the analysis I introduce a sequence of six longitudinal models, utilising the eight available observations for each individual. In each case, the parameters of interest are constrained to be the same across waves. Age differences in the results are thus inferred from differences in results between older and younger individuals, not from differences within the same individual over time. This is discussed below.

Figures 2 and 3 show each of the six models. Model 1 includes path a only, reproducing the initial descriptive model shown in Figure 1a. Model 2 adds the autoregressive parameters b and c ; this introduces the first-wave measures of income and GHQ, which are allowed to correlate (path d). With this addition, the residual correlations a are now conditional on income and GHQ at $t-1$ and thus capture the year-to-year change in mental health inequality.

Models 3, 4, and 5 each separately add paths to this baseline model (2). The paths added represent the three hypotheses of social causation, health selection, and indirect health selection respectively. In Model 6 these terms are all included simultaneously.

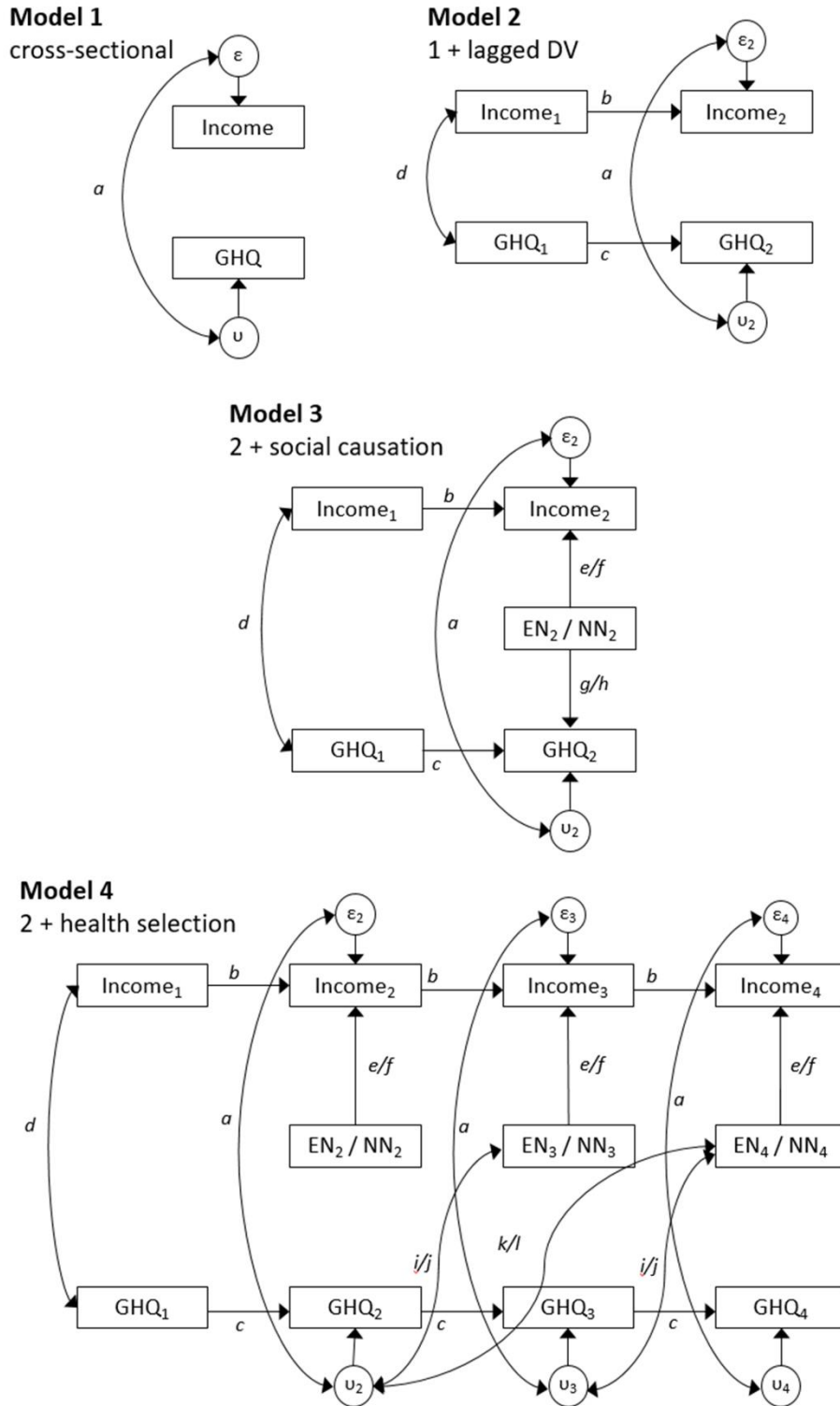
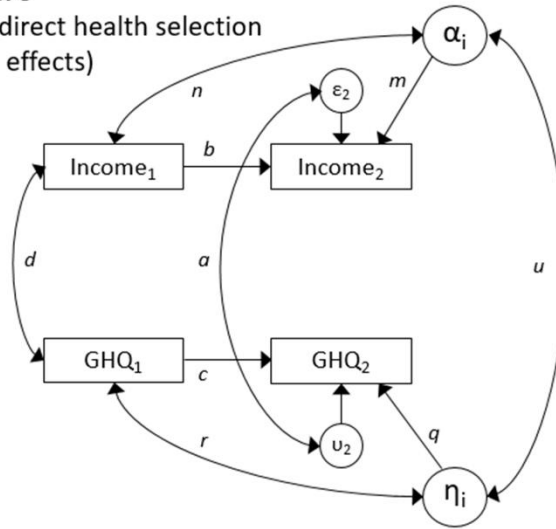


Figure 2 Longitudinal structural equation models (Models 1-4; Figure 3 shows 5 and 6)
 Note: minimum necessary number of waves shown. Employment transition indicators *EN* and *NN* combined for clarity; in the actual models they are separate (hence the two paths represented in one line by *e/f* etc.). Correlations between the two within and across time, and within each over time, are estimated but not shown here. Error variances and all paths except *o*, *p*, *s*, and *t* are constrained to be equal over time.

Model 5

2 + indirect health selection
(fixed effects)



Model 6 full model

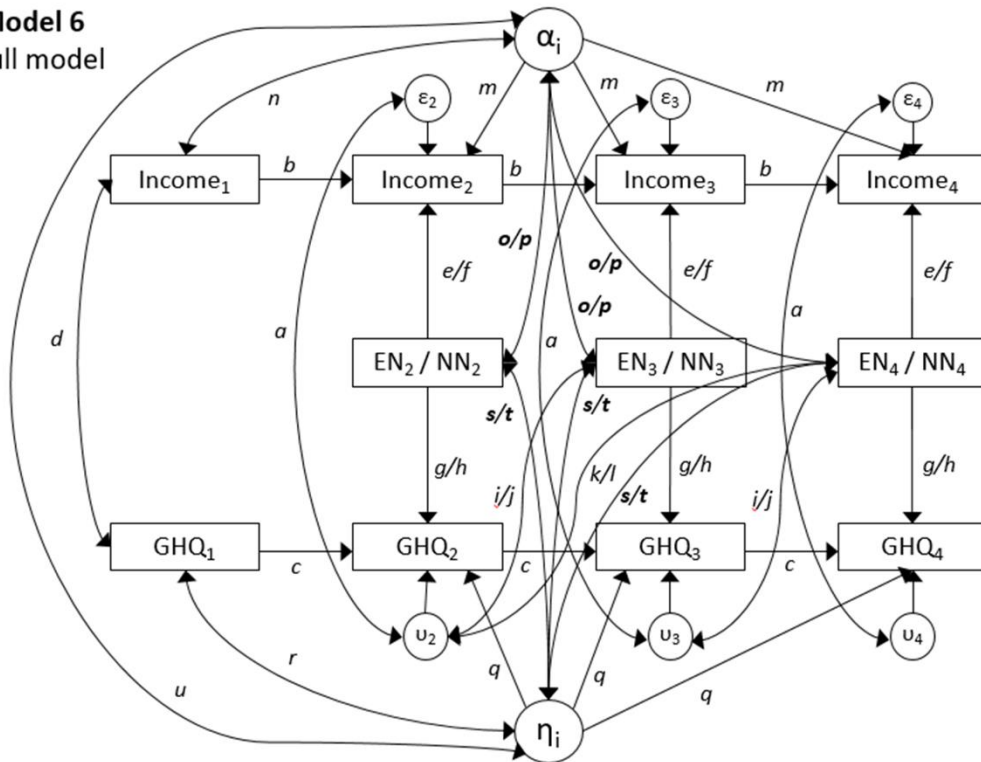


Figure 3 Longitudinal structural equation models (Models 5 and 6; Figure 2 shows 1-4)
Note: see note to Figure 2.

Model 3 adds the $t-1$ to t employment transition indicators as causes of both income (paths e and f) and mental health at t . The paths g and h from employment transitions to mental health represent the mechanism of social causation. For example, $EN_t \rightarrow GHQ_t$ represents the effect of moving out of employment between $t-1$ and t on mental health at t . Estimating the effect of employment transitions rather than the (lagged or contemporaneous) effect of economic status has some advantages for causal inference. Recent work has highlighted that ‘for continuously varying states ... relying on the temporal ordering of the data can be much worse than useless’ for determining causal relationships (Leszczensky & Wolbring, 2019; Vaisey & Miles, 2017: 63), requiring a condition which in many contexts it is impossible to know if one has met: that ‘the lags between [observations] match the real-world causal lags in the processes under study’ (Vaisey & Miles, 2017: 64). In addition, EN_t , for instance, captures both the effect of an event with a clear temporal position – between $t-1$ and t – and the likely synchronous influence of economic status on both income and mental health. Life *events* are prominent in models of mental health problems as potential triggers of onset (Brown & Harris, 1978; Goldberg & Huxley, 1992). By contrast, modelling just the effect of contemporaneous economic status (conditional on lagged economic status) would give an estimate that aggregates two substantively distinct groups: those who leave employment, and those who remain non-employed.

Leaving employment or remaining non-employed may of course be caused by a deterioration in mental health. Model 4 tests the health selection hypothesis. It includes employment transitions as determinants of income as in Model 3, but not of mental health (paths g and h are constrained to zero). Health selection is modelled by allowing covariances (i , j , k , and l) between the error terms for GHQ at $t-1$ and $t-2$ and the employment transition indicators at t . (Covariances representing selection at all greater lags are also estimated but not reported). This parametrisation enables estimation of Model 6 (see below), but a more straightforward version – in which there are directed paths from GHQ at $t-1$ and $t-2$ to each of the employment transition indicators at t – gives almost identical results (Model 4a, reported in the Appendix).

As a test of the indirect health selection hypothesis, according to which stable, unobserved differences between individuals generate an association between income and mental health, while the two do not causally affect one another, Model 5 adds individual fixed effects to Model 2 (α_i for income and η_i for GHQ). Under OLS, doing so introduces bias: whether the fixed effects are removed by demeaning or first-differencing, a correlation arises between the

transformed lagged dependent variable and the transformed error term (Angrist & Pischke, 2009: 245; Halaby, 2004: 539; Nickell, 1981). However, recent work has demonstrated the feasibility of maximum-likelihood estimation of models including both fixed effects and lagged dependent variables (Allison et al., 2017; Moral-Benito, 2013). This is accomplished within a SEM framework by modelling the fixed effects as latent variables allowed to correlate freely with the time-varying covariates (Bollen & Brand, 2010; Mundlak, 1978; Teachman et al., 2001). The time-invariant effects of α_i and η_i are represented by m and q ; n , r , and the vectors \mathbf{o} , \mathbf{p} , \mathbf{s} , and \mathbf{t} allow unrestricted correlations with the time-varying covariates. The two fixed effects are not assumed to be uncorrelated (path u).

Finally, Model 6 includes all these paths. Given prior theory and evidence in favour of all three (not mutually exclusive) hypotheses, this model is the most plausible from which to draw estimates of the magnitude of each parameter. Further, it explores the extent to which growth in mental health inequality can be accounted for by the combination of the three mechanisms.

Conceptually, both employment transitions and mental health are dependent variables in this full model (under the health selection and social causation hypotheses respectively). For simplicity I only refer to the *EN* transition:

$$GHQ_t = cGHQ_{t-1} + gEN_t + \eta_i + v_{it} \quad (1)$$

$$EN_t = \beta_1 GHQ_{t-1} + \beta_2 GHQ_{t-2} + v_{it} \quad (2)$$

However, modelling these simultaneously as causal paths violates the assumption of strict exogeneity – that the disturbance terms are mean independent of all past, current, and future values of the independent variables (Halaby, 2004). For instance, under this system of equations, GHQ_{t-1} in (2) is determined by EN_{t-1} and therefore may correlate with its idiosyncratic component v_{it-1} . Drawing on work developing a cross-lagged panel model with fixed effects estimated by maximum likelihood within a SEM framework (ML-SEM) (Allison et al., 2017; Williams et al., 2018), I instead allow for health selection in the full model by eliminating (2) and including the covariances i , j , k , and l . This represents an assumption that employment transitions are not strictly exogenous but rather sequentially exogenous or predetermined: ‘independent of all future values of [the time-varying error] but may be

correlated with past values of [it]' (Allison et al., 2017: 3); that is, prior shocks to mental health may affect subsequent employment transitions. In a more general context, independent simulation results indicate that this approach deals well with reverse causality under a range of scenarios (Leszczensky & Wolbring, 2019). It should be noted that the ML-SEM approach assumes no serial correlation in the error terms.

In interpreting paths g and h in Model 6 from $(t-1$ to $t)$ employment transitions to GHQ (at t) as effects of the former on the latter, I make the assumption that the employment transition is not caused by a change in GHQ at some point *between* $t-1$ and t . In support of this is the intuition that health selection effects are likely to be 'slow' relative to social causation effects. Whereas mood is susceptible to immediate influence, and income – though not determined wholly by employment – declines mechanically as a function of leaving employment, it may take some time for deteriorating mental health to reach the point where an individual opts or is forced to leave employment; this is likely to be a complex and relatively gradual social process.

Model 7, discussed and presented in the Appendix, offers an alternative specification for comparison: it is the same as Model 6 but without the lagged dependent variables, i.e. paths b and c are constrained to zero. Results are similar to Model 6 and do not alter the conclusions.

The third stage of the analysis is to run the six models described above, now stratified by sex and age group, and examine differences in the estimated parameters. Defining age groups is not straightforward in a longitudinal context: doing so by age-at-observation splits individuals between age groups. For example, to define an age group as all person-years between the ages of 30 and 40 would be to discard half the information from an individual followed from age 37 to age 44. To preserve the 'length' of the panel, I instead define age groups with reference to an individual's age in the middle of the observation period: that is (arbitrarily), their age at wave 4. To illustrate, the age group 30 to 40 now no longer 'cuts in half' the individual who is followed from age 37 to age 44, but includes them, using all eight observations. The label '30 to 40' is, correspondingly, not strictly accurate or entirely intuitive, and includes some observations of individuals when they are aged below 30 or over 40; however it maximises the use of information and most person-years do fall within the nominal age range. The age groups I use are 16-32, 33-49, and 50-66.

5 Results

5.1 Cross-sectional associations

Cross-sectional inequality in mental health at each age is shown in Figure 4. The pattern is generally similar among men (panel (a)) and women (panel (b)). The association between household income and GHQ score is close to zero at the youngest ages, gradually increasing until it becomes statistically significant for those aged 25, among both men and women. With a few exceptions, significant inequalities are evident at all ages past this. The strength of the association appears to increase steadily from its tentative emergence in the early 20s, up to the early 30s. From this age until the mid-50s there is volatility, especially among men, but in general the coefficients remain stable around -0.15 to -0.20 . From the mid-50s to age 70 the association declines steadily, becoming nonsignificant among women and nearly so among men. Even at their largest, around -0.22 , these correlations represent only modest associations. These descriptive findings offer cautious support for the idea that health inequalities are generated in the transition to adulthood.

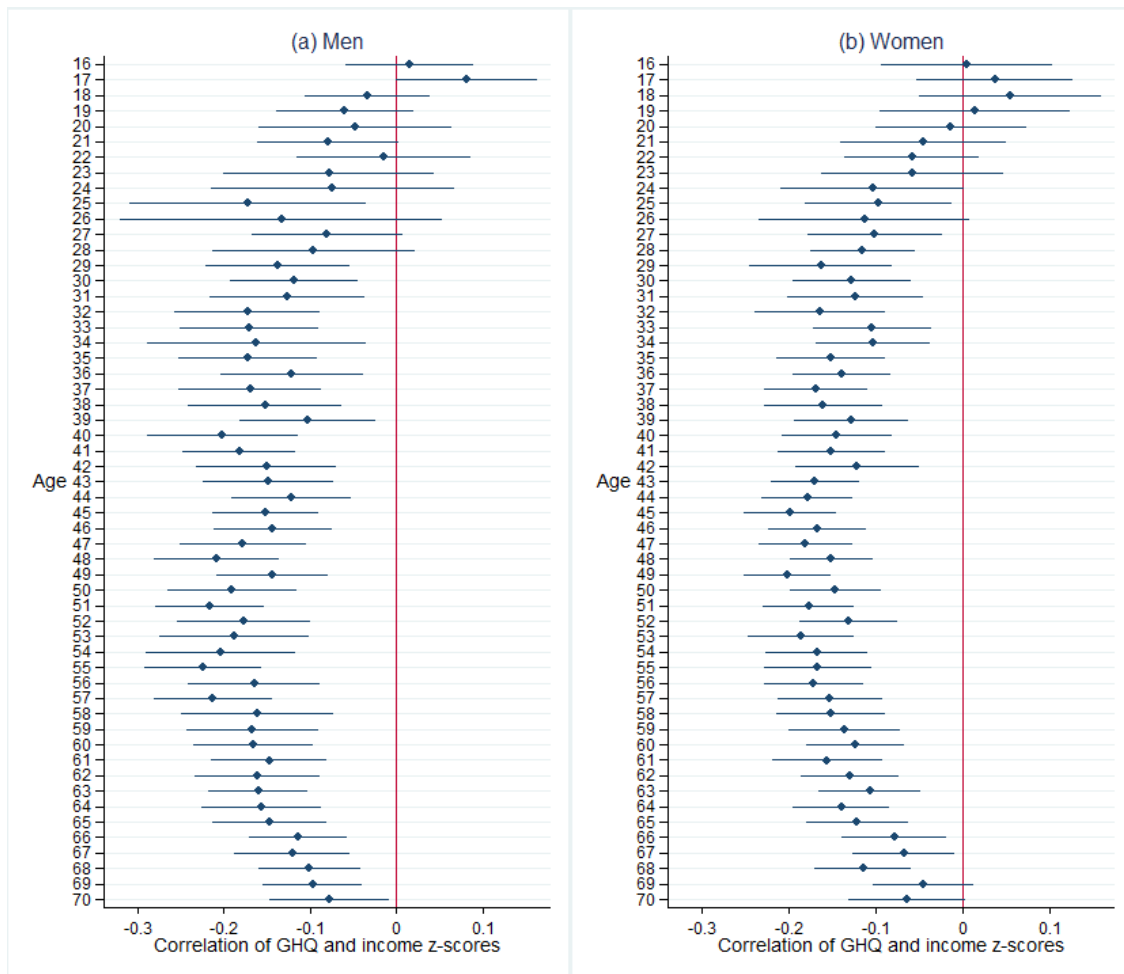


Figure 4 Mental health inequality by age and sex

Note: the correlations shown correspond to path *a* in Figure 1a. Higher GHQ score indicates worse mental health.

Figure 5 shows the same correlations, but with economic status (employed versus not employed) included in the model as in Figure 1b. Conditional on economic status the associations are substantially reduced, falling mostly in the range -0.05 to -0.10 and often not statistically significant among men (for whom the estimates are less precise due to a lower response rate). The association between income and depressive symptoms is largely accounted for by whether or not individuals are in work.

A simple explanation for the pattern observed in Figure 4 could be that the association varies as a function of the proportion of individuals at a given age who live with their parents. An individual's depressive symptoms are likely to be more closely related to their household income, the greater the extent to which that individual determines their household income – for instance by being a member of a smaller household. Where individuals live with their parents,

the latter are likely to primarily determine household income, and the two variables may be relatively ‘insulated’ from one another. Figure S2 in the Appendix replicates Figure 4, separately for men and women who do and do not live with their parents. The association is weaker when individuals live with their parents, but at age 22 (and up to 24 for men), there is no evidence for the inequality having emerged even among individuals who no longer live with their parents.

5.2 Longitudinal SEMs

Before contrasting results between men and women of different ages, I first present longitudinal results for the aggregated sample, shown in Table 1. *a* from Model 1 represents the synchronous association between GHQ score and household income before any other elements are added to the model. At -0.139 , this is in line with the correlations in Figure 4, though the latter highlights the heterogeneity concealed by aggregating the sample.

Model 2 adds the autoregressive paths *b* and *c*. These show large coefficients (0.717 and 0.551 respectively), reflecting a high degree of persistence in both variables and household income in particular. Though over half of the synchronous mental health inequality (path *a*) is accounted for by the association between these variables in the previous year, what remains is substantial (-0.060) and represents the year-to-year growth in the social gradient. This does not accumulate arithmetically because the autoregressive coefficients are less than 1; that is, individuals’ GHQ and income tend to regress to the mean, attenuating the association between the two and offsetting part of the year-to-year growth.

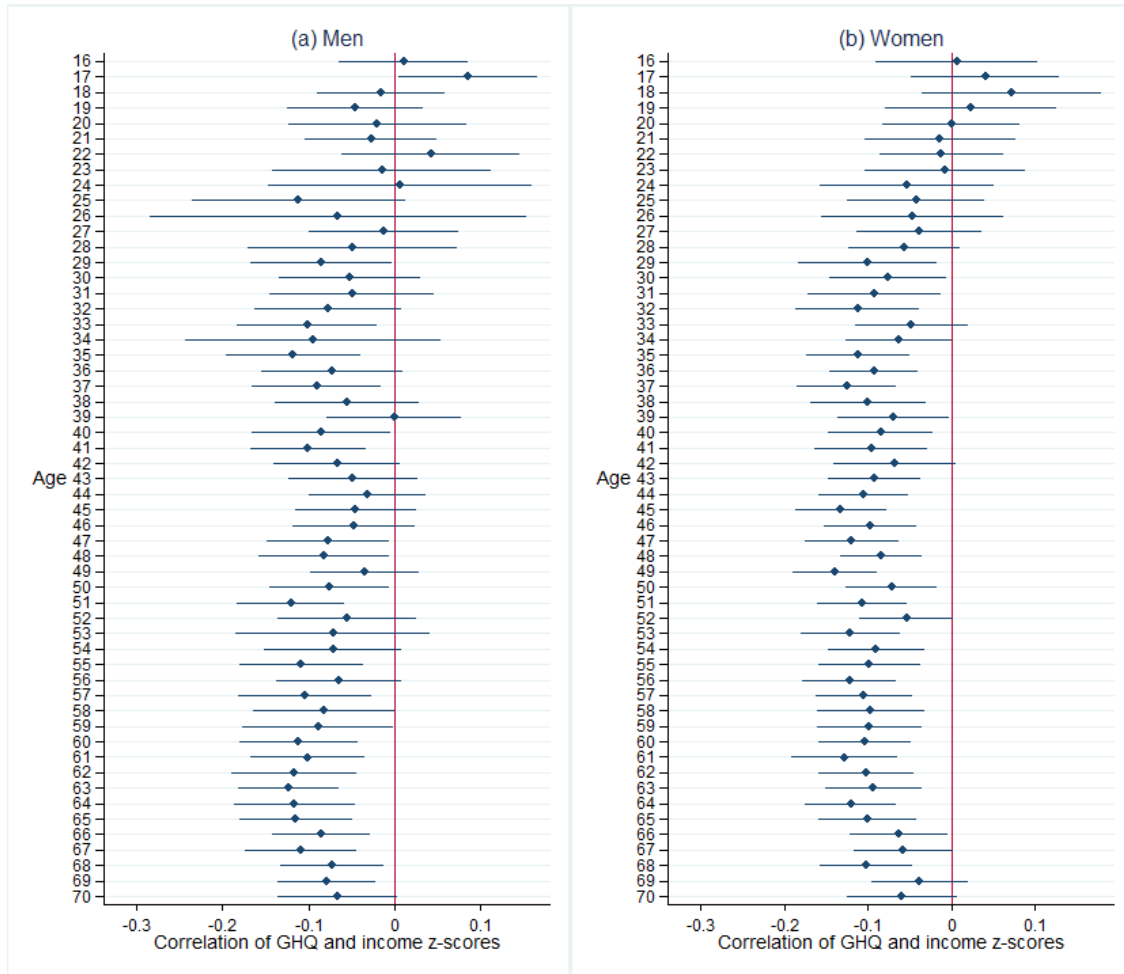


Figure 5 Mental health inequality conditional on economic status, by age and sex
 Note: the correlations shown correspond to path *a* in Figure 1b. Higher GHQ score indicates worse mental health.

Model 3 introduces employment transitions. It is the first to achieve acceptable model fit by any measure, if only for RMSEA (0.070) (Little, 2013). As expected, moving from employed to non-employed is negatively associated with household income (path *e*; -0.639), and those non-employed in consecutive waves have lower incomes than those employed at *t* (path *f*; -0.233). These two indicators are also associated with worse mental health relative to individuals employed at *t* (paths *g* and *h*; 0.154 , 0.161). Since GHQ is standardised and the population standard deviation is estimated at close to 5.5, each of these transitions is associated on average with close to one additional point on the 0-36 GHQ scale.

Model 4 adds covariance paths representing health selection to the baseline model. The unexplained portion of GHQ at both *t*-2 and *t*-1 covaries significantly with subsequent moves out of employment (paths *i* and *k*). In the alternative specification (Model 4a, see Table S3 in

the Appendix) some health selection is also evident, but in this case the association is between prior GHQ and remaining out of employment. Paths a - f are negligibly different in Model 4a compared to Model 4, while the latter shows clearly better fit.

Moving from Model 2 to Models 3 and 4, path a is reduced from -0.060 to -0.047 in both cases. Initially, this would appear to suggest that mechanisms of social causation and health selection each account for about a fifth of the year-to-year growth in inequality when looking at the aggregated sample. However, this reduction in path a in fact results simply from the addition of paths e and f (see Model 2a in Table S3); that is, from introducing employment transitions as determinants of income only (and unrelated to mental health). This addition makes the income residual ε conditional on one's recent employment history. This result indicates that employment plays a role in explaining mental health inequality. However, paths e and f and the consequences of their addition cannot be 'claimed' by any of the hypotheses being tested. This leads to the surprising result that social causation and health selection as operationalised here, while both evident, do not themselves appear to account for any of the year-to-year growth in the social gradient.

Table 1 Longitudinal SEM path estimates and model fit statistics, aggregated sample

	Model	1	2	3	4	5	6
<i>a</i>	$\varepsilon_t \leftrightarrow u_t$	-0.139 ***	-0.060 ***	-0.047 ***	-0.047 ***	-0.038 ***	-0.029 ***
<i>b</i>	$\text{Income}_{t-1} \rightarrow \text{income}_t$		0.717 ***	0.694 ***	0.694 ***	0.325 ***	0.303 ***
<i>c</i>	$\text{GHQ}_{t-1} \rightarrow \text{GHQ}_t$		0.551 ***	0.543 ***	0.552 ***	0.160 ***	0.160 ***
<i>d</i>	$\text{Income}_1 \leftrightarrow \text{GHQ}_1$		-0.147 ***	-0.144 ***	-0.147 ***	-0.004	-0.030 ***
<i>e</i>	$\text{EN}_t \rightarrow \text{income}_t$			-0.639 ***	-0.634 ***		-0.636 ***
<i>f</i>	$\text{NN}_t \rightarrow \text{income}_t$			-0.233 ***	-0.227 ***		-0.399 ***
<i>g</i>	$\text{EN}_t \rightarrow \text{GHQ}_t$			0.154 ***			0.202 ***
<i>h</i>	$\text{NN}_t \rightarrow \text{GHQ}_t$			0.161 ***			0.126 ***
<i>i</i>	$u_{t-1} \leftrightarrow \text{EN}_t$				0.026 ***		0.039 ***
<i>j</i>	$u_{t-1} \leftrightarrow \text{NN}_t$				0.000		-0.004
<i>k</i>	$u_{t-2} \leftrightarrow \text{EN}_t$				0.024 ***		0.030 ***
<i>l</i>	$u_{t-2} \leftrightarrow \text{NN}_t$				0.013 ***		0.010
	Fixed effects	No	No	No	No	Yes	Yes
	RMSEA	0.274	0.108	0.070	0.072	0.035	0.030
	AIC	484512	319345	269260	308645	364836	257240
	CFI	0.016	0.835	0.821	0.820	0.984	0.971
	TLI	0.122	0.847	0.838	0.832	0.984	0.970
	SRMR	0.370	0.179	0.106	0.106	0.068	0.066

Note: *a*, *b*, *c* etc. refer to the paths shown in Figures 2 and 3. Paths with double-headed arrows represent covariances, reported here as correlations. N(individuals) = 17,843. Fit statistics from models without adjustment for survey design or attrition (unavailable otherwise).

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Model 5 adds individual fixed effects to the equations for income and mental health, accounting for the effects of any time-invariant unobserved characteristics of individuals, on the assumption that those effects do not vary over time. 51% of individuals in the sample are continuously employed across the eight waves and thus show no variation in the employment transition variables.³ 26% of individuals experience at least one transition out of employment (*EN*), while 39% have at least one observation at which they are out of employment at both $t-1$ and t (*NN*). Unobserved characteristics of the sort highlighted by the indirect health selection hypothesis appear to play an important role. The residual association *a* is substantially reduced compared with Model 2, at -0.038 . As one might expect, the fixed effects also account for a large portion of the persistence in income and mental health, attenuating paths *b* and *c*.

³ N does not change with the introduction of fixed effects because estimation is not achieved by differencing out the fixed effects or time-demeaning the data. Rather, fixed effects estimation is achieved by estimating (rather than constraining to zero as in a random effects model) the correlations between the latent variables (α_i and η_i) which represent the fixed effects, and the time-varying covariates (Allison et al., 2017; Bollen & Brand, 2010; Teachman et al., 2001).

Model 6 allows all the paths previously described. The unexplained synchronous association between income and depressive symptoms, a , is reduced to -0.029 (as above, the further reduction is most likely due to the inclusion of paths e and f), but not however completely accounted for. Estimates for the paths corresponding to social causation and health selection are generally similar to those from earlier models. g , the effect of moving out of employment on GHQ, is around a third larger (0.202), and the selection effects associated with this transition are also larger than previously. Since the coefficient for NN is somewhat reduced (0.126), the two transition variables here clearly differ in their effects on depressive symptoms. Model 6 shows the best (and good) fit (RMSEA 0.030, CFI 0.971, TLI 0.970, SRMR 0.066) (Hu & Bentler, 1999).

The residual correlation between income and depressive symptoms net of the values of those variables from a year earlier (and the other influences in the model) is thus substantially smaller a) when income is made conditional on employment transition status, and b) when only within-individuals variation is considered. The former does not discriminate between the hypothesised mechanisms. There is minimal support for social causation and health selection in explaining growth in the association between income and mental health problems. However there is support for the indirect health selection hypothesis, which highlights between-individual differences from early life as important determinants of unequal SES and mental health trajectories.

5.3 Variation by age group

I now examine how these results vary when looking at men and women of different ages. I first present the disaggregated results from Model 6 (Table 2) and then examine change in path a across models for the different groups (Table 3).

Table 2 Longitudinal SEM path estimates for Model 6, by sex and age group

		Men			Women		
		16-32	33-49	50-66	16-32	33-49	50-66
<i>a</i>	$\varepsilon_t \leftrightarrow u_t$	-0.022	-0.031 **	-0.026 **	-0.011	-0.051 ***	-0.018 *
<i>b</i>	$\text{Income}_{t-1} \rightarrow \text{income}_t$	0.356 ***	0.251 ***	0.301 ***	0.340 ***	0.267 ***	0.311 ***
<i>c</i>	$\text{GHQ}_{t-1} \rightarrow \text{GHQ}_t$	0.135 ***	0.138 ***	0.168 ***	0.173 ***	0.145 ***	0.168 ***
<i>d</i>	$\text{Income}_1 \leftrightarrow \text{GHQ}_1$	-0.005	-0.098 ***	-0.026	-0.046	-0.030 *	-0.017
<i>e</i>	$\text{EN}_t \rightarrow \text{income}_t$	-0.687 ***	-1.120 ***	-0.830 ***	-0.371 ***	-0.453 ***	-0.580 ***
<i>f</i>	$\text{NN}_t \rightarrow \text{income}_t$	-0.458 ***	-0.597 ***	-0.473 ***	-0.353 ***	-0.357 ***	-0.324 ***
<i>g</i>	$\text{EN}_t \rightarrow \text{GHQ}_t$	0.241 ***	0.513 ***	0.261 ***	0.070	0.250 ***	0.122
<i>h</i>	$\text{NN}_t \rightarrow \text{GHQ}_t$	0.125 **	0.473 ***	0.172 **	0.043	0.200 ***	0.115 **
<i>i</i>	$u_{t-1} \leftrightarrow \text{EN}_t$	0.023	0.057 ***	0.074 ***	0.001	0.030 **	0.037 ***
<i>j</i>	$u_{t-1} \leftrightarrow \text{NN}_t$	-0.020 *	0.008	-0.009 *	-0.006	0.005	0.004
<i>k</i>	$u_{t-2} \leftrightarrow \text{EN}_t$	0.001	0.040 **	0.055 ***	0.014	0.018	-0.022
<i>l</i>	$u_{t-2} \leftrightarrow \text{NN}_t$	-0.021	0.032	0.020	-0.009	0.014	0.011
N(individuals)		1461	2851	3143	2144	4114	4130

Note: see note to Table 1.

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

What do the disaggregated Model 6 results tell us about the strength of social causation and health selection across different lifecourse stages for men and women? Moving out of employment has a greater impact on household income (*e*) among men and particularly those in the 33-49 group (−1.120). The impact on GHQ (*g*) is also largest for men (0.513) and women (0.250) in this age group, and generally larger for men than women across all age groups. Path *h* too is greatest for the 33-49 age group among both men (0.473) and women (0.200). These results confirm the prediction that social causation would be especially strong in mid-adulthood. By contrast, there is little evidence for health selection in the transition to adulthood: selection out of employment (*i* and *k*) is evident only among the older two age groups, while the significant correlation *j* for younger men is not in the expected direction.

How does the addition of processes of social causation, selection, and the introduction of fixed effects differentially affect the parameter *a* for the year-to-year growth of inequality in depressive symptoms? Table 3 presents estimates of *a* for each model by sex and age group.

Table 3 Path a across Models 1-6, by sex and age group

	Men			Women		
	16-32	33-49	50-66	16-32	33-49	50-66
Model 1	-0.092 ***	-0.149 ***	-0.171 ***	-0.093 ***	-0.161 ***	-0.129 ***
Model 2	-0.052 ***	-0.070 ***	-0.066 ***	-0.038 ***	-0.085 ***	-0.037 ***
Model 3	-0.031 *	-0.034 **	-0.061 ***	-0.021 *	-0.065 ***	-0.035 ***
Model 4	-0.029 *	-0.035 **	-0.050 ***	-0.019 *	-0.065 ***	-0.034 ***
Model 5	-0.043 **	-0.057 ***	-0.041 ***	-0.015	-0.060 ***	-0.019 *
Model 6	-0.022	-0.031 **	-0.026 **	-0.011	-0.051 ***	-0.018 *

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

The results for Model 1 again reflect the findings presented in Figure 4: the cross-sectional association between household income and depressive symptoms is larger in the older age groups, and similar for men and women of the same ages. As described above, Model 2 adds the autoregressive parameter and thus changes the interpretation of a . This produces somewhat surprising results for the 50-66 age group: a , though reduced to a greater degree than among the other age groups, remains substantial in magnitude and negative, suggesting that the gradient continues to grow across these ages when in fact it attenuates.

a is generally substantially smaller in Models 3 and 4 compared with Model 2. As above, this reduction is mainly due to the addition of paths e and f , from employment transitions to income (see Model 2a in Table S4). Model 5, which includes fixed effects for income and mental health, shows a pattern of reduction in a compared with Model 2 which varies across groups. The reduction is most pronounced among the 50-66 age group and among women in the 16-32 age group.

In the full model (6), a is generally reduced by more than 50% compared with Model 2. An exception to the pattern is women aged 33-49, for whom a remains relatively large in magnitude (-0.051 in Model 6, -0.085 in Model 2). This aside, differences between groups in Model 6 are rather small. What remains of a in Model 6 represents the synchronous residual association of income and GHQ, conditional on the previous year's values and individuals' means of those variables; what might account for this is discussed below.

6 Discussion

The aim of this paper was to use nationally representative longitudinal UK data to explore when and how social inequality in mental health is generated. This inequality is operationalised here as the association between an individual's household income and depressive symptoms. Previous work has generally focused on testing for the *presence* of two mechanisms which each contribute to the formation of this inequality – social causation and health selection. The contribution of this paper is to additionally explore the extent to which they *actually contribute* to the generation of this inequality, and at what stages of the adult lifecourse they may do so. Further, in light of past research, this paper uses a method that adjusts for unobserved heterogeneity – enabling direct testing of the indirect health selection hypothesis – and focuses on employment transitions due to their close links with health.

In line with past research on social inequality in mental health, descriptive results showed the absence of an association between household income and depressive symptoms at age 16. Looking at person-year observations at increasing ages, the gradient becomes clearly evident by the mid-20s, continues to grow into the mid-30s, and stays relatively stable, before declining approximately through the mid-50s to 70. The same picture emerges for men and women.

Some previous work had found the association between income and depressive symptoms to be substantially rather small (and/or statistically nonsignificant) conditional on economic status (Hounkpatin et al., 2015; Oswald & Powdthavee, 2008). Adding an indicator for being employed to the initial cross-sectional model attenuated the association to a substantial degree. Accordingly, the longitudinal analysis focused on employment transitions as drivers of the social gradient in mental health.

A series of longitudinal SEMs was fit to test the extent to which the social causation, health selection, and indirect health selection mechanisms explain growth in mental health inequality – that is, the synchronous association between income and mental health conditional on the values of those variables in the previous year. Social causation was operationalised as the impact of leaving employment or remaining non-employed over the course of a year, health selection as the impact of shocks to mental health on employment transitions over the following year(s), and indirect health selection as stable unobserved differences between individuals, modelled with individual fixed effects.

Theory and past research suggested that social causation would be especially important across mid-adulthood, while individuals would be particularly susceptible to health selection during the transition to adulthood. With regard to whether these mechanisms were evident, and the magnitude of effects of employment transitions on mental health and vice versa, the full model (including individual fixed effects) yielded support only for the prediction that social causation would be particularly strong in mid-adulthood. For individuals in the age range 33-49, being non-employed or, particularly, leaving employment, had a relatively large impact both on income and mental health. For men in this age group, leaving employment was associated with a substantial increase in symptoms of depression and anxiety relative to the same individual when in employment: the coefficient 0.513 is equivalent to approximately three points on the 0-36 GHQ scale. With regard to health selection, there were modest associations between negative shocks to mental health and leaving employment over the following year, but these were only evident among the mid-aged and older groups in the sample. An important caveat is that conclusions about differences in the strength of the different processes across lifecourse stages rely on an assumption of no cohort or period effects.

With regard to whether and for which groups the three mechanisms explain growth in the association between income and mental health, there was surprisingly minimal support for either social causation or health selection, for men or women in any of the age groups. Apparent reductions upon introducing these mechanisms into the model in fact owed to the inclusion of paths from employment transitions to income. On the other hand, indirect health selection explained a modest portion of this growth – just over a third across the whole sample.

These results hint at the importance of the relatively neglected indirect health selection hypothesis. A role for stable unobserved differences between individuals observed in adulthood points to the complex range of genetic and early environmental factors which influence individuals' socioeconomic and wellbeing trajectories. However, the finding that neither social causation nor health selection contribute to generation of the social gradient in mental health through adulthood is paradoxical. There is strong prior evidence that employment transitions both affect and are affected by mental health, and these associations were clearly evident here too. Given this, one would expect these mechanisms to act to increase the cross-sectional association between income and mental health. It may be that these processes are more complex than the data and methods applied here can capture.

A full model including all three hypothesised pathways did not completely account for the observed increases in mental health inequality. This residual association could arise from a combination of some kind of time-varying confounding (for instance, changes in marital status (Sbarra et al., 2014)), a causal effect of income on depressive symptoms (Hounkpatin et al., 2015) and, less plausibly, a relatively ‘fast’ causal effect of depressive symptoms on income.

The hypothesis that social inequality in mental health would increase over the course of the transition to adulthood was derived in part from earlier work which had reported ‘equalisation’ in adolescence. The equalisation hypothesis has been the subject of methodological debate. The issues raised are worth briefly reviewing for their possible relevance to the finding reported here – that indeed there is relative equality in mental health across household income levels in late adolescence. Criticism of the evidence for the equalisation hypothesis has centred on using occupation-based measures of SES to assess health inequalities in adolescence (Judge & Benzeval, 1993; Macintyre & West, 1991). One approach has been to use geographical deprivation indices; this has yielded further evidence of equalisation with respect to mortality and self-esteem, the latter outcome being the closest conceptually to mental health which has been examined in this way (Dibben & Popham, 2013; Fagg et al., 2013; Green, 2013). Others have continued to focus on the family but argued for household income as preferable to occupation-based measures. ‘Important tasks, therefore’, write Emerson et al., ‘are to examine the negative patterning of health ... using measures of [SES] which are inclusive and equivalent ... Household income provides such a measure’ (2006: 354). These authors also note that evidence for the equalisation hypothesis ‘appears strongest ... for some rather than all dimensions of young people’s health, including ... self-reported symptoms of psychological distress’ (*ibid.*). It is perfectly plausible that the claim of equalisation varies in strength across health outcomes; this is not grounds for a wholesale rejection.

One limitation of this paper is that the relative strength of social causation and health selection are not *directly* compared in terms of their magnitude, as for instance in the work of Chandola et al. (2003). This reflects the operationalisation of the hypotheses through employment transitions, which are necessarily modelled as binary variables. Nevertheless, the approach does allow a comparison of the two hypotheses with respect to their contribution to explaining growth in the association between income and mental health.

A second limitation is the lack of specificity in the category ‘not employed’, and consequently in the meaning of the employment transition variables. Further analysis shows that for men aged 16-49, *EN* represents a transition into unemployment in most cases, while for women in these age groups the most common non-employment state is maternity leave or looking after the home or family; among 50-66 year-olds of both sexes retirement is the most common transition represented by *EN*. The approach taken here is the consequence of a trade-off: specifying more categories leads to an exponentially greater number of transition indicators. Some convergence problems occurred even when including the *NE* transition to differentiate between individuals employed at both $t-1$ and t , and those employed at t but not $t-1$. The evidence of Steele et al. (2013) does at least indicate that for men (women were not analysed), effects on GHQ are in the same direction when contrasting transitions out of employment and into unemployment (2.22) or into economic inactivity (0.59), and of very similar magnitude when looking at individuals who are unemployed (0.40) or economically inactive (0.45) in successive waves. Further evidence from the BHPS indicates that transitioning into retirement is associated with a (nonsignificant) increased risk of ‘caseness’ under a dichotomised operationalisation of the GHQ (OR 1.23 for men, 1.41 for women), while transitioning from employment into long term illness (ORs 3.16, 2.09), maternity leave (1.52, women only), and family care (1.56, women only) were all significantly associated with a greater risk of high GHQ score (Thomas et al., 2005). This suggests that leaving employment is experienced as a *loss of benefits* (Anaf et al., 2013), as well as an introduction to the challenges particular to the new economic status.

The measure also obscures variation within the category ‘employed’ – but recall Richards and Paskov’s (2016) finding of no social class gradient in mental health conditional on economic status. While gaining poor quality employment has been linked to worse physical health outcomes even than remaining unemployed, this conclusion did not hold for a measure of mental health (Chandola & Zhang, 2018). Further, a recent analysis of Understanding Society data finds that ‘even a small number of working hours (between one and [eight hours] a week) generates significant mental health and well-being benefits for previously unemployed or economically inactive individuals’ (Kamerāde et al., 2019: 1). On the positive side, this relatively low-resolution measurement of economic status and transitions therein allows for a methodological approach that is robust to unobserved heterogeneity, and to reverse causality acting over the course of a year or more.

The findings reported here reaffirm the importance of employment for depressive symptoms and suggest a partial redistribution of attention towards the relatively neglected – and not much elaborated – mechanism of indirect health selection. Past work has highlighted factors measured early in life which make large contributions to explaining social inequality in mental health in adulthood. For instance, teacher-rated ability across five domains at age 7 explained as much of the social gradient in mental health at age 33 as did job insecurity at age 33 (Power et al., 2002). Rather than consign early life influences and individual differences to ‘unobserved heterogeneity’, a priority for future work on social inequality in health should be to more clearly delineate what sorts of relatively fixed characteristics of individuals might matter, through what sorts of pathways, and how these characteristics are moulded across childhood and adolescence. At the same time, another clear conclusion is that individuals leaving work – for whatever reason – are vulnerable to deterioration in their mental health and should be supported where possible, and in particular to find new employment if desired.

7 References

- Aittomäki, A., Martikainen, P., Laaksonen, M., Lahelma, E., & Rahkonen, O. (2012). Household economic resources, labour-market advantage and health problems—A study on causal relationships using prospective register data. *Social Science & Medicine*, 75(7), 1303–1310. <https://doi.org/10.1016/j.socscimed.2012.05.015>
- Allison, P. D. (1990). Change Scores as Dependent Variables in Regression Analysis. *Sociological Methodology*, 20, 93–114.
- Allison, P. D., Williams, R., & Moral-Benito, E. (2017). Maximum Likelihood for Cross-lagged Panel Models with Fixed Effects. *Socius*, 3, 1–17. <https://doi.org/10.1177/2378023117710578>
- Alloy, L. B., Albright, J. S., Abramson, L. Y., & Dykman, B. M. (1990). Depressive Realism and Nondepressive Optimistic Illusions: The Role of the Self. In R. E. Ingram (Ed.), *Contemporary Psychological Approaches to Depression: Theory, Research, and Treatment* (pp. 71–86). Boston, MA: Springer US. https://doi.org/10.1007/978-1-4613-0649-8_6
- Anaf, J., Baum, F., Newman, L., Ziersch, A., & Jolley, G. (2013). The interplay between structure and agency in shaping the mental health consequences of job loss. *BMC Public Health*, 13, 110. <https://doi.org/10.1186/1471-2458-13-110>
- Anderson, T. W., & Hsiao, C. (1982). Formulation and estimation of dynamic models using panel data. *Journal of Econometrics*, 18(1), 47–82. [https://doi.org/10.1016/0304-4076\(82\)90095-1](https://doi.org/10.1016/0304-4076(82)90095-1)
- Aneshensel, C. S., Pearlin, L. I., Mullan, J. T., Zarit, S. H., & Whitlatch, C. J. (1995). *Profiles in caregiving: The unexpected career*. San Diego, CA, US: Academic Press.
- Angrist, J. D., & Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Woodstock: Princeton University Press.
- Blane, D., Harding, S., & Rosato, M. (1999). Does social mobility affect the size of the socioeconomic mortality differential?: Evidence from the Office for National Statistics Longitudinal Study. *Journal of the Royal Statistical Society. Series A, (Statistics in Society)*, 162(1), 59–70.
- Bollen, K. A., & Brand, J. E. (2010). A General Panel Model with Random and Fixed Effects: A Structural Equations Approach. *Social Forces*, 89(1), 1–34. <https://doi.org/10.1353/sof.2010.0072>
- Brand, J. E. (2015). The Far-Reaching Impact of Job Loss and Unemployment. *Annual Review of Sociology*, 41(1), 359–375. <https://doi.org/10.1146/annurev-soc-071913-043237>
- Brown, G. W., & Harris, T. O. (1978). *Social Origins of Depression*. London: Tavistock.
- Case, A., & Paxson, C. (2011). The Long Reach of Childhood Health and Circumstance: Evidence from the Whitehall II Study. *The Economic Journal*, 121(554), F183–F204.
- Chandola, T., Bartley, M., Sacker, A., Jenkinson, C., & Marmot, M. (2003). Health selection in the Whitehall II study, UK. *Social Science & Medicine*, 56(10), 2059–2072.
- Chandola, T., & Zhang, N. (2018). Re-employment, job quality, health and allostatic load biomarkers: Prospective evidence from the UK Household Longitudinal Study. *International Journal of Epidemiology*, 47(1), 47–57. <https://doi.org/10.1093/ije/dyx150>
- Clark, A. E., & Oswald, A. J. (1994). Unhappiness and Unemployment. *The Economic Journal*, 104(424), 648–659. <https://doi.org/10.2307/2234639>
- Cook, T. D., & Furstenberg, F. F. (2002). Explaining Aspects of the Transition to Adulthood in Italy, Sweden, Germany, and the United States: A Cross-Disciplinary, Case

- Synthesis Approach. *The Annals of the American Academy of Political and Social Science*, 580, 257–287.
- Costello, E. J., Compton, S. N., Keeler, G., & Angold, A. (2003). Relationships between poverty and psychopathology: A natural experiment. *Journal of the American Medical Association*, 290(15), 2023–2029.
- Dahl, E. (1993). Social inequality in health—The role of the healthy worker effect. *Social Science & Medicine*, 36(8), 1077–1086. [https://doi.org/10.1016/0277-9536\(93\)90126-O](https://doi.org/10.1016/0277-9536(93)90126-O)
- Department for Work and Pensions, & Department of Health. (2009). *Working our way to better mental health: A framework for action*. Norwich: Stationery Office.
- Dibben, C., & Popham, F. (2013). Are health inequalities evident at all ages? An ecological study of English mortality records. *European Journal of Public Health*, 23(1), 39–45. <https://doi.org/10.1093/eurpub/cks019>
- Dohrenwend, B. P., Levav, I., Shrout, P. E., Schwartz, S., Naveh, G., Link, B. G., ... Stueve, A. (1992). Socioeconomic Status and Psychiatric Disorders: The Causation-Selection Issue. *Science*, 255(5047), 946–952.
- Dooley, D., & Prause, J. (1997). Effect of Students' Self-esteem on Later Employment Status: Interactions of Self-esteem with Gender and Race. *Applied Psychology*, 46(2), 175–198. <https://doi.org/10.1111/j.1464-0597.1997.tb01223.x>
- Dooley, D., Prause, J., & Ham-Rowbottom, K. A. (2000). Underemployment and depression: Longitudinal relationships. *Journal of Health and Social Behavior*, 41(4), 421–436.
- Egan, M., Daly, M., & Delaney, L. (2016). Adolescent psychological distress, unemployment, and the Great Recession: Evidence from the National Longitudinal Study of Youth 1997. *Social Science & Medicine*, 156, 98–105. <https://doi.org/10.1016/j.socscimed.2016.03.013>
- Elovainio, M., Ferrie, J. E., Singh-Manoux, A., Shipley, M., Batty, G. D., Head, J., ... Kivimäki, M. (2011). Socioeconomic differences in cardiometabolic factors: Social causation or health-related selection? Evidence from the Whitehall II Cohort Study, 1991-2004. *American Journal of Epidemiology*, 174(7), 779–789. <https://doi.org/10.1093/aje/kwr149>
- Elstad, J. I. (2004). Health and Status Attainment: Effects of Health on Occupational Achievement among Employed Norwegian Men. *Acta Sociologica*, 47(2), 127–140. <https://doi.org/10.1177/0001699304043824>
- Emerson, E., Graham, H., & Hatton, C. (2006). Household income and health status in children and adolescents in Britain. *European Journal of Public Health*, 16(4), 354–360. <https://doi.org/10.1093/eurpub/cki200>
- Erikson, R., & Goldthorpe, J. H. (1992). *The constant flux: A study of class mobility in industrial societies*. Oxford: Clarendon Press.
- Fagg, J. H., Curtis, S. E., Cummins, S., Stansfeld, S. A., & Quesnel-Vallée, A. (2013). Neighbourhood deprivation and adolescent self-esteem: Exploration of the 'socio-economic equalisation in youth' hypothesis in Britain and Canada. *Social Science & Medicine*, 91, 168–177. <https://doi.org/10.1016/j.socscimed.2013.02.021>
- Finkel, S. E. (1995). *Causal analysis with panel data*. London: Sage.
- Fisher, P., Fumagalli, L., Buck, N., & Avram, S. (2019). *Understanding Society and its income data* (Understanding Society Working Paper Series No. 2019– 08). Colchester: University of Essex. Retrieved from <https://www.understandingsociety.ac.uk/sites/default/files/downloads/working-papers/2019-08.pdf>
- Foverskov, E., & Holm, A. (2016). Socioeconomic inequality in health in the British household panel: Tests of the social causation, health selection and the indirect selection

- hypothesis using dynamic fixed effects panel models. *Social Science & Medicine*, 150, 172–183. <https://doi.org/10.1016/j.socscimed.2015.12.021>
- Friedrich, M. J. (2017). Depression Is the Leading Cause of Disability Around the World. *JAMA*, 317(15), 1517–1517. <https://doi.org/10.1001/jama.2017.3826>
- Fryers, T., Melzer, D., & Jenkins, R. (2003). Social inequalities and the common mental disorders: A systematic review of the evidence. *Social Psychiatry and Psychiatric Epidemiology*, 38(5), 229–237. <https://doi.org/10.1007/s00127-003-0627-2>
- Fryers, T., Melzer, D., Jenkins, R., & Brugha, T. (2005). The distribution of the common mental disorders: Social inequalities in Europe. *Clinical Practice and Epidemiology in Mental Health*, 1, 14. <https://doi.org/10.1186/1745-0179-1-14>
- Furstenberg, F. F., Kennedy, S., McLoyd, V. C., Rumbaut, R. G., & Settersten, R. A. (2004). Growing Up is Harder to do. *Contexts*, 3(3), 33–41. <https://doi.org/10.1525/ctx.2004.3.3.33>
- Glendinning, A., Hendry, L., & Shucksmith, J. (1995). Lifestyle, health and social class in adolescence. *Social Science & Medicine*, 41(2), 235–248. [https://doi.org/10.1016/0277-9536\(94\)00316-1](https://doi.org/10.1016/0277-9536(94)00316-1)
- Goldberg, D. P., Gater, R., Sartorius, N., Ustun, T. B., Piccinelli, M., Gureje, O., & Rutter, C. (1997). The validity of two versions of the GHQ in the WHO study of mental illness in general health care. *Psychological Medicine*, 27(1), 191–197.
- Goldberg, D. P., & Huxley, P. (1992). *Common Mental Disorders: A Bio-Social Model*. London: Tavistock/Routledge.
- Gottfredson, L. S. (2004). Intelligence: Is it the epidemiologists’ elusive ‘fundamental cause’ of social class inequalities in health? *Journal of Personality and Social Psychology*, 86(1), 174–199. <https://doi.org/10.1037/0022-3514.86.1.174>
- Green, M. A. (2013). The equalisation hypothesis and changes in geographical inequalities of age based mortality in England, 2002–2004 to 2008–2010. *Social Science & Medicine*, 87, 93–98. <https://doi.org/10.1016/j.socscimed.2013.03.029>
- Halaby, C. N. (2004). Panel Models in Sociological Research: Theory into Practice. *Annual Review of Sociology*, 30(1), 507–544. <https://doi.org/10.1146/annurev.soc.30.012703.110629>
- Hammarström, A., & Janlert, U. (1997). Nervous and depressive symptoms in a longitudinal study of youth unemployment—Selection or exposure? *Journal of Adolescence*, 20(3), 293–305. <https://doi.org/10.1006/jado.1997.0086>
- Hoffmann, R., Kröger, H., & Pakpahan, E. (2018). Pathways between socioeconomic status and health: Does health selection or social causation dominate in Europe? *Advances in Life Course Research*, 36, 23–36. <https://doi.org/10.1016/j.alcr.2018.02.002>
- Houkpatin, H. O., Wood, A. M., Brown, G. D. A., & Dunn, G. (2015). Why does Income Relate to Depressive Symptoms? Testing the Income Rank Hypothesis Longitudinally. *Social Indicators Research*, 124(2), 637–655. <https://doi.org/10.1007/s11205-014-0795-3>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Jackson, C. (2007). The General Health Questionnaire. *Occupational Medicine*, 57(1), 79–79. <https://doi.org/10.1093/occmed/kql169>
- Judge, K., & Benzeval, M. (1993). Health inequalities: New concerns about the children of single mothers. *BMJ*, 306(6879), 677–680. <https://doi.org/10.1136/bmj.306.6879.677>
- Kameråde, D., Wang, S., Burchell, B., Balderson, S. U., & Coutts, A. (2019). A shorter working week for everyone: How much paid work is needed for mental health and well-

- being? *Social Science & Medicine*, Online first.
<https://doi.org/10.1016/j.socscimed.2019.06.006>
- Kessler, R. C. (2012). The Costs of Depression. *Psychiatric Clinics of North America*, 35(1), 1–14. <https://doi.org/10.1016/j.psc.2011.11.005>
- Kessler, R. C., Foster, C. L., Saunders, W. B., & Stang, P. E. (1995). Social Consequences of Psychiatric Disorders, I: Educational Attainment. *American Journal of Psychiatry*, 152(7), 1026–1032.
- Kessler, R. C., McLaughlin, K. A., Green, J. G., Gruber, M. J., Sampson, N. A., Zaslavsky, A. M., ... Williams, D. R. (2010). Childhood adversities and adult psychopathology in the WHO World Mental Health Surveys. *The British Journal of Psychiatry*, 197(5), 378–385. <https://doi.org/10.1192/bjp.bp.110.080499>
- Ki, M., Sacker, A., Kelly, Y., & Nazroo, J. (2011). Health selection operating between classes and across employment statuses. *Journal of Epidemiology & Community Health*, 65(12), 1132–1139. <https://doi.org/10.1136/jech.2009.107995>
- Klein, D. N. (2008). Classification of depressive disorders in the DSM-V: Proposal for a two-dimension system. *Journal of Abnormal Psychology*, 117(3), 552–560. <https://doi.org/10.1037/0021-843X.117.3.552>
- Knies, G. (Ed.). (2018). *Understanding Society: Waves 1-7, 2009-2016 and harmonised BHPS: Waves 1-18, 1991-2009, User Guide, July 2018*. Colchester: University of Essex.
- Köhler, S., Wiethoff, K., Ricken, R., Stamm, T., Baghai, T. C., Fisher, R., ... Adli, M. (2015). Characteristics and differences in treatment outcome of inpatients with chronic vs. Episodic major depressive disorders. *Journal of Affective Disorders*, 173, 126–133. <https://doi.org/10.1016/j.jad.2014.10.059>
- Kröger, H., Pakpahan, E., & Hoffmann, R. (2015). What causes health inequality? A systematic review on the relative importance of social causation and health selection. *The European Journal of Public Health*, 25(6), 951–960. <https://doi.org/10.1093/eurpub/ckv111>
- Krug, G., & Eberl, A. (2018). What explains the negative effect of unemployment on health? An analysis accounting for reverse causality. *Research in Social Stratification and Mobility*, 55, 25–39. <https://doi.org/10.1016/j.rssm.2018.03.001>
- Lachman, M. E. (2004). Development in midlife. *Annual Review of Psychology*, 55, 305–331. <https://doi.org/10.1146/annurev.psych.55.090902.141521>
- Lachman, M. E., Teshale, S., & Agrigoroaei, S. (2015). Midlife as a pivotal period in the life course: Balancing growth and decline at the crossroads of youth and old age. *International Journal of Behavioral Development*, 39(1), 20–31. <https://doi.org/10.1177/0165025414533223>
- Leszczensky, L., & Wolbring, T. (2019). How to Deal With Reverse Causality Using Panel Data? Recommendations for Researchers Based on a Simulation Study. *Sociological Methods & Research*, 0049124119882473. <https://doi.org/10.1177/0049124119882473>
- Link, B. G., Phelan, J. C., Miech, R., & Westin, E. L. (2008). The resources that matter: Fundamental social causes of health disparities and the challenge of intelligence. *Journal of Health and Social Behavior*, 49(1), 72–91. <https://doi.org/10.1177/002214650804900106>
- Little, T. D. (2013). *Longitudinal Structural Equation Modeling*. New York: The Guilford Press.
- Lorant, V., Delige, D., Eaton, W., Robert, A., Philippot, P., & Ansseau, M. (2003). Socioeconomic Inequalities in Depression: A Meta-Analysis. *American Journal of Epidemiology*, 157(2), 98–112.

- Lundin, A., Hallgren, M., Theobald, H., Hellgren, C., & Torgén, M. (2016). Validity of the 12-item version of the General Health Questionnaire in detecting depression in the general population. *Public Health*, 136, 66–74. <https://doi.org/10.1016/j.puhe.2016.03.005>
- Macintyre, S., & West, P. (1991). Lack of class variation in health in adolescence: An artefact of an occupational measure of social class? *Social Science & Medicine*, 32(4), 395–402. [https://doi.org/10.1016/0277-9536\(91\)90340-I](https://doi.org/10.1016/0277-9536(91)90340-I)
- McLaughlin, K. A., Green, J. G., Gruber, M. J., Sampson, N. A., Zaslavsky, A. M., & Kessler, R. C. (2012). Childhood Adversities and First Onset of Psychiatric Disorders in a National Sample of US Adolescents. *Archives of General Psychiatry*, 69(11), 1151–1160. <https://doi.org/10.1001/archgenpsychiatry.2011.2277>
- McLeod, J. D., & Fettes, D. L. (2007). Trajectories of Failure: The Educational Careers of Children with Mental Health Problems. *American Journal of Sociology*, 113(3), 653–701. <https://doi.org/10.1086/521849>
- Minton, J. W., Pickett, K. E., & Dorling, D. (2012). Health, employment, and economic change, 1973-2009: Repeated cross sectional study. *BMJ*, 344, e2316. <https://doi.org/10.1136/bmj.e2316>
- Moral-Benito, E. (2013). Likelihood-Based Estimation of Dynamic Panels With Predetermined Regressors. *Journal of Business & Economic Statistics*, 31(4), 451–472. <https://doi.org/10.1080/07350015.2013.818003>
- Mulatu, M. S., & Schooler, C. (2002). Causal Connections between Socio- Economic Status and Health: Reciprocal Effects and Mediating Mechanisms. *Journal of Health and Social Behavior*, 43(1), 22–41.
- Mundlak, Y. (1978). On the Pooling of Time Series and Cross Section Data. *Econometrica*, 46, 69–85.
- Murphy, J. A., & Byrne, G. J. (2012). Prevalence and correlates of the proposed DSM-5 diagnosis of Chronic Depressive Disorder. *Journal of Affective Disorders*, 139(2), 172–180. <https://doi.org/10.1016/j.jad.2012.01.033>
- Newman, K. S. (1988). *Falling from grace: The experience of downward mobility in the American middle class*. New York: Free Press.
- Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica*, 49, 1417–1426.
- Oswald, A. J., & Powdthavee, N. (2008). Death, Happiness, and the Calculation of Compensatory Damages. *The Journal of Legal Studies*, 37(S2), S217–S251. <https://doi.org/10.1086/595674>
- Power, C., Atherton, K., Strachan, D. P., Shepherd, P., Fuller, E., Davis, A., ... Stansfeld, S. A. (2007). Life-course influences on health in British adults: Effects of socio-economic position in childhood and adulthood. *International Journal of Epidemiology*, 36(3), 532–539.
- Power, C., Stansfeld, S. A., Matthews, S., Manor, O., & Hope, S. (2002). Childhood and adulthood risk factors for socio-economic differentials in psychological distress: Evidence from the 1958 British birth cohort. *Social Science & Medicine*, 55(11), 1989–2004. [https://doi.org/10.1016/S0277-9536\(01\)00325-2](https://doi.org/10.1016/S0277-9536(01)00325-2)
- Prause, J., & Dooley, D. (2001). Favourable Employment Status Change and Psychological Depression: A Two-year Follow-up Analysis of the National Longitudinal Survey of Youth. *Applied Psychology*, 50(2), 282–304. <https://doi.org/10.1111/1464-0597.00059>
- Reeves, A., McKee, M., Mackenbach, J., Whitehead, M., & Stuckler, D. (2017). Introduction of a National Minimum Wage Reduced Depressive Symptoms in Low-Wage Workers: A Quasi-Natural Experiment in the UK. *Health Economics*, 26(5), 639–655. <https://doi.org/10.1002/hec.3336>
- Richards, L., & Paskov, M. (2016). Social class, employment status and inequality in psychological well-being in the UK: Cross-sectional and fixed effects analyses over

- two decades. *Social Science & Medicine*, 167, 45–53. <https://doi.org/10.1016/j.socscimed.2016.08.042>
- Sbarra, D. A., Emery, R. E., Beam, C. R., & Ocker, B. L. (2014). Marital Dissolution and Major Depression in Midlife: A Propensity Score Analysis. *Clinical Psychological Science*, 2(3), 249–257. <https://doi.org/10.1177/2167702613498727>
- Steele, F., French, R., & Bartley, M. (2013). Adjusting for selection bias in longitudinal analyses using simultaneous equations modeling: The relationship between employment transitions and mental health. *Epidemiology*, 24(5), 703–711. <https://doi.org/10.1097/EDE.0b013e31829d2479>
- Swartz, T. T., Kim, M., Uno, M., Mortimer, J., & O'Brien, K. B. (2011). Safety Nets and Scaffolds: Parental Support in the Transition to Adulthood. *Journal of Marriage and Family*, 73(2), 414–429. <https://doi.org/10.1111/j.1741-3737.2010.00815.x>
- Sweeting, H., Green, M., Benzeval, M., & West, P. (2016). The emergence of health inequalities in early adulthood: Evidence on timing and mechanisms from a West of Scotland cohort. *BMC Public Health*, 16(1), 41. <https://doi.org/10.1186/s12889-015-2674-5>
- Teachman, J., Duncan, G. J., Yeung, W. J., & Levy, D. (2001). Covariance Structure Models for Fixed and Random Effects. *Sociological Methods & Research*, 30(2), 271–288.
- Thomas, C., Benzeval, M., & Stansfeld, S. A. (2005). Employment transitions and mental health: An analysis from the British household panel survey. *Journal of Epidemiology & Community Health*, 59(3), 243–249. <https://doi.org/10.1136/jech.2004.019778>
- Vaisey, S., & Miles, A. (2017). What You Can—And Can't—Do With Three-Wave Panel Data. *Sociological Methods & Research*, 46(1), 44–67. <https://doi.org/10.1177/0049124114547769>
- Vallejo-Torres, L., Hale, D., Morris, S., & Viner, R. M. (2014). Income-related inequality in health and health-related behaviour: Exploring the equalisation hypothesis. *Journal of Epidemiology & Community Health*, 68(7), 615–621. <https://doi.org/10.1136/jech-2013-203306>
- van de Mheen, H., Stronks, K., Schrijvers, C. T. M., & Mackenbach, J. P. (1999). The influence of adult ill health on occupational class mobility and mobility out of and into employment in The Netherlands. *Social Science & Medicine*, 49(4), 509–518. [https://doi.org/10.1016/S0277-9536\(99\)00140-9](https://doi.org/10.1016/S0277-9536(99)00140-9)
- Warren, J. R. (2009). Socioeconomic Status and Health across the Life Course: A Test of the Social Causation and Health Selection Hypotheses. *Social Forces*, 87(4), 2125–2153. <https://doi.org/10.1353/sof.0.0219>
- Wickrama, K. A. S., Noh, S., & Elder, G. H. (2009). An investigation of family SES-based inequalities in depressive symptoms from early adolescence to emerging adulthood. *Advances in Life Course Research*, 14(3). <https://doi.org/10.1016/j.alcr.2010.04.001>
- Wickrama, T., & Wickrama, K. A. S. (2010). Heterogeneity in Adolescent Depressive Symptom Trajectories: Implications for Young Adults' Risky Lifestyle. *Journal of Adolescent Health*, 47(4), 407–413. <https://doi.org/10.1016/j.jadohealth.2010.02.013>
- Williams, R., Allison, P. D., & Moral-Benito, E. (2018). Linear dynamic panel-data estimation using maximum likelihood and structural equation modeling. *The Stata Journal*, 18(2), 293–326.
- Wisco, B. E. (2009). Depressive cognition: Self-reference and depth of processing. *Clinical Psychology Review*, 29(4), 382–392. <https://doi.org/10.1016/j.cpr.2009.03.003>

8 Appendix

Contents

Missing data approach

Figure S1 Binned scatterplots showing the association between GHQ score and monthly household income in its raw form (a) and transformed (b)

Figure S2 Mental health inequality by sex, age, and whether ‘living at home with parent(s)’

General Health Questionnaire (GHQ-12) items and response categories

Discussion of alternative specification with no lagged dependent variable (Model 7)

Table S1 Descriptive statistics, by subsample

Table S2 Disaggregation of *EN* and *NN* transitions, percentages by subsample

Table S3 Longitudinal SEM path estimates and model fit statistics, aggregated sample; including supplementary models 2a, 4a, and 7

Table S4 Path α across all models, by sex and age group

Missing data approach

Understanding Society provides a range of weights to account for sampling design as well as non-response and attrition. The latter is accomplished through modelling based on a wide range of information collected at various levels (including region, neighbourhood, household, and individual) and from various sources including, if applicable, previous waves of the survey. Extensive detail is provided in the Understanding Society *User Guide* (Knies, 2018).

For the cross-sectional description of mental health inequality by age and sex, I pool observations from all waves, using the cross-sectional individual-level self-completion weight. These results can be thought of as pooling eight samples which are each representative of the UK at the time that that wave's data was collected. 'Self-completion' refers to the fact that these weights adjust for the differential probability of individuals to complete the self-completion portion of the individual interview, which contains the questionnaire modules on mental health.

In the longitudinal structural equation models, I combine the use of longitudinal individual-level self-completion weights with full-information maximum likelihood estimation (FIML). This is in order to strike a balance between robust adjustment for selective attrition on the one hand, and maximising use of the available data on the other. The longitudinal weights provided by Understanding Society are constructed such that an individual who does not respond at a given wave receives a zero weight for that and all subsequent waves. Using the longitudinal weights from wave 8 thus drastically reduces sample size, discarding information even from individuals who have only missed one wave (except insofar as they inform the weighting of the minority of individuals who are successfully interviewed at all waves). FIML utilises all available data, including from non-complete cases, but does so under a missing-at-random assumption. FIML is however highly computationally intensive.

I therefore use the longitudinal weights from wave 3 and thus include only individuals who respond in waves 1-3 (or 2-3 in the case of the BHPS sample). The sample is further restricted to individuals who responded to at least three of the remaining five survey waves, because FIML failed to converge in some cases without this restriction. The weights used adjust for survey design and initial non-response, and for predictors of attrition over three waves. I therefore make the assumption that the predictors of attrition across three waves are similar to

the predictors of attrition across eight waves. The longitudinal sample is representative of the UK population at wave 2, since this is the wave at which this combined sample (Understanding Society and BHPS) enters the study.

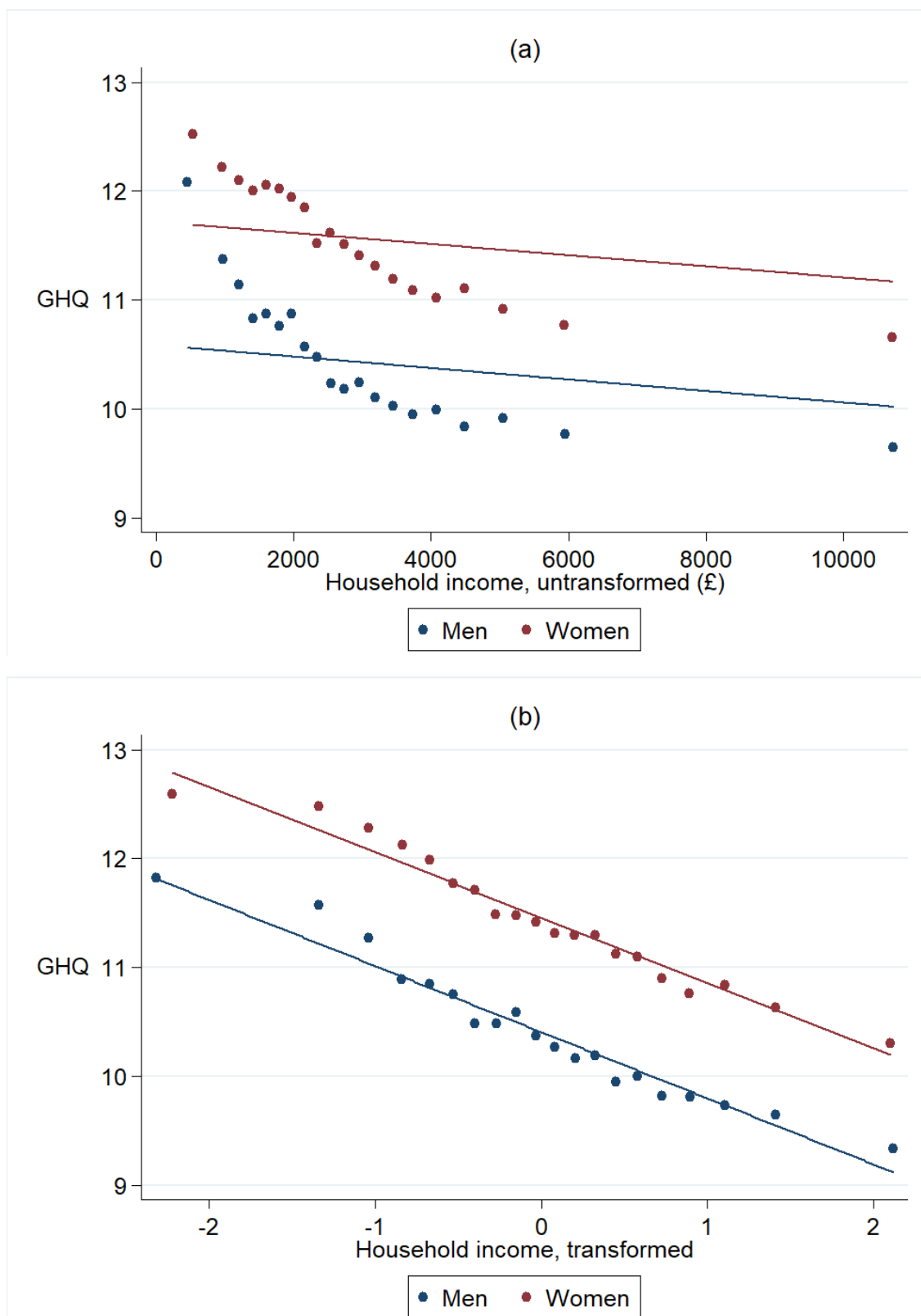


Figure S1 Binned scatterplots showing the association between GHQ score and monthly household income in its raw form (a) and transformed (b)

Note: see *Household income* above for details. Higher GHQ score indicates worse mental health.

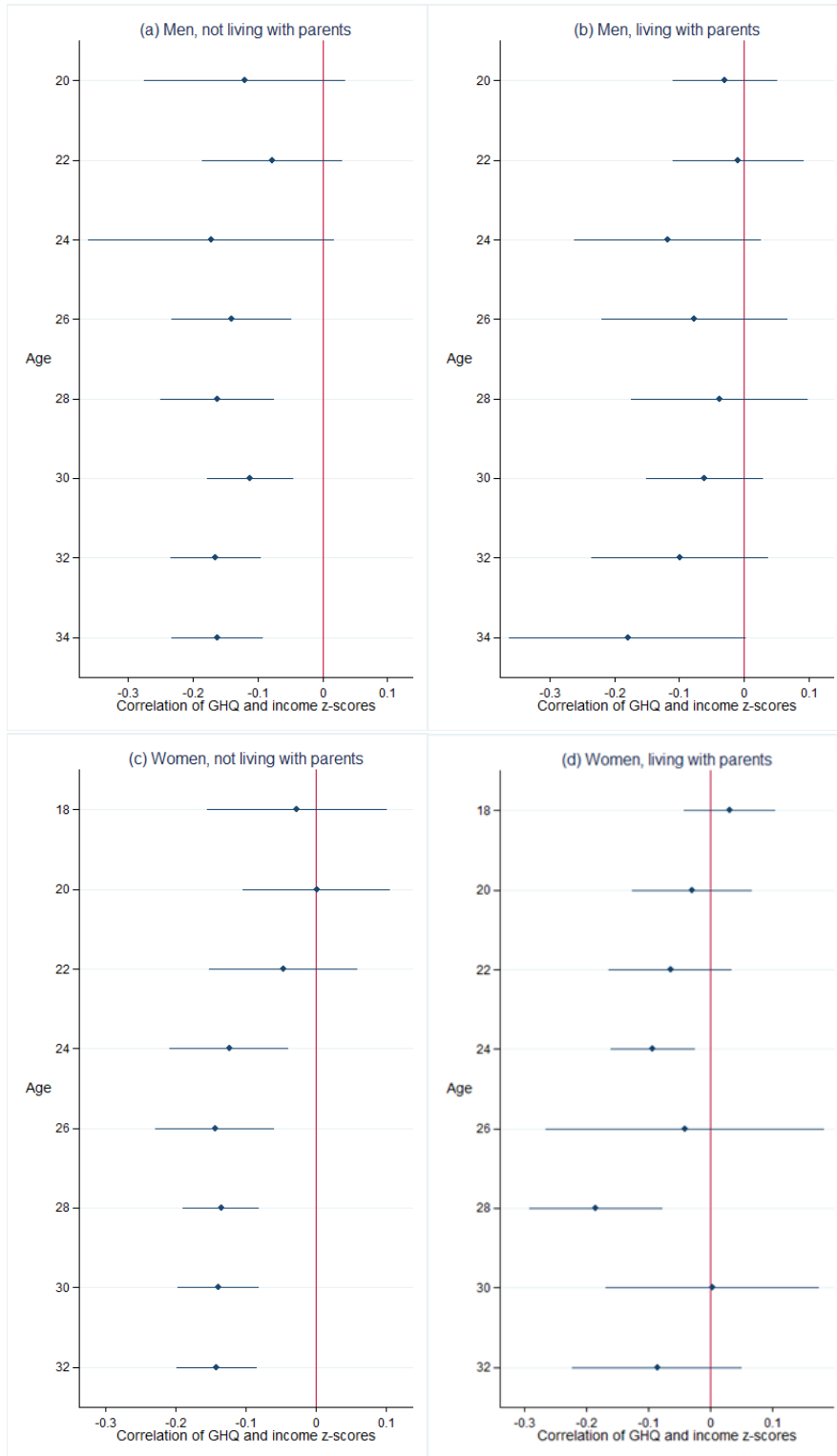


Figure S2 Mental health inequality by sex, age, and whether ‘living at home with parent(s)’

Note: the correlations shown correspond to path *a* in Figure 1a. Higher GHQ score indicates worse mental health. Men aged 16-19 and women aged 16-17 omitted because very few not living with parents; women aged 33-34 omitted because very few living with parents.

General Health Questionnaire (GHQ-12) items and response categories

‘The next questions are about how you have been feeling recently.

Have you recently...

- been able to concentrate on whatever you're doing?^a
- lost much sleep over worry?^b
- felt that you were playing a useful part in things?^a
- felt capable of making decisions about things?^a
- felt constantly under strain?^b
- felt you couldn't overcome your difficulties?^b
- been able to enjoy your normal day-to-day activities?^a
- been able to face up to problems?^a
- been feeling unhappy or depressed?^b
- been losing confidence in yourself?^b
- been thinking of yourself as a worthless person?^b
- been feeling reasonably happy, all things considered?^a

^a Response categories: *More [or better] than usual / Same as usual / Less so than usual / Much less than usual*

^b Response categories: *Not at all / No more than usual / Rather more than usual / Much more than usual*

Note: from a self-completion questionnaire module.

Discussion of alternative specification with no lagged dependent variable (Model 7)

Model 7 offers an alternative specification for comparison: it is the same as Model 6 but without the lagged dependent variables (LDVs), i.e. paths b and c are constrained to zero. Results from Model 7 are included in Tables S3 and S4. Results are similar to Model 6 and do not alter the conclusions.

Should Model 6 or Model 7 be preferred? There are arguments for each. One might prefer Model 7 – with no LDVs in the equations for income and mental health – under the assumption that income and mental health at $t-1$ are not causally related to themselves at t , and that persistence in these two variables owes solely to time-invariant individual characteristics captured by the fixed effects. On this account, income and mental health are ‘created anew’ with each wave (Finkel, 1995: 8). The interpretation of individual parameters in the presence of fixed effects is also simplified by exclusion of the LDV terms.

While Finkel (1995) generally advocates the LDV model for the analysis of change using panel data, Allison (1990) argues that in most situations the first-difference estimator is preferable.⁴ The two exceptions are in the case of a causal effect of Y_{t-1} on Y_t , and ‘when the *transient* (period-specific) components of $[Y_{t-1}]$... are correlated with X ’ (emphasis in original) (Allison, 1990: 109), X being a treatment that occurs between waves. Finkel (1995) further argues that even in the absence of a causal effect of Y_{t-1} on Y_t , an LDV is preferable in the presence of regression to the mean.

Even from the position of Allison (1990) however, there are grounds for including LDVs for income and depressive symptoms. For the former, causal mechanisms are plausible though likely to play a relatively weak role in determining future income. Greater income may allow individuals to make investments which generate higher future income; employees may accept jobs only on the condition that firms match or exceed their existing salary, regardless of other considerations; or employers may view past salary as the best available signal of an employee’s value. Indeed income has been used as an illustrative case both of past values being likely to influence future values (Bollen & Brand, 2010; Finkel, 1995) and of the opposite (Allison,

⁴ Though cf. Allison et al. (2017) which develops a cross-lagged panel model with fixed effects as an alternative to econometric first-differencing-based methods, and which outperforms them in simulations (Leszczensky & Wolbring, 2019). Such cross-lagged models include lagged dependent variables.

1990). The other justifications for an LDV are more relevant for depressive symptoms. Correlations between the transient components of depressive symptoms (i.e. mental health shocks) and the ‘treatment’ here, employment transitions, are precisely what has been found, though not in every case. Regression to the mean is also particularly likely with respect to depressive symptoms. High GHQ score is a validated predictor of case-level depression (Lundin et al., 2016), which tends to present episodically. Recent psychiatric epidemiological research has argued for a distinction between the more common episodic or non-chronic depression, and chronic depression, with the latter representing about 20-30% of cases (Köhler et al., 2015; Murphy & Byrne, 2012). Even chronic depression, however, ‘typically exhibit[s] a waxing and waning course’ (Klein, 2008: 553).

In sum, one can make a reasonable case for both Models 6 and 7; as it turns out, they yield the same substantive conclusions.

Table S1 Descriptive statistics, by subsample

	Aggregated sample		16-32		Men 33-49		50-66		16-32		Women 33-49		50-66	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Income	7.36	0.53	7.31	0.53	7.40	0.52	7.43	0.57	7.25	0.50	7.34	0.51	7.39	0.54
Between SD		0.44		0.41		0.44		0.47		0.40		0.43		0.45
Within SD		0.30		0.33		0.28		0.31		0.31		0.28		0.30
GHQ	11.15	5.53	10.32	5.21	10.78	5.13	10.35	5.09	11.71	5.94	11.72	5.82	11.47	5.60
Between SD		4.22		3.77		3.85		4.06		4.26		4.41		4.36
Within SD		3.60		3.62		3.44		3.07		4.16		3.84		3.53
Employed	0.69		0.74		0.90		0.61		0.63		0.78		0.55	
<i>Employment transitions</i>														
NN (not employed - not employed)	0.26		0.23		0.07		0.33		0.30		0.18		0.40	
EN (employed - not employed)	0.05		0.04		0.02		0.06		0.07		0.04		0.05	
NE (not employed - employed)	0.04		0.08		0.03		0.02		0.11		0.05		0.02	
EE (employed - employed)	0.65		0.65		0.88		0.59		0.52		0.73		0.52	
Individuals always NE or EE	0.51		0.50		0.80		0.44		0.30		0.60		0.39	
Individuals ever EN	0.26		0.19		0.13		0.33		0.35		0.21		0.30	
Individuals ever NN	0.39		0.44		0.12		0.47		0.53		0.28		0.53	
Individuals always NN	0.16		0.09		0.05		0.20		0.13		0.11		0.29	
N(observations)	111,835		8,746		17,863		20,060		12,940		25,874		26,352	
N(individuals)	17,843		1,461		2,851		3,143		2,144		4,114		4,130	

Note: means (and proportions) are given at the observation level unless specified. Income and GHQ are standardised in the analysis (mean 0, SD 1) but presented here prior to standardisation. Income is presented here after the transformations described in the text have been applied. Unweighted statistics.

Table S2 Disaggregation of *EN* and *NN* transitions, percentages by subsample

	Aggregated sample	16-32	Men 33-49	50-66	16-32	Women 33-49	50-66
<i>EN</i> : employed ->							
unemployed	25.3	55.6	68.2	21.0	18.6	25.6	13.7
retired	34.0	0.0	3.0	69.8	0.0	1.5	65.7
student	7.9	38.6	6.3	0.2	22.0	5.6	0.4
long-term sick/disabled	5.4	1.6	9.8	5.9	1.4	6.4	6.4
looking after home/family	27.4	4.3	12.8	3.1	58.1	60.8	13.8
<i>NN</i> : continuously...							
unemployed	7.4	19.8	33.6	6.2	8.6	7.2	2.1
retired	39.9	0.0	1.4	68.4	0.0	0.2	66.8
student	9.6	60.2	2.3	0.0	37.0	2.3	0.1
long-term sick/disabled	10.1	2.9	27.6	11.8	2.6	14.1	9.3
looking after home/family	18.0	0.8	11.3	2.2	32.7	57.6	9.0
(transitions between non- employed states)	15.0	16.3	23.8	11.4	19.1	18.6	12.7

Note: numbers are percentages of the *EN* (or, separately, *NN*) observations within each subsample in which the *N* corresponds to each self-reported non-employment status.

Table S3 Longitudinal SEM path estimates and model fit statistics, aggregated sample; including supplementary models 2a, 4a, and 7

Model		1	2	2a	3	4	5	6	7	4a		
<i>a</i>	$\varepsilon_t \leftrightarrow u_t$	-0.139 ***	-0.060 ***	-0.047 ***	-0.047 ***	-0.047 ***	-0.038 ***	-0.029 ***	-0.030 ***	<i>a</i>	$\varepsilon_t \leftrightarrow u_t$	-0.047 ***
<i>b</i>	$\text{Income}_{t-1} \rightarrow \text{income}_t$		0.717 ***	0.694 ***	0.694 ***	0.694 ***	0.325 ***	0.303 ***		<i>b</i>	$\text{Income}_{t-1} \rightarrow \text{income}_t$	0.696 ***
<i>c</i>	$\text{GHQ}_{t-1} \rightarrow \text{GHQ}_t$		0.551 ***	0.552 ***	0.543 ***	0.552 ***	0.160 ***	0.160 ***		<i>c</i>	$\text{GHQ}_{t-1} \rightarrow \text{GHQ}_t$	0.552 ***
<i>d</i>	$\text{Income}_1 \leftrightarrow \text{GHQ}_1$		-0.147 ***	-0.147 ***	-0.144 ***	-0.147 ***	-0.004	-0.030 ***		<i>d</i>	$\text{Income}_1 \leftrightarrow \text{GHQ}_1$	-0.153 ***
<i>e</i>	$\text{EN}_t \rightarrow \text{income}_t$			-0.634 ***	-0.639 ***	-0.634 ***		-0.636 ***	-0.595 ***	<i>e</i>	$\text{EN}_t \rightarrow \text{income}_t$	-0.629 ***
<i>f</i>	$\text{NN}_t \rightarrow \text{income}_t$			-0.227 ***	-0.233 ***	-0.227 ***		-0.399 ***	-0.507 ***	<i>f</i>	$\text{NN}_t \rightarrow \text{income}_t$	-0.219 ***
<i>g</i>	$\text{EN}_t \rightarrow \text{GHQ}_t$				0.154 ***			0.202 ***	0.188 ***			
<i>h</i>	$\text{NN}_t \rightarrow \text{GHQ}_t$				0.161 ***			0.126 ***	0.109 ***			
<i>i</i>	$u_{t-1} \leftrightarrow \text{EN}_t$					0.026 ***		0.039 ***	0.033 ***		$\text{GHQ}_{t-1} \rightarrow \text{EN}_t$	0.002 *
<i>j</i>	$u_{t-1} \leftrightarrow \text{NN}_t$					0.000		-0.004	-0.007 **		$\text{GHQ}_{t-1} \rightarrow \text{NN}_t$	0.040 ***
<i>k</i>	$u_{t-2} \leftrightarrow \text{EN}_t$					0.024 ***		0.030 ***	0.020 ***		$\text{GHQ}_{t-2} \rightarrow \text{EN}_t$	0.000
<i>l</i>	$u_{t-2} \leftrightarrow \text{NN}_t$					0.013 ***		0.010	0.005		$\text{GHQ}_{t-2} \rightarrow \text{NN}_t$	0.034 ***
Fixed effects		No	No	No	No	No	Yes	Yes	Yes	Fixed effects		No
RMSEA		0.274	0.108	0.085	0.070	0.072	0.035	0.030	0.046	RMSEA		0.187
AIC		484512	319345	316139	269260	308645	364836	257240	243582	AIC		447382
CFI		0.016	0.835	0.746	0.821	0.820	0.984	0.971	0.938	CFI		0.349
TLI		0.122	0.847	0.771	0.838	0.832	0.984	0.970	0.929	TLI		0.354
SRMR		0.370	0.179	0.129	0.106	0.106	0.068	0.066	0.018	SRMR		0.211

Note: see note to Table 1.

Table S4 Path α across all models, by sex and age group

	Men			Women		
	16-32	33-49	50-66	16-32	33-49	50-66
Model 1	-0.092 ***	-0.149 ***	-0.171 ***	-0.093 ***	-0.161 ***	-0.129 ***
Model 2	-0.052 ***	-0.070 ***	-0.066 ***	-0.038 ***	-0.085 ***	-0.037 ***
Model 2a	-0.032 *	-0.035 **	-0.057 ***	-0.021 *	-0.065 ***	-0.035 ***
Model 3	-0.031 *	-0.034 **	-0.061 ***	-0.021 *	-0.065 ***	-0.035 ***
Model 4	-0.029 *	-0.035 **	-0.050 ***	-0.019 *	-0.065 ***	-0.034 ***
Model 4a	-0.032 *	-0.036 **	-0.056 ***	-0.021 *	-0.066 ***	-0.035 ***
Model 5	-0.043 **	-0.057 ***	-0.041 ***	-0.015	-0.060 ***	-0.019 *
Model 6	-0.022	-0.031 **	-0.026 **	-0.011	-0.051 ***	-0.018 *
Model 7	-0.027	-0.034 **	-0.025 **	-0.013	-0.049 ***	-0.014

Note: compare Table 3.