Article:

Rice, N. orcid.org/0000-0003-0312-823X, Roberts, J. orcid.org/0000-0003-2883-7251 and Sechel, C. orcid.org/0000-0002-8653-3443 (2025) Mental health and labour productivity. Journal of Economic Behavior & Organization, 236. 107075. ISSN 0167-2681

https://doi.org/10.1016/j.jebo.2025.107075

Mental health and labour productivity

Nigel Rice[a,b,*], Jennifer Roberts[c], Cristina Sechel[c]

[a] *Centre for Health Economics, University of York, UK*
[b] *Department of Economics and Related Studies, University of York, UK*
[c] *Department of Economics, University of Sheffield, UK*

Abstract

We present novel evidence of the effect of mental health on productivity using a direct measure of productivity from the COVID-19 modules of the UK Household Longitudinal Study. We employ spatial variation in COVID-19 deaths as an instrumental variable and supplement results by computing bounds by considering coefficient stability to observable factors to infer the influence of unobservables. Our findings reveal a substantive positive relationship between poor mental health and decreased productivity. Our estimates suggest productivity losses of around 54 minutes per week (on average) for individuals who report a decline in mental health.

Keywords: mental health, productivity, UKHLS.

JEL: I12, J14, J24.

Acknowledgements

The UK Household Longitudinal Study (UKHLS) is an initiative by the Economic and Social Research Council, with scientific leadership by the Institute for Social and Economic Research, University of Essex, and survey delivery by the National Centre for Social Research. The data were accessed via the UK Data Service. We are grateful to Pinka Chatterji for providing code to adapt to assess coefficient stability in the presence of unobservables, and to Dacheng Huo for data and research assistance. For the purpose of open access, the author(s) has applied a creative commons attribution (CC BY) licence to any author accepted manuscript version arising.

* Corresponding author

Email addresses: `nigel.rice@york.ac.uk` (Nigel Rice), `j.r.roberts@sheffield.ac.uk` (Jennifer Roberts), `c.sechel@sheffield.ac.uk` (Cristina Sechel)

1. Introduction

"Productivity isn't everything, but, in the long run, it is almost everything. A country's ability to improve its standard of living over time depends almost entirely on its ability to raise its output per worker". Paul Krugman in The Age of Diminished Expectations (Krugman 1994).

Productivity growth is recognised as the most important contributory factor to sustainable gains in living standards. Understanding the drivers of productivity, and in particular, reasons underlying recent global trends towards a slowdown in productivity growth is fundamental to securing future well-being. However, the relationship between well-being and productivity is complex, especially when considering the potential impact of health (see Sharpe & Mobasher Fard (2022) for an overview). Health is a key component of human capital and an important factor of production (Layard 2013). We know, for example, that healthier workers are, on average, more productive (Burton et al. 2005) and this is particularly true in the case of mental health (MH) (Lerner & Henke 2008). MH problems, especially anxiety and depression, have a greater impact on ability to work than any other group of disorders.[1] In addition to a large gap in employment rates between those with and without these problems (Banerjee et al. 2017, Bryan et al. 2022), workers with MH problems also earn lower wages than those without (Contoyannis & Rice 2001). An increasing prevalence of MH conditions may be partly responsible for the persistently low productivity levels that characterise the UK economy and may also contribute to the 'productivity puzzle' that the UK has experienced following the 2008 global financial crisis (Pessoa & Van Reenen 2014), defined by falling real wages along with static (or declining) output per worker and rising employment. Changes in the composition of the workforce have been suggested as a possible explanation for this puzzle, but this debate has neglected health, focusing instead on education, skills and job type (Emmerson et al. 2013). However, the prevalence of MH problems is increasing and welfare-to-work policies have increased the incentives for workers with MH problems to participate in

[1] Around half of Employment Support Allowance claimants, and a third of Personal Independence Payments claimants in the UK, have a mental or behavioural disorder as their main condition or disability for their claim (House of Commons 2023).

work.[2]

Empirical validation of the relationship between MH and productivity, particularly at an individual level, is hampered by limited availability of direct measures of productivity in large secondary data sets. Given the fundamental issues we outline above, this is an important evidence gap. In a rare empirical study in this area Oswald et al. (2015) use an experimental design to explore the relationship between happiness and productivity, and find that happier individuals have approximately 12% greater productivity in a piece rate setting.[3] In this study we employ a different approach, and attempt to address the evidence gap by exploring whether changes in MH contribute to changes in productivity in the UK. We exploit the COVID-19 modules of the UK Household Longitudinal Study which include a direct (self-reported) measure of productivity change relative to pre-COVID levels. This type of measure is rarely available and, as far as we are aware, has never been used to study the relationship between health and productivity. Rather, the vast majority of evidence relies on wages to proxy productivity (e.g. Meyer & Mok 2019); or estimates the effects of MH problems directly on employment outcomes (e.g. Bryan et al. 2022).[4] As such, our analysis contributes a unique perspective that complements existing work on this topic, which relies mainly on proxy measures of productivity.

We measure MH via the General Health Questionnaire (GHQ) both before and during the pandemic. This allows us to identify a continuum of MH states that are not limited to having (or not having) diagnosed MH conditions. The relationship between MH and productivity is likely endogenous. We deal with this in part by considering changes in MH from pre-COVID-19 to during the pandemic with a measure of changes in productivity across the same period. In addition, we assess the extent of additional endogeneity bias through the use of an exclusion restriction (instrumental

[2] The employment rate of people with MH problems has been increasing steadily in recent years; from around 37% at the start of 2018 to around 45% in early 2020 (Roberts et al. 2021).

[3] Happiness is not generally considered to be equivalent to MH; the former is a measure of an emotional affective state, while the latter is more evaluative. Nevertheless Oswald et al. (2015) is one of the few studies that directly explores the relationship between either of these concepts and productivity.

[4] There is also a related strand of literature that estimates the effect of MH on absenteeism and presenteeism, for example Bubonya et al. (2017) and Bryan et al. (2021). Also see a review by De Oliveira et al. (2023). However, the measures used in this work are not closely related to the concept of productivity that we adopt.

variable) and supplement this by considering coefficient stability to the influence of unobservable factors correlated with both productivity and MH using the approach of Altonji et al. (2005). In particular, we bound estimates using an upper bound assuming no unobserved selection (into changes in MH) and a lower bound where unobserved selection is constrained to be equivalent to selection through observables.

Our results show a positive relationship between MH and productivity. Although women experienced larger reductions in MH as well as larger falls in productivity compared to men, we find no statistically significant gender differences in the relationship between MH and productivity in our baseline ordered probit results. Instrumenting for MH using a measure of local area COVID-19 deaths suggests no evidence of endogeneity bias in baseline estimates. Supplementing the IV approach by restricting bias due to unboservables to that defined by observables (via the approach of Altonji et al. (2005)) reaffirms this finding for men. For women, this approach leads to a lower bound estimate that includes a null effect. The 'true' effect likely resides between the null of zero and the positive baseline ordered probit result.

Taken as a whole, our results suggest that while individual productivity changes as a result of MH deterioration are relatively small, these aggregate to substantial productivity losses for the economy as a whole where population MH is deteriorating. This link between MH and productivity strengthens the case for public policy to invest in MH prevention and treatment programmes. Not only will such measures improve population well-being directly through better health, but they also have the potential to enhance productivity, and hence, indirectly lead to further well-being gains through increased living standards.

2. Data

We use data from the Understanding Society COVID-19 Study, which consists of a series of 9 surveys conducted between April 2020 and October 2021 (ISER 2021). The eligible sample consists of all individuals aged 16 years and above in April 2020 from households who participated in waves 8 and 9 of the main UK Household Longitudinal Study (UKHLS).[5] The surveys were administered online with some telephone

[5]Excluding individuals who refused to take part in the main UKHLS questionnaire or where mentally or physically unable to make informed decisions about participating in the survey. Individuals with unknown or foreign postal addresses were also excluded from the COVID-19 study.

interviews for households without internet access. We focus on module 3 collected at the end of June 2020 for our main results, and supplement these with sensitivity analyses using data from the September 2020 module.[6] These modules cover the first phase of the pandemic where policies on social distancing, lockdown, and working from home were most prevalent. The announcement of the first lockdown in the UK ordering people to stay at home and including the closure of pubs, restaurants, gyms and other social venues took place on 23 March 2020.[7] Following a peak of daily confirmed cases in April 2020, lockdown restrictions began to be eased with separate announcements from mid-May (plan for easing restrictions) through to the beginning of July when hospitality venues were reopened. Importantly the imposition of major restrictions and their durations were uniformly applied across England and Wales. The first exception to this was the 4 July (after the collection of the June Module 3 data) where local lockdown restrictions came into force in one city and parts of its county.[8] A reintroduction of restrictions was ordered on 14 September (socialising in groups of up to 6), and a return to working from home on 22 September 2020. In total 31,964 people were invited to take part in the June COVID-19 module and 19,372 invited for the September module. Response rates were 44.2% in June and 66.5% in September. Importantly, the COVID-19 modules can be linked to data from past (and future) UKHLS waves, which allows us to link baseline information about respondents prior to the start of the pandemic.

Our measure of productivity is based on responses to the question "Please think about how much work you get done *per hour* these days. How does that compare to how much you would have got done *per hour* in January/February 2020?".[9] The

[6]Productivity data was only collected in modules 3, 5, 7 and 9. Module 3 interviews took place between 25 June and 1 July, and module 5 interviews between 24 September and 1 October. Later surveys took place in January and September 2021. It is possible that other factors (beyond the initial shock of the pandemic) came into play during the later months for which we are unable to control. We focus on the first module in order to minimise bias arising from confounding factors and recall error.

[7]Exceptions were made for essential workers, people who were unable to work remotely, shopping for essential goods, accessing medical care, and undertaking exercise outdoors.

[8]Leicester and parts of Leicestershire were subject to local restrictions in early July 2020. For a timeline of UK Government Coronavirus lockdowns and measures, see https://www.instituteforgovernment.org.uk/sites/default/files/timeline-lockdown-web.pdf

[9]It is clear from the module and question routing that the question relates to market work and is only asked of respondents if they are either employed or self-employed.

response categories available are: 1. "I get much more done", 2. "I get a little more done", 3. "I get about the same done", 4. "I get a little less done", 5. "I get much less done" (see Table A1). As a follow-up question in the September wave respondents are asked to quantify how long it previously took to get done what currently (at the time of the module questionnaire) takes an hour. The responses available and the frequency of answers are provided in Table A2.

While it is the case that people in declining MH might over- or under-report the changes in their productivity, we have no evidence to support any systematic bias in one direction or the other. The relatively short time comparisons involved here (June and September, vs. the previous February) mean that the recall bias often associated with self-reported measures, should not be a serious problem.[10] However, we do have reason to suspect that measurement error exists. For each set of respondents, i.e. those who report getting more done, and those who report getting less done, there is considerable overlap in the categories used to quantify how long it would have previously taken to complete one hour of work. For example, of the 44% of respondents claiming they get much more done in an hour now than previously, 31% say that an hour of work now would have taken up to 75 minutes previously, and 44% say it would have taken between 75 and 90 minutes. However, of respondents who say they get a little more done now, 56% reported it would have taken up to 75 minutes previously and 34% between 75 and 90 minutes. This illustrates the likely existence of non-negligible measurement error in the original 5-point scale used to categorise productivity changes. To reduce concerns over measurement errors we construct the following three point categorical variable: 1. "much less or a little less done", 2. "about the same", 3. "much more or a little more done". In the June module, the productivity question is only asked of respondents who work from home at least some of the time. This restriction was dropped in the September module, so all working respondents answered the question. Importantly, this question asks about the amount of work achieved per hour and represents a change in productivity benchmarked against the two months immediately preceding the first lockdown in March 2020.

[10]Other biases arising from social desirability or justification of certain behaviours should be minimised as there is no reason for the respondent to think there is a 'right' or 'acceptable' answer to this question, especially given the context of the pandemic disruption and uncertainty.

MH is measured using the 12-item General Health Questionnaire (GHQ). GHQ is a widely used screening instrument for common mental disorders and a general measure of psychiatric well-being in the population (Goldberg et al. 1998). It has been validated in a number of international studies (Sartorius & Ustün 1995, Goldberg et al. 1997, Schmitz et al. 1999), and has been shown to be predictive of face-to-face clinical diagnosis of MH problems (Anjara et al. 2020). It has gained much attention as a measure of psychological health in studies of the relationship between MH and labour market outcomes (for example, see Garcia-Gomez et al. 2010, Mavridis 2015, Lagomarsino & Spiganti 2020, Bryan et al. 2022). The GHQ consists of 12 items intended to assess the severity of mental problems in the last few weeks; as such, this is a measure of current MH. Each item is scored using a 4-point Likert scale. These are then aggregated to generate a total score ranging from 0 to 36. For ease of interpretation, we reverse code this score such that higher values indicate better MH. GHQ scores are collected in the COVID-19 modules as well as in the main UKHLS waves. This allows us to measure the changes in MH relative to a pre-COVID baseline. Accordingly, our measure of MH, GHQdiff, is the difference between the GHQ score in June (or September) and the baseline GHQ score for the same individual in the final UKHLS interview taken before 2020.[11] Positive values of GHQdiff denote an improvement in MH, while negative values signal a deterioration. It is the case that on average women generally report worse GHQ scores than men (see for example Madden 2010), and our results reflect this with a mean at baseline for men of 25.41 and 24.03 for women (both increasing in good health). However, our models use the change in the GHQ score, and there is very little evidence available on whether or not there are gender differences in the ability of the GHQ to detect changes in MH. In a rare study Schlechter & Neufeld (2024) explore this for young adults (using data from the UK Next Steps survey) and find that gender differences in the GHQ over time exhibit measurement invariance, and are truly attributable to differences in the latent MH construct, and not measurement differences. This finding is reassuring for our interpretation that the changes in GHQ are equally reliable measures of MH change for both genders.

Data on COVID-19 deaths are available for England and Wales from the Office for

[11] The vast majority of baseline interviews (approximately 95%) occurred in 2019, the rest took place in early 2020.

National Statistics at the level of Middle Layer Super Output Areas (MSOAs).[12,13] We use total COVID-19 related deaths during the months of March-July 2020 as our identifying instrument, which is excluded from the productivity outcome equation, as set out in Section 3. Due to a more limited geography available for the deaths data compared to the UKHLS sample, merging the datasets results in a smaller sample size for analyses when using the instrument (n = 2,538) than available for non-IV estimation approaches (2,902).

Table 1: Summary statistics: June 2020

	Full sample		Males		Females	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Productivity change						
(Much or a little) less done	0.318	0.466	0.286	0.452	0.346	0.476
Same	0.426	0.495	0.463	0.499	0.395	0.489
(Much or a little) more done	0.255	0.436	0.251	0.434	0.259	0.438
Health						
GHQdiff	-1.676	6.113	-1.477	5.605	-1.845	6.513
Health condition	0.409	0.492	0.408	0.492	0.410	0.492
Socio-demographic						
Female	0.539	0.499				
Age	44.81	12.44	46.08	12.42	43.72	12.36
Ethnicity	0.077	0.266	0.091	0.288	0.064	0.245
Degree	0.698	0.459	0.723	0.448	0.676	0.468
Employment						
Employed	0.808	0.394	0.777	0.417	0.834	0.372
Self-employed	0.149	0.356	0.179	0.384	0.122	0.328
Both self & employed	0.044	0.204	0.044	0.205	0.043	0.204
Household						
Couple	0.731	0.444	0.779	0.415	0.690	0.463
Kids04	0.116	0.320	0.125	0.331	0.108	0.311
Kids515	0.300	0.458	0.305	0.461	0.295	0.456
Kids1618	0.116	0.320	0.118	0.323	0.114	0.318
N	2,902		1,201		1,701	
Instrumental variable						
COVID-19 deaths	6.961	5.825	6.909	5.401	6.998	5.379
N	2,538		1,057		1,481	

Sample summary statistics for the full sample, and for males and females, weighted using cross-sectional weights.

We include a number of personal and household characteristics as control variables, based on a standard Mincerian wage equation (Mincer 1958). Detailed definitions of all variables are given in Table A1. For the analysis of the June 2020 data we also

[12] https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/datasets/deathsinvolvingcovid19bylocalareaanddeprivation

[13] MSOAs are a statistical geography and have a resident population of between 5000 and 15000 people; there are 7264 in England and Wales.

include a set of industry indicators, which capture the main sector in which the individual currently works.

Our analysis consists of 18-64 year olds who are employed or self-employed, excluding workers who were furloughed at the time of the interview, or before.[14] After dropping individuals for whom data was missing on the set of key explanatory variables, the June sample consists of $N = 2,902$ individuals, of which $1,201$ were male and $1,701$ were female (2,538 (1,057 males, 1,481 females) for models that involve the exclusion restriction).[15] Table 1 reports summary statistics for the main variables in our June sample for men and women separately, excluding the set of industry categories which are separately reported in Table A3. All statistics are weighted by the cross-sectional weights provided with the data to increase sample representativeness to the national population at baseline.

Approximately 43% of the sample reported no change in productivity relative to Jan/Feb 2020; 25.5% reported an increase (with 13.9% saying they got a little more done and 11.6% much more, not shown in table); 31.8% reported a decrease (with 21.5% getting a little less done and 10.3% getting much less done). Female respondents were more polarised in reporting changes in productivity compared to males, particularly in terms of getting less done (28.6% of men compared to 34.6% of women).

On average there was a decrease in MH during the pandemic; a result that has been documented elsewhere, for example Banks & Xu (2020) and Daly et al. (2020). The average decrease across the full sample was -1.675.[16] This represents an approximate one third of a standard deviation of the pre-COVID baseline GHQ score. The decrease was, however, larger for females (-1.845) than for males (-1.477). This finding is in line with Orefice & Quintana-Domeque (2021) who show that gender differences in

[14] At the beginning of the lockdown period the UK Government introduced the Coronavirus Job Retention Scheme. This was designed so that employers could furlough workers but receive a subsidy equivalent to 80% of employee's salaries, ensuring that furloughed workers received the majority of their pay.

[15] This constitutes approximately two thirds of respondents who were employed or self-employed in June 2020, not furloughed, and were asked about their change in productivity (individuals who reported never working from home in that month were not asked this question).

[16] Looking at the historical modules from the main Understanding Society survey (ISER 2022), there was a downward trend in mean GHQ scores prior to COVID. As expected, the deterioration in MH between 2019 and 2020 was more pronounced compared to the average annual change in mean GHQ scores between 2014 and 2019.

MH effects were associated with increased childcare and housework responsibilities for women, as well as the difference in COVID-19 related health concerns between men and women. Similarly Cheng et al. (2021) show that COVID-19 disrupted work-life balance in the household through increases in childcare, homeschooling and financial insecurity, and that the burden of these fell disproportionately on women, resulting in larger deterioration in their MH. A similar proportion, around 41%, of men and women reported having a long-term health condition. The majority, 81%, of the sample were employed. The proportion was higher for women (83%) than men (78%), and conversely a greater proportion of men than women were self-employed (18% versus 12%).

3. Empirical approach

Our measure of productivity consists of responses on a categorical (*Likert*) scale, which we model using an ordered probit (OP) (Greene & Hensher 2010). Underlying this model is a latent variable, y_i^*, which is assumed to be a linear (in unknown parameters, λ and $\boldsymbol{\beta}_x$) function of the observed MH variable (MH_i), a vector of exogenous characteristics $\boldsymbol{x}_i$ (with no constant term), and a standard normal disturbance term, ε_i, such that

$$y_i^* = \lambda MH_i + \boldsymbol{x}_i' \boldsymbol{\beta}_x + \varepsilon_i, \qquad i = 1, \ldots, N, \tag{1}$$

where λ is the parameter of interest. y_i^* is mapped onto observed $j = 0, \ldots, J-1$ outcomes as follows

$$y = j \text{ if } \mu_{j-1} \leq y_i^* < \mu_j \qquad \text{for } j = 0, \ldots, J-1,$$

where $\mu_{-1} = -\infty$ and $\mu_{J-1} = +\infty$ (to ensure well-defined probabilities, we also assume $\mu_{j-1} < \mu_j, \forall j$.) λ is our parameter of interest. The expressions for the resulting probabilities and likelihood functions are well-known (for example, see Greene & Hensher (2010)). We estimate Equation (1) separately for the June and September modules, and for men and women. Although we use cross-sectional data, our model, in part, controls for potential time-invariant measurement error and individual unobserved effects in both productivity and MH as both measures represent changes from the baseline (pre-COVID) wave. Identification of the effect of MH on the outcome in

Equation (1) requires that there are no unobservable factors that determine changes to MH and that are also related to changes in productivity.

3.1. Endogeneity of mental health

Should MH be endogenous to productivity in Equation (1) standard OP estimation will lead to biased estimates of the parameter of interest, λ. Causal interpretation requires independent variation in MH. Here we rely on exogenous variation derived from an exclusion restriction (instrumental variable) and estimate the following bivariate model:

$$\begin{aligned} y_i^* &= \lambda MH_i + \boldsymbol{x}_i'\boldsymbol{\beta}_x + \varepsilon_i, \\ MH_i &= \alpha + \boldsymbol{x}_i'\boldsymbol{\gamma}_x + \gamma_z z + \epsilon_i, \\ \text{and } \begin{bmatrix} \varepsilon_i \\ \epsilon_i \end{bmatrix} &\sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1, \rho \\ \rho, \sigma^2 \end{bmatrix}\right) \end{aligned} \tag{2}$$

for $i = 1, \ldots, N$. y_i^* is mapped onto observed $j = 0, \ldots, J-1$ outcomes as in Equation (1). z is a measure of total COVID-19 deaths in the MSOA of residence of individual i between 1 March and 31 July 2020. While non-linearity (in the outcome equation) of the bivariate model lends itself to formal identification through functional form, this alone cannot be relied upon and hence the motivation for including an exclusion restriction, z_i. Local COVID-19 deaths are likely to raise fear and anxiety about contracting the virus and hence impact MH. However, they were unlikely to directly impact on productivity conditional on the set of controls.[17] In the context of our main analysis relying on data collected in June 2020, social distancing restrictions put in place to mitigate the spread of COVID-19 were uniformly applied across the country with little discretion for variation in policies across areas that might have led to differential lockdown measures. Importantly, lockdown measures did not respond to local deaths from COVID-19. For all of these reasons, we argue that our instrument is superior to those commonly used in studies of the impact of MH on outcomes.

[17] These include industry sector, type of employment (self versus employee), working from home, education, age, gender, health conditions, number of children in household.

Instruments such as family bereavement, death or illness of a close friend or child health (Black et al. 2016, Le & Nguyen 2018, Persson & Rossin-Slater 2018, Frijters et al. 2014, von Hinke et al. 2022) are unlikely to be valid in our context, due to their potential for direct effects on productivity.

Nevertheless, claims for identification via instruments are always open to question. We therefore supplement our use of an exclusion restriction by assessing the extent of endogeneity bias in estimates of λ exploiting methods developed by Altonji et al. (2002*a*, 2005), and extended by Oster (2013, 2019). These methods rely on selection on observable characteristics to provide information on selection along unobservable factors and do not require exogenous variation brought about through the use of instruments. In the context of investigating the impact of psychiatric disorders on employment and labour force participation, Chatterji et al. (2011) employ this approach to a bivariate probit specification. We follow a similar approach by adapting the method to the bivariate model specified in Equation (2) with the absence of an exclusion restriction such that:

$$\begin{aligned} y_i^* &= \lambda MH_i + \boldsymbol{x}_i'\boldsymbol{\beta}_x + \varepsilon_i, \\ MH_i &= \alpha + \boldsymbol{x}_i'\boldsymbol{\gamma}_x + \epsilon_i, \\ \text{and } \begin{bmatrix} \varepsilon_i \\ \epsilon_i \end{bmatrix} &\sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1, \rho \\ \rho, \sigma^2 \end{bmatrix}\right) \end{aligned} \tag{3}$$

While identification in the absence of the exclusion restriction is aided by non-linearity in the outcome model, we treat the model as under-identified, in particular with respect to the parameter ρ and follow the method set out by Altonji et al. (2002*b*, 2005). First, we impose assumptions on the strength of the correlation between the unobservables determining productivity changes and unobservables determining changes to MH. This translates to placing restrictions on the value of ρ to identify λ and is complementary to specifying an exclusion restriction. The approach is informative by providing bounds on the value of λ by imposing assumptions on the role of unobservables based on the role of observables. In this respect, Altonji et al. (2005) propose a framework based on assumptions related to selection into treatment (the endogenous variable) due to unobservables and how these assumptions influence the treatment parameter of interest.[18] The idea is that the degree of selection into treatment based on observables provides a guide to the degree of selection due to unobservables (see also Oster (2019) for further insight into this approach). Assuming no selection on unobservables amounts to restricting the correlation between the errors, ρ, to zero. Assuming equal selection into treatment based on the observable and unobservable characteristics provides a point estimate for the alternative bound.

Altonji et al. (2002*b*, 2005) discuss the conditions under which selection on observable and unobservable characteristics can be assumed to be equal. The intuition behind the assumption is as follows. The outcome is assumed to be determined by a large set of characteristics (implying that no characteristic dominates the outcome or endogenous variable distributions) for which only a subset of randomly selected factors are observed. This implies that observed and unobserved characteristics can be treated symmetrically, and that treatment has an equal relationship with the explained part of the outcome (via observables) as with the unexplained part (via unobservables). A final assumption is that the regression of MH on $y^* - \lambda MH$ is equal to the regression of the part of MH that is orthogonal to x on the corresponding component of $y^* - \lambda MH$ (see Altonji et al. (2002*b*) for a fuller discussion of this condition).

Under the above assumptions Altonji et al. (2005) derive the condition for selection on unobservables to be equal to selection on observables for a bivariate probit specification. Adapting this to the context of the bivariate OP model of Equations (3) the condition is given by:[19]

$$\rho = \frac{Cov\left(\boldsymbol{x}_i'\boldsymbol{\gamma}_x \boldsymbol{x}_i'\boldsymbol{\beta}_x\right)}{\sigma Var\left(\boldsymbol{x}_i'\boldsymbol{\beta}_x\right)} \tag{4}$$

We estimate Equation (3) for the June module, and for men and women separately by firstly allowing ρ to vary and secondly by imposing condition (4).

[18] Altonji et al. (2005) set out the approach using a bivariate probit model where the outcome is determined in part by an endogenous binary treatment variable. When describing the approach we refer to selection into treatment.

[19] See Dujardin & Goffette-Nagot (2010) for a similar application in the context of joint estimation of a binary outcome and latent endogenous variable when investigating the effects neighbourhood deprivation on unemployment.

Table 2: Changes in productivity by changes to MH: June 2020

	GHQdiff					
	Less than -10	-10 to less than -5	-5 to less than 0	0 to less than 5	5 to less than 10	10 and upwards
Productivity change						
(Much or a little) less done	54.6	41.5	29.8	22.4	22.5	25.8
Same	31.7	36.5	44.6	48.2	43.2	37.6
(Much or a little) more done	13.7	22.1	25.7	29.4	34.4	36.6
N	183	340	1,041	1,018	227	93

Associations between changes to mental health and reported productivity changes. Columns report the frequencies of reporting categories of changes to productivity for discrete changes to GHQdiff.

4. Results

Our interest lies in the effect of MH on productivity. Prior to more formal analysis we consider the bivariate association between MH and productivity. Table 2 reports the percentages of respondents reporting each of the three categories of changes to productivity by changes in GHQ scores. This is provided for various levels of GHQdiff (for example, where a change in MH is observed to be: $-5 \leq \text{GHQdiff} < 0$ etc.). Greater negative (positive) changes to MH are associated with higher (lower) proportions of individuals reporting getting little or much less work done, and a lower (higher) proportion of individuals reporting getting more or much more work done. This occurs across the range of observed changes in MH and provides support for further analysis.

4.1. Ordered probit estimation

Table 3 reports OP estimates for the June sample, and for males and females separately.[20] For the pooled sample the coefficient estimate for GHQdiff, on a latent scale, is 0.039 (top panel). This is positive, indicating that increases (decreases) in MH are associated with reporting of increases (decreases) in work productivity. The result is statistically significant at conventional levels.[21]

The effect of MH on productivity is larger for males than females. While females reported a larger decrease in MH during the pandemic, and were more likely to report a decrease in productivity than men, there is some indication that the relationship between changes in MH and changes in productivity is stronger for men

[20]Table A4 reports estimates for all covariates corresponding to regressions reported in Table 3.

[21]This result also holds when restricting the sample to employed respondents only (i.e. excluding individuals who are wholly or partially self-employed).

Table 3: Effect of MH on productivity in June 2020: Ordered probit

	June 2020		
	Pooled Coef (S.E.)	Males Coef (S.E.)	Females Coef (S.E.)
Coefficient estimates on a latent scale			
GHQdiff	0.039 ***	0.042 ***	0.037 ***
	(0.005)	(0.008)	(0.006)
N	2,902	1,201	1,701
Marginal effects on change in productivity			
Much or little less done	-0.013 ***	-0.013 ***	-0.013 ***
	(0.002)	(0.003)	(0.002)
Same	0.001 ***	0.0007	0.002 **
	(0.0004)	(0.0006)	(0.0005)
Much or little more done	0.012 ***	0.013 ***	0.011 ***
	(0.001)	(0.002)	(0.002)
N	2,902	1,201	1,701

* p<0.10, ** p<0.05, *** p<0.01. All regressions contain variables for age, age-squared, educational attainment, ethnicity, employment, living as a couple, number of children in household, health conditions, whether always working from home, and industry of employment. Regressions are weighted using cross-sectional weights.

than for women. However, the coefficients (and marginal effects) are not statistically significantly different between men and women.

The coefficient estimates provided in Table 3 are difficult to interpret beyond the general direction of effect, and strictly are not comparable across sub-samples (e.g. gender) due to different scaling of the estimates.[22] Table 3 also reports average marginal effects.[23] These provide the effect of a unit change in GHQdiff on the probability of reporting each of the three ordered outcomes representing the change in productivity. As expected, increases (decreases) in GHQdiff increase (decrease) the probability of reporting getting more done and decrease (increase) the probability of reporting getting less done. In the full sample, a unit increase (improvement) in GHQdiff leads to a 1.2 percentage point (ppt) increase in getting more done and a 1.3 ppt decrease in getting less done in June 2020 relative to January/February 2020. These effects are significant at the 1% level. There are no discernible differences by gender in these estimates.

While these effects appear modest, they represent notable proportions of the over-

[22]For an OP model location and scale are not separately identified (see Greene & Hensher (2010)).
[23]Marginal effects are computed at observed values of the covariates and averaged over individuals.

all sample means of reporting lower (higher) productivity. Overall, 32% of respondents reported getting less done and 26% reported getting more done (see Table 1). The corresponding marginal effects presented in Table 3 represent a change of just under 5% of these means.

The estimates presented in Table 3 represent the average effect on productivity due to a one unit change in GHQdiff. However, one unit on the GHQ scale is relatively small and does not typically reflect substantial changes in the underlying MH of the individual. The GHQ scale ranges from 0 to 36 and for the individuals in our June sample the average GHQ score pre-COVID is 25 with a standard deviation of 5. It is therefore relevant to consider the effects of larger changes in MH. For example, we can compare someone who experienced no change in MH with someone who experienced a drop of 5 points on the GHQ scale. For accuracy, we compare predicted probabilities rather than a multiplier of the average marginal effects. For someone with no change in MH, the predicted probability of reporting getting less done is on average 29% for our sample, while the predicted probability of getting more done is 27%. The respective probabilities for someone reporting a 5 point drop in MH are 36% and 21%. These estimates suggest large differences in productivity for these two types of individuals.

Table 4: Effect of MH on productivity in June 2020: bivariate ordered probit

	June 2020		
	Pooled Coef (S.E.)	Males Coef (S.E.)	Females Coef (S.E.)
Productivity outcome equation			
GHQdiff	0.096 **	-0.031	0.099 ***
	(0.046)	(0.638)	(0.036)
Mental health outcome equation			
Exclusion restriction (COVID-19 deaths)	-0.091 **	-0.012	-0.148 ***
	(0.037)	(0.034)	(0.055)
ρ	-0.363	0.372	-0.413
	(0.310)	(3.184)	(0.256)
N	2,538	1,057	1,481

* p<0.10, ** p<0.05, *** p<0.01. Results are obtained from estimation of the bivariate probit model in Equation (2). The top panel reports coefficients and associated standard errors for MH in the outcome equation for productivity; the bottom panel provides results for the exclusion restriction or instrument for Local COVID-19 deaths in the MH equation. Coefficients are reported on a latent scale. All regressions contain variables for age, age-squared, educational attainment, ethnicity, employment, living as a couple, number of children in household, health conditions, whether always working from home, and industry of employment. Regressions are weighted using cross-sectional weights.

4.2. Endogeneity of mental health

Table 4 presents results from estimation of the bivariate model in Equation (2) including the exclusion restriction via local COVID-19 deaths. Estimated coefficients for GHQdiff remain positive and significant (at the 5% and 1% level respectively) for the pooled sample and for the sub-sample of females. However, the coefficient for men is negative and not statistically significant. Estimates for the pooled sample are 2.5 times higher in the bivariate model than the corresponding OP results, and for the female sub-sample 2.75 times greater.

Local COVID-19 deaths are negatively related to MH in the MH equation and statistically significant at the 5% level ($p = 0.037$) for the pooled sample and at the 1% level ($p = 0.007$) for females. The exclusion restriction is not significantly different to zero for men indicating a potential lack of identification in this sub-sample. Across all models, the estimate of ρ is not statistically significantly different to zero.[24]

The results from Table 4 suggest that selection on unobservables (and reverse causality) may not be a problem for this model. However, it is also clear that for the male sub-sample identification from the exclusion restriction is weak. This may also apply to the pooled and sub-sample of females. Identification via instruments (exclusion restrictions) has been studied less in the context of non-linear bivariate models than their linear counterparts (see Mourifié & Méango (2014) and Han & Vytlacil (2017) for a discussion) and it is not clear how strong an exclusion restriction in the context of the model specified in Equation (2) needs to be to attain identification.[25] In the absence of additional exclusion restrictions we now consider results from applying the approach of Altonji et al. (2005).

4.3. Coefficient stability

Table 5 presents results from estimating the OP model in the first column. This model assumes no selection on unobservables. The following five columns report estimates from the bivariate model by placing restrictions on ρ ranging from -0.2 to 0.2. Coefficient estimates, respective standard errors (in parentheses) and marginal

[24] The corresponding p-values are: $p = 0.239$ in the pooled sample; $p = 0.907$ for men, and $p = 0.107$ for women.

[25] Note that in the model for MH F-statistics for the inclusion of the exclusion restriction are 6.12, 0.11, and 7.24 respectively for the pooled, men and women samples.

effects are reported. Specifying $\rho = 0$ implies no selection of unobservables and is equivalent to the standard OP model.

It might be hypothesised that unobservables positively correlated with increasing (good) MH would also be positively correlated with higher labour productivity. This would imply a positive value of ρ. The relationship between unobservables and MH and productivity is, however, unknown a priori and it is feasible that individual characteristics or preferences that improve MH might well be negatively correlated with productivity (for example, preferences for leisure).

Assuming higher negative correlation across common unobservables in the bivariate equations increases the estimated effect of MH on productivity, such that the coefficient on GHQdiff, $\lambda = 0.071$ when $\rho = -0.2$. Higher positive values decrease the impact of MH on productivity (to $\hat{\lambda} = 0.005$ when $\rho = 0.2$). These can be contrasted against the corresponding coefficient estimate $\hat{\lambda} = 0.039$ when $\rho = 0.0$ (the standard OP model). For the pooled sample $\hat{\lambda}$ becomes insignificant for values of ρ between 0.1 and 0.2. Accordingly small positive values for ρ lead to a null effect of MH on productivity changes. Figure A1 displays in greater detail how the estimated coefficient, $\hat{\lambda}$, on MH varies with ρ. The estimate decreases monotonically with increasing values of ρ and its confidence interval crosses zero when $\rho = 0.18$. We observe a similar pattern when considering male and female sub-samples.

The final column of Table 5 assumes selection on MH based on unobservables is equivalent to that based on observables given by the condition in Equation (4). Estimates are provided for the effect of MH ($\hat{\lambda}$) and the corresponding marginal effects. For the pooled sample, the condition in Equation (4) leads to $\hat{\rho} = 0.16$. Imposing this restricting results in an estimate of λ on GHQdiff of 0.012 (s.e. 0.005). The corresponding marginal effects, while smaller than the values assuming no selection ($\rho = 0$), are statistically significant. Results by gender differ markedly. For men, assuming selection on unobservables is equivalent to that on observables results in an estimate of $\rho = 0.013$. The corresponding estimate fro GHQdiff and the implied marginal effects are virtually equivalent to the results for the standard OP (which assumes no selection on unobservables). In contrast, for women, imposing the same condition results in the estimate, $\hat{\rho} = 0.21$. This is sufficient to render the impact of MH on productivity indistinguishable from zero.

The assumption of equal selection on unobservables as selection on observables can be seen as a conservative position given the set of characteristics used as covariates $\boldsymbol{x}$.

Table 5: Effect of mental health on productivity

	June 2020: Pooled sample (N = 2,902)						
	Ordered Probit	Bivariate ordered Probit					
		Coef. (S.E.) on Mental health					
		ρ					Equal Selection
		-0.2	-0.1	0	0.1	0.2	0.16
GHQdiff	0.039	0.071	0.055	0.039	0.022	0.005	0.012
S.E.	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
	Marginal effects (S.E.) on change in productivity						
Much of little less done	-0.013	-0.024	-0.019	-0.013	-0.007	-0.002	-0.004
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Same	0.001	0.002	0.002	0.001	0.0007	0.0002	0.0004
	(0.0004)	(0.001)	(0.0006)	(0.0004)	(0.0003)	(0.0002)	(0.0002)
Much or little more done	0.012	0.022	0.017	0.012	0.007	0.002	0.004
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
	Males (N = 1,201)						
		-0.2	-0.1	0	0.1	0.2	0.013
GHQdiff	0.042	0.078	0.060	0.042	0.024	0.005	0.041
S.E.	(0.008)	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)
	Marginal effects (S.E.) on change in productivity						
Much of little less done	-0.013	-0.025	-0.019	-0.013	-0.008	-0.002	-0.013
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)
Same	0.0007	0.001	0.001	0.0007	0.0004	0.0001	0.0007
	(0.0006)	(0.001)	(0.001)	(0.0006)	(0.0004)	(0.0001)	(0.0006)
Much or little more done	0.013	0.024	0.018	0.013	0.007	0.001	0.012
	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
	Females (N = 1,701)						
		-0.2	-0.1	0	0.1	0.2	0.210
GHQdiff	0.036	0.067	0.052	0.037	0.021	0.005	0.003
S.E.	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
	Marginal effects (S.E.) on change in productivity						
Much of little less done	-0.013	-0.024	-0.018	-0.013	-0.007	-0.002	-0.0009
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Same	0.002	0.003	0.002	0.002	0.0009	0.0002	0.0001
	(0.0005)	(0.0008)	(0.0007)	(0.0004)	(0.0003)	(0.0003)	(0.0003)
Much or little more done	0.011	0.021	0.016	0.011	0.006	0.001	0.0008
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)

All regressions contain variables for sex, age, age-squared, educational attainment, ethnicity, employment, living as a couple, number of children in household, health conditions, whether work from home always, and industry of employment. Regressions are weighted using cross-sectional weights. $\rho = corr(\varepsilon_i, \epsilon_i)$. We constrain the upper limit of ρ to 0.2 since we fail to reject the null hypothesis: GHQdiff = 0 at $\rho = 0.18$ in the full sample (see Figure A1). At larger values of ρ the estimate becomes negative and counter-intuitive. For both males and females, similarly the coefficient estimate of GHQdiff is non-significant in the range: $0.1 < \rho < 0.2$. Equal selection is when the condition in Equation (4) is maintained.

We consider the true amount of selection to be lower than that observed from observables in the sample of women and that the MH effect estimated on the assumption of equal selection to be a lower bound. Estimates derived from the standard OP model, which assumes no selection on unobservables is our estimate of an upper bound.

Taken as a whole, considering results from both imposing an exclusion restriction and from assessing coefficient stability by imposing the level of selection on unobservables to be equal to observables, the results suggest that for men our OP estimate ($\lambda = 0.042$) is robust. For women, if we assume the more extreme position that selection on unobservables is equivalent to selection on observables, then the effect of MH on productivity is null. However, it is likely that selection on unobservables is less than observables, and negative deviations away from $\hat{\rho} = 0.21$ will result in a significant impact of MH on productivity, albeit modest.

4.4. Robustness checks

In this and the following section we consider robustness checks and heterogeneity analyses under the assumption of no meaningful endogeneity bias by considering OP estimates.

It might be hypothesised that the relationship between MH and productivity is confounded by working patterns and contractual arrangements for different groups of workers. For example, the self-employed may have more flexibility to factor in leave or work less demanding hours when faced with a MH issue. Alternatively, they might not feel able to take sickness leave for fear of losing business and consequently continue to work at a less productive rate. Individuals who are salaried may experience stronger labour market attachment and greater job security manifesting in a different relationship between MH and productivity than hourly paid workers. We estimate the OP model (Equation (1)) on sub-samples of employed workers and separately salaried workers. Results are presented in the first two columns of Table A5. Coefficient estimates on GHQdiff for salaried only workers (0.039) are indistinguishable from our main full-sample result. The coefficient on the employed subsample is slightly lower at 0.035, potentially indicating a stronger relationship between MH and productivity for the self-employed, however the results are not statistically distinguishable.[26]

[26]Throughout this discussion we note that the scaling of the various models will differ, affecting direct comparison of coefficients across OP models. The marginal effects, however, can be compared.

In a perfectly competitive market without frictions, worker productivity would be compensated through wage rates. Higher wage earners might also be more susceptible to productivity loses due to MH problems than lower wage earners due to being more productive prior to the onset of a period of ill health. The third column of Table A5 conditions estimation on pre-pandemic levels of wages (weekly wage). The estimated coefficient on wages in the outcome equation is 0.0003 (se 0.000091) implying that respondents with higher (lower) levels of wages were more (less) productive. However, the impact of a change in MH on changes in productivity in this specification is equivalent to our main result (0.039).

In a further check we also attempt to leave out individuals who may have misinterpreted the productivity question. Participants who reported a rise or fall in productivity were asked the reason for this change in the September wave. Responses include items such as 'lack of motivation', 'childcare/home-schooling', 'interrupted less'. We exclude individuals who reported a fall in productivity due to 'having had less work to do' and individuals who reported a rise in productivity due to 'having had more work to do' or 'not needed to commute/travel to work' because these responses may indicate changes in the number of hours worked rather than productivity per hour. In the full September module with no work from home restrictions, these categories make up 11.1% of respondents who report a fall in productivity, and 22.2% (more work to do) and 17.7% (no commute/travel) of respondents reporting a rise in productivity. Excluding these individuals has little effect on the MH coefficient in the main results using the sample of respondents from the September module.

The use of local COVID-19 deaths as an exclusion restriction in the bivariate probit models of Equation (2) is based on the assumption that deaths affect productivity only through their impact on MH conditional on the covariates included. This assumption may not hold if local deaths led to wider community-level impacts which also affected productivity. We do not believe that local deaths were at a scale to meaningfully impact productivity via local disruptions.[27] It is, however, likely that higher local deaths created a general sense of insecurity, but this should be reflected in our measure of changes to MH. If a family member or close neighbour passed away or was severely impacted by COVID-19, which directly affected a respondent's productivity,

[27] Public Health England suggest that in England there were 72,178 deaths (in laboratory confirmed cases of COVID-19) in 2020 in a population of approximately 56 million (PHE 2023).

this would also potentially lead to a lack of validity in the instrument. To investigate this issue, we exploit data within the June module of the UKHLS that contains information on whether, and if so how, a respondent is impacted by the pandemic. We construct a set of binary variables that represent whether the respondent: has had symptoms that could be coronavirus; had symptoms that could be coronavirus at interview; sought medical help for COVID-19 symptoms; very likely had coronavirus[28]; had contact with a COVID-19 case (close contact in the last two weeks); had household members who had coronavirus symptoms; had household members who tested for coronavirus; or whether the respondent had tested for coronavirus. Incorporating the above information into the ordered probit model as additional covariates, results in no discernible change to the estimated coefficient on GHQdiff. This is reported in Table A5. The same conclusion holds for the bivariate probit model with deaths included as an exclusion restriction when these additional variables are included.[29] The coefficient on GHQdiff remains unaltered, and none of the COVID-related variables are statistically significant (at 5%) in the productivity outcome equation.

We investigate the impact of MH on productivity using the COVID modules of UKHLS. It is possible that the changes in the GHQ scores from pre-pandemic to June 2020 were influenced by particular component items of the GHQ that, in turn, were more responsive to the pandemic than might be experienced more generally (see also Serrano-Alarcon et al. 2022). To investigate this Table A6 displays the overall change in the GHQ score observed in our estimation samples together with changes in its component items. All changes are negative implying a worsening of MH. The largest mean changes are observed for the component items: 'enjoy day-to-day activities'; 'loss of sleep, problems overcoming difficulties'; and 'constantly under pressure'. The first of these is associated with the largest change both overall and by gender. It might be argued that this item is the most likely to be influenced by disruptions to peoples' lives that the pandemic caused. To assess whether this component of GHQdiff is driving our results, we construct a version of GHQdiff which excludes this item.[30]

[28]Constructed from questions on whether tested positive for COVID-19, or the likelihood of having had COVID-19 given had symptoms and not tested, or tested and inconclusive.

[29]Results available on request.

[30]Accordingly, instead of GHQ having a range of 0-36 with 12 items, it ranges from 0-33 with 11 items. GHQdiff is then calculated as the change from pre-pandemic to June 2020 in the constructed score.

The results for the full sample are reported in column (5) of Table A5; they are close to those reported in Table 3.[31] This provides reassurance that the results reported are not driven by component items of GHQ that were particularly susceptible to the pandemic.

4.5. Heterogeneity analyses

Table A7 considers heterogeneity in the relationship between changes in MH and changes in productivity by considering different ages of individuals and by MH status at baseline. These are compared for the full sample of respondents who reported working from home at least sometimes in the June wave. The results indicate that the relationship between MH and productivity is marginally stronger for older compared to younger workers (less than 50 years) for women, but the opposite for men. However, the magnitudes of the differentials are not particularly notable and given the reduced sample sizes and accordingly precision upon which these estimates are based, the differences are not statistically significantly different.

The final two columns of Table A7 present results separately for individuals with good and poor MH at baseline. These categories are defined by the caseness score for GHQ. Each of the 12 items for the GHQ is scored on a four point scale (0-3), such that the overall measure ranges from 0 to 36, with 36 (on the original scale) being most distressed. The caseness scale recodes values of 0 and 1 on each of the 12 items to 0 and values of 2 and 3 to 1, resulting in a scale from 0 (least distressed) to 12 (most distressed). We use a cut-off of 4 and above to represent poor initial MH, and less than 4 as good initial MH.[32] Using this threshold, the estimated relationship between the change in MH and productivity is slightly larger for individuals who report good MH at baseline than for those reporting poor MH, but the differences are small.

The September wave of the COVID questionnaire also contained the productivity question. To be consistent with the June wave, the benchmark used to report changes in productivity was also January/February of 2020. This was asked of all respondents

[31] Ordered probit coefficient estimates for the male and female sub-samples using this version of GHQdiff are 0.044 (0.009) and 0.039 (0.006) respectively. Again, these are close to the main results presented in Table 3.

[32] The National Health Service in England uses a cutoff of 4 as the threshold to monitor the percentage of people suffering from poor MH and an indicator of possible psychiatric disorder in the general population, see https://files.digital.nhs.uk/BA/46AF8E/Spec_03J_321VSP2_10_V1.pdf. In our sample, 19% have a baseline MH score greater than the cutoff.

as long as they were employed, self-employed or both. There was no restriction on respondents working from home.[33] We further restrict the sample to individuals who had not been furloughed. The September wave omitted to collect information on industry sector and accordingly, we include the set of conditioning variables used in the June analysis with the exception of these. Table A8 reports these results.[34] The effects of MH on productivity remain positive and statistically significant. However, they are lower compared to the corresponding estimate for June reported in Table 3. Effects for males remain larger than for females, but differences are not statistically significant. While the pandemic restrictions were more relaxed in September compared to June 2020, the relationship between changes in MH and changes in productivity persists.

Taken as a whole, robustness and heterogeneity analyses do not reveal strong evidence of differential effects.

4.6. Quantification of productivity changes

A follow-up to the change in productivity question aimed at quantifying the gains or losses reported was included in the September module. This was in the form of the question "Thinking about how much less (more) you get done these days, would you say that what you can do in an hour now would previously have taken you: ..." Table 6 summarises the responses on the 3-point scale used in the paper.[35] The sample consists of 3,601 respondents, of which 2,179 report no change in productivity. Of the remaining 1,421 respondents, 512 report a fall in productivity (367 "A little less" and 145 "Much less"), and 909 report a rise (477 "A little more" and 432 "Much more").

The second (sixth) column of Table 6 summarises the fall (rise) in productivity in minutes per hour by taking the midpoint of the reported category. For example, if prior to the pandemic it took between 45 and 60 minutes to complete what could be completed at the September wave in one hour, then the estimated productivity loss is 8 minutes (approximate midpoint between a loss of 15 minutes and 0 minutes). We do likewise for all other categories including responses to a rise in productivity.[36]

[33]By September the UK Government advice on working-from-home had relaxed since the June wave and accordingly it would have been less relevant to only question respondents who were at least working form home some of the time.

[34]For comparison omitting industry information from the June wave results in little difference in the estimate for GHQdiff of 0.038 (0.005).

[35]Note Table A2 provides the responses on the original 5-point scale.

[36]The final category for a rise in productivity where the response is "More than an hour and a

This enables us to compute the estimated loss or gain in productivity per hour. For example, for respondents who report that they get less done, the loss in minutes is $[(230 \times 8) + (183 \times 23) + (89 \times 45)]/502 = 20.03$ minutes per hour. The gain for those reporting much or a little more done is 19.84 minutes per hour. For men reporting a decrease in productivity, the estimated loss is 17.7 minutes per hour, and for women it is 21.6 minutes.

A unit decrease (increase) in GHQ leads to an increase in the probability of reporting getting less (more) done of 0.013 (0.012) (Table 3). This equates to an estimated change in productivity of approximately 16 seconds per hour ($20mins \times 0.013$), or 2.6 minutes a day (assuming a 7.5 hour work day). This figure appears small, but considering that the average change in the GHQ score observed in the sample is -1.675, then across the full sample this equates to an aggregate expected loss of productivity of 1,266 minutes for every hour worked.[37] This is equivalent to 791 hours[38] (106 days) per week in lost productivity across the full sample of 2,902 individuals. If this sample were representative of the population of workers in June 2020, then total productivity losses would have been substantial.

If we focus on the sub-sample of men (for which estimates appear robust across the set of modelling approaches) the expected loss in productivity for a unit decrease in the GHQ score is approximately 1.73 minutes a day.[39] The average change in the GHQ score observed for men is, however, smaller than the pooled sample at -1.477, which equates to an aggregate loss of productivity of 408 minutes for every hour worked. Across the sample of men, this computes to a total of 255 hours (34 days) per week, or 13 minutes per individual per week in lost productivity.

A similar calculation for women reveals greater productivity losses than for men totalling 551 hours (73 days) per week across the sample, or approximately 19 minutes per week per person. This estimate is based on the relationship between MH and productivity from the ordered probit model reported in Table 3 and while representing an upper bound is useful for comparative purposes.

The observed average change in MH across the samples do not, however, reflect the

half", is truncated at the upper limit at an hour and three-quarters when estimating the midpoint.

[37] $1.675 \times 0.013 \times 20.03 \times 2902 = 1,266$ minutes.

[38] $\frac{1266 \times 37.5}{60} = 791$.

[39] $17.7 \times 0.013 \times 7.5 = 1.73$.

Table 6: September wave: Quantification of changes in productivity

Fall in productivity	Mins	Little or much less Freq.	%	*Rise in productivity*	Mins	Little or much more Freq.	%
Pooled sample							
Don't know		10	1.95	Don't know/refusal		12	1.32
Between 45 and 60 minutes	8	230	44.92	Up to 75 minutes	8	400	44.00
Between 30 and 45 minutes	23	183	35.74	Between 75 and 90 minutes	23	353	38.83
Less than 30 minutes	45	89	17.38	More than 90 minutes	45	144	15.84
N		512	100	N		909	100
Males							
Don't know		4	1.93	Don't know/refusal		4	1.16
Between 45 and 60 minutes	8	110	53.14	Up to 75 minutes	8	169	48.99
Between 30 and 45 minutes	23	67	32.37	Between 75 and 90 minutes	23	131	37.97
Less than 30 minutes	45	26	12.56	More than 90 minutes	45	41	11.88
N		207	100	N		345	100
Females							
Don't know		6	1.97	Don't know/refusal		8	1.42
Between 45 and 60 minutes	8	120	39.34	Up to 75 minutes	8	231	40.96
Between 30 and 45 minutes	23	116	38.03	Between 75 and 90 minutes	23	222	39.36
Less than 30 minutes	45	63	20.66	More than 90 minutes	45	103	18.26
N		305	100	N		564	100

Responses to the question "Thinking about how much less (more) you get done these days, would you say that what you can do in an hour now would previously have taken you:"

extent of decline at an individual level. For example, for women (men) who report a decrease in MH, the average decline is -5.737 (-5.117). At an individual level these changes in MH on the GHQ scale will lead to a loss of productivity equating to approximately one hour (three quarters of an hour) a week for a female (male), or 54 minutes across both males and females. These appear meaningful declines in productivity at the individual level.

The above estimates point to differences by gender in estimated productivity losses. Two factors contribute to this difference. First, the average decline in MH is greater for females than males. Second, when quantifying productivity losses in terms of time lost a higher proportion of women than men report larger losses. For example, of respondents who reported getting less done, 38% (21%) of women compared to 32% (13%) of men report that what they achieve in a hour would previously have taken between 30 and 45 minutes (less than 30 minutes) (see Table 6). Estimates for women, however, are best viewed as an upper bound, whereas results for men appear more robust to estimation. A caveat to these estimates is that the quantification of productivity losses reported in Table 6 is based on the question fielded in the September module of the questionnaire and applied here to productivity changes observed in the June module.[40] Given that MH declines in September where not as great as those observed in the June wave, this is likely to lead to more conservative estimates of time losses.

5. Discussion and conclusion

The World Health Organisation (in 2019) estimated an annual cost of depression and anxiety to the global economy of around US$1 trillion in lost productivity.[41] MH problems can make it difficult to carry out day-to-day work activities, leading to decreased efficiency and reduced capacity to work. For example, depression has been found to limit physical job functioning 20% of the time, and to impair cognitive performance at least 35% of the time (Lerner & Henke 2008). The prevalence of common mental disorders is high in the UK, with one in six people reporting these in 2014, and has been increasing since the early 1990s (Baker 2021). If this trend continues, we anticipate MH will be an increasing factor in explaining worker productivity. Given the well-documented low levels of productivity in the UK, understanding the link between MH and productivity is a particularly important aspect of managing the UK economy and for future labour market policy. Our analysis provides much needed evidence that quantifies the effect of MH on productivity.

Unlike most studies that rely on wages to proxy productivity, we exploit unique data, which includes a direct measure of productivity change. We find that a change in MH has a statistically significant (albeit modest) effect on a change in individual productivity. Deteriorating MH leads to a higher probability of getting less done at work relative to baseline. The opposite is true for improvements in MH. Despite women reporting a greater reduction in both MH and productivity relative to the pre-COVID period, we find no evidence that the effect differs statistically by gender.

Using additional self-reported data on how much more or less each individual got done, we estimate that a deterioration in MH of 1.675 GHQ units (the observed average decline in these data) led to an aggregate expected loss of 2,531 minutes for every hour worked across the sample of respondents in this survey. If these estimates are applicable to the UK population of workers, the implied productivity losses are

[40] Note that quantification of productivity losses was not fielded in the June module.

[41] https://www.who.int/news-room/commentaries/detail/mental-health-in-the-workplace

substantial. We expect the effects to be much larger if one considers changes in MH that lead to diagnosed conditions. This link between MH and productivity strengthens the case for public policy to invest in MH prevention and treatment programmes. Not only will such measures improve population well-being directly through improved health, but will also enhance productivity, and hence, indirectly lead to further well-being gains through increased living standards.

References

Altonji, J., Elder, T. & Taber, C. (2002*a*), An evaluation of instrumental variable strategies for estimating the effects of catholic schools, Technical report, NBER Working Paper no. 9358.

Altonji, J., Elder, T. & Taber, C. (2002*b*), Selection on observed and unobserved variables: assessing the effectiveness of catholic schools, Technical report, Manuscript, Dept. Economics, Northwestern University.

Altonji, J., Elder, T. & Taber, C. (2005), 'Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools', *Journal of Political Economy* **113**(1), 151–184.

Anjara, S., Bonetto, C., Van Bartel, T. & Brayne, C. (2020), 'Using the GHQ-12 to screen for mental health problems among primary care patients: psychometrics and practical considerations', *International Journal of Mental Health Systems* **15**, 62.
URL: *https://doi.org/10.1186/s13033-020-00397-0*

Baker, C. (2021), 'Mental health statistics (England)', *UK House of Commons Library* **CBP-06988**.

Banerjee, S., Chatterji, P. & Lahiri, K. (2017), 'Effects of psychiatric disorders on labour market outcomes: A latent variable approach using multiple clinical indicators', *Health Economics* **26**(2), 184–205.

Banks, J. & Xu, X. (2020), 'The mental health effects of the first two months of lockdown during the COVID-19 pandemic in the UK', *Fiscal Studies* **41**(3), 685–708.

Black, S., Devereux, P. & Salvanes, K. (2016), 'Does grief transfer across generations? Bereavements during pregnancy and child outcomes', *American Economic Journal: Applied Economics* **8**(1), 193–223.

Bryan, M., Bryce, A. & Roberts, J. (2021), 'The effects of mental and physical health problems on sickness absence', *The European Journal of Health Economics* **22**, 1519–1533.

Bryan, M., Rice, N., Roberts, J. & Sechel, C. (2022), 'Mental health and employment: A bounding approach using panel data', *Oxford Bulletin of Economics and Statistics* **84**(5), 1018–1051.

Bubonya, M., Cobb-Clark, D. A. & Wooden, M. (2017), 'Mental health and productivity at work: does what you do matter?', *Labour Economics* **46**, 150–165.

Burton, W. N., Chen, C. Y., Conti, D. J., Schultz, A. B., Pransky, G. & Edington, D. W. (2005), 'The association of health risks with on-the-job productivity', *Journal of Occupational and Environmental Medicine* **47**(8), 769–777.

Chatterji, P., Alegria, M. & Takeuchi, D. (2011), 'Psychiatric disorders and labour market outcomes: Evidence from the National Comorbidity Survey - replication', *Journal of Health Economics* **30**, 858–868.

Cheng, Z., Mendolia, S., Paloyo, A. R., Savage, D. A. & Tani, M. (2021), 'Workig parents, financial insecurity, and childcare: mental health in the time of COVID-19 in the UK', *Review of Economics of the Household* **19**, 123–144.

Contoyannis, P. & Rice, N. (2001), 'The impact of health on wages: Evidence from the british household panel survey', *Empirical Economics* **26**(4), 599–622.

Daly, M., Sutin, A. R. & Robinson, E. (2020), 'Longitudinal changes in mental health and the COVID-19 pandemic: evidence from the UK Household Longitudinal Study', *Psychological Medicine* pp. 1–10. https:doi:10.1017/S0033291720004432.

De Oliveira, C., Saka, M., Bone, L. & Jacobs, R. (2023), 'The role of mental health on workplace productvity: a critical review of the literature', *Applied Health Economics and Health Policy* **21**, 167–193.

Dujardin, C. & Goffette-Nagot, F. (2010), 'Neighborhood effects on unemployment? A test *á la* Altonji', *Regional Science and Urban Economics* **40**, 380–396.

Emmerson, C., Johnson, P. & Miller, H. (2013), 'The IFS Green Budget'. https://ifs.org.uk/green-budget/2013.

Frijters, P., Johnston, D. W. & Shields, M. A. (2014), 'The effect of mental health on employment: evidence from Australian panel data', *Health Economics* **23**, 1058–1071.

Garcia-Gomez, P., Jones, A. M. & Rice, N. (2010), 'Health effects on labour market exits and entries', *Labour Economics* **17**, 62–76.

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., Oldehinkel, T. & Ormel, J. (1998), 'Why GHQ threshold varies from one place to another', *Psychological Medicine* **28**(4), 915–921.

Greene, W. & Hensher, D. (2010), *Modeling ordered choices*, Cambridge University Press.

Han, S. & Vytlacil, E. (2017), 'Identification in a generalization of bivariate probit models with dummy endogenous regressors', *Journal of Econometrics* **199**, 63–73.

House of Commons (2023), Health assessments for benefits. Fifth report of session 2022-12, Technical report, House of Commons Work and Pensions Committee.

ISER (2021), 'Understanding Society: COVID-19 Study, 2020-2021. [data collection]. 11th Edition. UK Data Service'. SN: 8644, 10.5255/UKDA-SN-8644-11; Accessed May 16, 2022 4:10 PM.

ISER (2022), 'Understanding Society: Waves 1-11, 2009-2020 and harmonised BHPS: Waves 1-18, 1991-2009. [data collection]. 15th Edition. UK Data Service'.

Krugman, P. (1994), *The Age of Diminished Expectations*, The MIT Press, Cambridge, Massachusetts.

Lagomarsino, E. & Spiganti, A. (2020), 'No gain in pain: psychological well-being, participation, and wages in the BHPS', *European Journal of Health Economics* **21**, 1375–1389.

Layard, R. (2013), 'Mental health: the new frontier for labour economics', *Journal of Labour Policy* **2**(2), 1–17.

Le, H. & Nguyen, H. (2018), 'The impact of maternal mental health shocks on child health', *American Journal of Health Economics* **4**(2), 185–225.

Lerner, D. & Henke, R. (2008), 'What does research tell us about depression, job performance, and work productivity?', *Occupational and Environmental Medicine* **50**(4), 401–410.

Madden, D. (2010), 'Gender differences in mental well-being: a decomposition analysis', *Social Indicators Research* **99**, 101–114.

Mavridis, D. (2015), 'The unhappily unemployed return to work faster', *Journal of Labor Economics* **4**(1), 1–22.

Meyer, B. D. & Mok, W. K. C. (2019), 'Disability, earnings, income and consumption', *Journal of Public Economics* **171**, 51–69.

Mincer, J. (1958), 'Investment in human capital and personal income distribution', *Journal of Political Economy* **66**(4), 281–302.

Mourifié, I. & Méango, R. (2014), 'A note on the identification in two equation probit model with dummy endogenous regressor', *Economics Letters* **125**, 360–363.

Orefice, S. & Quintana-Domeque, C. (2021), 'Gender inequality in COVID-19 time: evidence from UK prolific participants', *Journal of Demographic Economics* **87**, 261–87.

Oster, E. (2013), Unobservable selection and coefficient stability: theory and validation, Technical Report No. 19054, NBER Working Paper.

Oster, E. (2019), 'Unobservable selection and coefficient stability: theory and evidence', *Journal of Business & Economic Statistics* **37**(2), 187–204.

Oswald, A. J., Proto, E. & Sgroi, D. (2015), 'Happiness and productivity', *Journal of Labor Economics* **33**(4), 789–822.

Persson, P. & Rossin-Slater, M. (2018), 'Family ruptures, stress, and the mental health of the next generation', *American Economic Review* **108**(4-5), 1214–1252.

Pessoa, J. & Van Reenen, J. (2014), 'The UK productivity and jobs puzzle: does the answer lie in wage flexibility?', *Economic Journal* **124**(576), 433–452.

PHE (2023), Covid-19 confirmed deaths in England (to 31 December 2020): report, Technical report, Public Health England COVID-19 Epidemiology Cell.
URL: *https://www.gov.uk/government/publications/covid-19-reported-sars-cov-2-deaths-in-england/covid-19-confirmed-deaths-in-england-to-31-december-2020-report*

Roberts, J., Bryan, M., Bryce, A., Rice, N. & Sechel, C. (2021), 'Disability Employment Gap: Written evidence to the Work and Pensions Committee Inquiry 2020'. https://committees.parliament.uk/writtenevidence/19219/pdf/.

Sartorius, N. & Ustün, T. B. (1995), *Mental Illness in General Health Care: an International Study*, John Wiley & Sons, New York.

Schlechter, P. & Neufeld, S. (2024), 'Longitudinal and gender measurement invariance of the general health questionnaire-12 (ghq-12) from adolescence to emerging adulthood', *Assessment* **31**(8), 1687–1701.

Schmitz, N., Kruse, J. & Tress, W. (1999), 'Psychometric properties of the general health questionnaire (GHQ-12) in a German primary care sample.', *Acta Psychiatrica Scandinavica* **100**(6), 462–468.

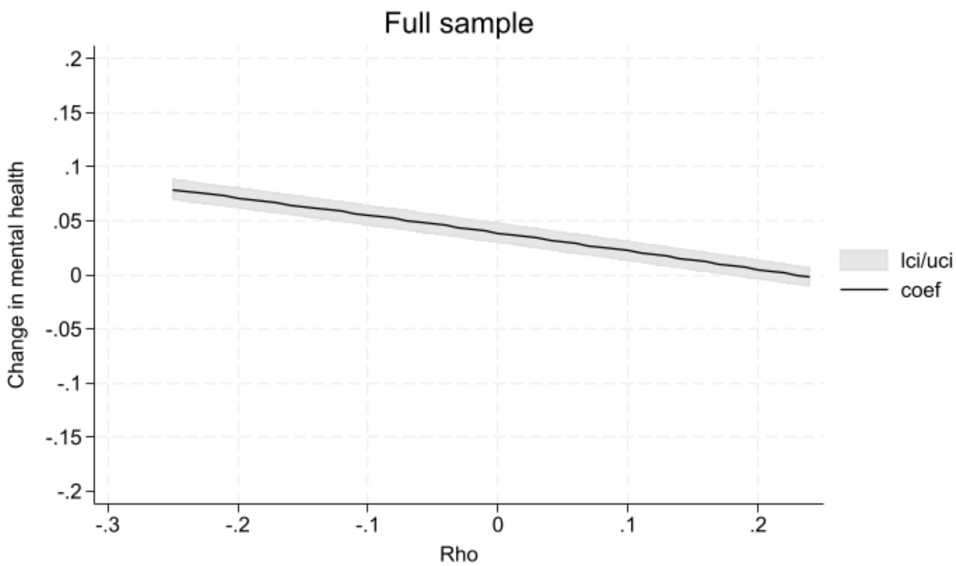
Serrano-Alarcon, M., Kentikelenis, A., McKee, M. & Stuckler, D. (2022), 'Impact of COVID-19 lockdowns on mental health: evidence from a quasi-natural experiment in England and Scotland', *Health Economics* **31**, 284–296.

Sharpe, A. & Mobasher Fard, S. (2022), ‘The current state of research on the two-way linkages between productivity and well-being’, ILO Working Paper 56, (Geneva, ILO).

von Hinke, S., Rice, N. & Tominey, E. (2022), ‘Mental health around pregnancy and child development from early childhood to adolescence’, *Labour Economics* **78**, 102245.

Appendix

Figure A1: Bivariate ordered probit: estimated coefficient on MH when varying ρ



Estimated coefficient on MH from a bivariate probit regression (Equations (3)) on the full sample by varying the selection parameter ρ. coef is the estimated coefficient, lci is the lower confidence interval, and uci is the upper confidence interval. We fail to reject $H_0 : MH = 0$ when $\rho \leq 0.18$.

Table A1: Variable definitions

Variable	Definition
Dependent variable	
Productivity change	”Please think about how much work you get done per hour these days. How does that compare to how much you would have got done per hour back in January/February 2020?” Possible answers: ”much less”, ”a little less”, ”same”, ”a little more”, ”much more”. We group into three categories: 1 - much or a little less, 2 - same, 3 - much or a little more.
Explanatory variables	
GHQdiff	Change in GHQ score between relevant COVID module and pre-COVID baseline. The original GHQ score is measured on a scale ranging from 0 to 36 where higher values denote worse health. We reverse code so that higher values denote better health.
Health condition	Dummy variable = 1 if individual has a long term health condition at the time of interview.
Female	Dummy variable = 1 if individual is female.
Age	Age of respondent in years at the time of interview.
Ethnicity	Dummy variable = 1 if ethnicity of individual is non-white.
Degree	Dummy variable = 1 if highest qualification attained is degree or equivalent.
Employment status	Three categories: employed (omitted), self-employed, both employed and self-employed.
Couple	Dummy variable = 1 if living as a couple in one household.
Kids 0-4	Dummy variable = 1 if there are kids aged 0-4 living in the household.
Kids 5-15	Dummy variable = 1 if there are kids aged 5-15 living in the household.
Kids 16-18	Dummy variable = 1 if there are kids aged 16-18 living in the household.
Industry	Industry dummies UK Standard Industrial Classification 2007 (omitted category Education).
Instrumental variable	
COVID-19 deaths	Number of COVID-19 related deaths in each Middle Layer Super Output Area (MSOA) occurring between 1 March and 31 July 2020

The table reports definitions of the key variables used throughout the analysis. Health condition is a binary variable representing whether a respondent has one or more of 22 listed health conditions. All conditions with the exception of one relate to physical health. Results reported in Table 3 are indifferent to the removal of the single condition relating to an emotional, nervous or psychiatric problem. Industry information is not available in the September 2020 module due to a routing error.

Table A2: September wave: Quantification of changes in productivity

		September 2020					
Fall in productivity	Midpoint loss	A little less		Much less		Little or much less	
		Freq.	%	Freq.	%	Freq.	%
Don't know	-	3	0.82	7	4.83	10	1.95
Between 45 and 60 minutes	8	202	55.04	28	19.31	230	44.92
Between 30 and 45 minutes	23	139	37.87	44	30.34	183	35.74
Less than 30 minutes	45	23	6.27	66	45.52	89	17.38
N		367	100	145	100	512	100
Rise in productivity	Midpoint gain	A little more		Much more		Little or much more	
		Freq.	%	Freq.	%	Freq.	%
Refusal	-	1	0.21	0	0	1	0.11
Don't know	-	5	1.05	6	1.39	11	1.21
Up to 75 minutes	8	268	56.18	132	30.56	400	44.00
Between 75 and 90 minutes	23	163	34.17	190	43.98	353	38.83
More than 90 minutes	45	40	8.39	104	24.07	144	15.84
N		477	100	432	100	909	100

The overall sample size is 3,601. 2,180 (60.5%) individuals reported no change in productivity. The question asked of respondents who reported a change in productivity is; top (bottom) panel: "Thinking about how much less (more) you get done these days, would you say that what you can do in an hour now would previously have taken you:" Note that the responses were provided as hours for 60 minutes and above. For example, between 75 minutes and 90 minutes was presented as being "Between an hour and a quarter and an hour and a half". Likewise for other responses. These are presented in minutes in the above to conserve text space.

Table A3: Industry summary statistics (June 2020)

	Full sample		Males		Females	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Industry						
Agriculture/forestry/fishing/mining/quarrying	0.010	0.099	0.011	0.105	0.009	0.093
Manufacturing	0.032	0.176	0.051	0.220	0.016	0.126
Utilities	0.025	0.157	0.030	0.170	0.021	0.145
Construction	0.039	0.193	0.055	0.229	0.025	0.155
Wholesale/retail	0.041	0.197	0.040	0.196	0.041	0.199
Repair of motor vehicles/transportation/storage	0.020	0.138	0.022	0.146	0.018	0.131
Accommodation/Food/Other services/households as employers	0.115	0.319	0.123	0.328	0.108	0.310
Information & communication	0.081	0.273	0.121	0.326	0.047	0.211
Financial & insurance	0.091	0.287	0.101	0.301	0.082	0.275
Real estate	0.014	0.118	0.019	0.137	0.010	0.098
Professional/scientific/technical	0.080	0.271	0.101	0.302	0.061	0.240
Admin/support services	0.048	0.213	0.021	0.144	0.070	0.256
Public administration & defense	0.063	0.243	0.082	0.275	0.046	0.210
Education	0.203	0.402	0.124	0.329	0.270	0.444
Human health/social work	0.097	0.296	0.060	0.237	0.129	0.335
Arts/entertainment/recreation	0.044	0.204	0.040	0.196	0.047	0.211
N	2,902		1,201		1,701	

Sample summary statistics weighted using cross-sectional weights.

Table A4: Effect of MH on productivity in June 2020, by gender. Full results

	Full sample Coef (s.e.)	Males Coef (s.e.)	Females Coef (s.e.)
GHQdiff	0.039	0.042	0.037
	(0.005)	(0.008)	(0.006)
Health condition	0.037	-0.051	0.097
	(0.057)	(0.084)	(0.078)
Socio-demographic			
Male	0.023		
	(0.062)		
Age	0.041	0.043	0.039
	(0.017)	(0.024)	(0.023)
Age^2	-0.0004	-0.0004	-0.0004
	(0.0002)	(0.0003)	(0.0002)
Degree	0.008	-0.158	0.114
	(0.068)	(0.102)	(0.083)
Ethnicity	0.133	0.100	0.114
	(0.117)	(0.190)	(0.183)
Employment			
Self-employed	-0.369	-0.410	-0.363
	(0.084)	(0.123)	(0.110)
Both self & employed	-0.163	-0.045	-0.268
	(0.151)	(0.242)	(0.190)
Household			
Couple	0.115	0.133	0.086
	(0.075)	(0.120)	(0.089)
Kids 0-4 yo	0.016	0.104	-0.084
	(0.096)	(0.143)	(0.127)
Kids 5-15 yo	-0.192	-0.145	-0.242
	(0.067)	(0.102)	(0.089)
Kids 16-18 yo	-0.020	0.068	-0.089
	(0.085)	(0.140)	(0.102)
Industry (Education omitted)			
Agriculture/Forestry/Fishing/Mining/Quarrying	0.255	-0.057	0.539
	(0.159)	(0.220)	(0.218)
Manufacturing	0.330	0.206	0.345
	(0.133)	(0.191)	(0.213)
Utilities	0.257	0.229	0.146
	(0.182)	(0.247)	(0.269)
Construction	0.290	0.304	0.015
	(0.253)	(0.194)	(0.588)
Wholesale/Retail	0.641	0.626	0.593
	(0.184)	(0.291)	(0.236)
Repair of Motor Vehicles/Transportation/Storage	0.552	0.259	0.729
	(0.187)	(0.263)	(0.283)
Accommodation/Food services/Other service activities/HH as employers	0.381	0.257	0.386
	(0.109)	(0.172)	(0.147)
Information & communication	0.343	0.280	0.285
	(0.126)	(0.190)	(0.164)
Financial & Insurance	0.331	0.234	0.294
	(0.113)	(0.183)	(0.147)
Real estate	0.236	0.210	0.108
	(0.174)	(0.246)	(0.259)
Professional/Scientific/Technical	0.233	0.114	0.258
	(0.112)	(0.175)	(0.153)
Admin/Support services	0.030	-0.169	0.103
	(0.124)	(0.258)	(0.141)
Public administration & defense	-0.007	-0.258	0.186
	(0.151)	(0.239)	(0.148)
Human health/Social work	0.489	0.237	0.572
	(0.100)	(0.213)	(0.111)
Arts/Entertainment/Recreation	-0.146	-0.421	0.012
	(0.149)	(0.224)	(0.193)
Always work from home	0.207	0.211	0.240
	(0.064)	(0.093)	(0.084)
N	2,902	1,201	1,701

Full results corresponding to Table 3. Dependent variable is change in productivity. Regressions weighted using cross-sectional weights.

Table A5: Robustness checks: June 2020 sample

	Employed Coef (s.e.) (1)	Salaried Coef (s.e.) (2)	With baselines wages Coef (s.e.) (3)	COVID contact Coef (s.e.) (4)	Alternative GHQdiff Coef (s.e.) (5)
Ordered Probit - coefficient estimates on a latent scale					
GHQdiff	0.036***	0.039***	0.039***	0.039***	0.041***
	(0.005)	(0.006)	(0.005)	(0.005)	(0.005)
Marginal effects (S.E.) on change in productivity					
Much of little less done	-0.012	-0.013	-0.013	-0.013	-0.014
	(0.002)	(0.002)	(0.005)	(0.002)	(0.002)
Same	0.0005	0.0005	0.001	0.001	0.001
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Much or little more done	0.013	0.012	0.012	0.012	0.012
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)
N	2,314	2,096	2,646	2,902	2,902

* p<0.10, ** p<0.05, *** p<0.01. Regressions reported in columns (1) and (2) contain controls for sex, age, age-squared, educational attainment, ethnicity, employment, living as a couple, number of children in household, health conditions, whether always worked from home, and industry of employment. Regression results (3) also contain baseline wages. Column (4) includes an additional set of controls to represent the extent of contact with COVID virus either directly, within the household, or through contact with another individual with COVID-19. Column (5) reports results from constructing a measure of GHQdiff without the item representing 'Enjoy day-to-day activities'. Regressions are weighted using cross-sectional weights.

Table A6: Breakdown of changes to GHQ score by changes in component item: June 2020

	Full sample		Males		Females	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Productivity change						
GHQdiff	-1.676	6.113	-1.477	5.605	-1.846	6.513
Component items						
Concentration	-0.154	0.724	-0.137	0.692	-0.168	0.751
Loss of sleep	-0.204	0.877	-0.175	0.842	-0.229	0.905
Playing a useful role	-0.087	0.712	-0.078	0.651	-0.095	0.761
Capable of making decisions	-0.099	0.619	-0.061	0.567	-0.132	0.659
Constantly under pressure	-0.161	0.850	-0.159	0.824	-0.163	0.872
Problems overcoming difficulties	-0.167	0.833	-0.158	0.787	-0.175	0.870
Enjoy day-to-day activities	-0.223	0.787	-0.188	0.743	-0.254	0.822
Ability to face problems	-0.086	0.553	-0.090	0.506	-0.082	0.590
Unhappy or depressed	-0.140	0.907	-0.145	0.847	-0.137	0.955
Losing confidence	-0.105	0.867	-0.085	0.811	-0.122	0.912
Believe worthless	-0.152	0.778	-0.130	0.692	-0.170	0.845
General happiness	-0.097	0.754	-0.070	0.729	-0.120	0.774
N	2,902		1,201		1,701	

Sample summary statistics weighted using cross-sectional weights.

Table A7: Heterogeneity Analysis: June 2020 sample

	All		Men		Women		All	
	Under 50	50 plus	under 50	50 plus	under 50	50 plus	MH good	MH bad
	Coef (s.e)	Coef (s.e)	Coef (s.e)	Coef (s.e)	Coef (s.e)	Coef (s.e)	Coef (s.e)	Coef (s.e)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ordered Probit - coefficient estimates on a latent scale								
GHQdiff	0.037 ***	0.045 ***	0.048 ***	0.038 ***	0.033 ***	0.049 ***	0.049 ***	0.041 ***
	(0.006)	(0.008)	(0.011)	(0.012)	(0.008)	(0.010)	(0.007)	(0.009)
Marginal effects (S.E.) on change in productivity								
Much of little less done	-0.012	-0.012	-0.016	-0.010	-0.012	-0.015	-0.016	-0.014
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)
Same	0.003	0.0002	0.001	-0.0003	0.002	0.0008	0.001	0.002
	(0.0005)	(0.0006)	(0.001)	(0.0008)	(0.0006)	(0.0009)	(0.0006)	(0.001)
Much or little more done	0.012	0.013	0.014	0.011	0.010	0.015	0.015	0.012
	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)
N	1,534	1,368	585	616	949	752	2,349	553

* p<0.10, ** p<0.05, *** p<0.01. Regressions reported in columns (1) and (2) contain controls for sex, age, age-squared, educational attainment, ethnicity, employment, living as a couple, number of children in household, health conditions, whether always worked from home, and industry of employment. Regressions are weighted using cross-sectional weights.

Table A8: Effect of MH on productivity in September 2020, by gender

	September 2020		
	Pooled	Males	Females
	Coef	Coef	Coef
	(S.E.)	(S.E.)	(S.E.)
Ordered Probit - coefficient estimates on a latent scale			
GHQdiff	0.028 ***	0.035 ***	0.024 ***
	(0.006)	(0.010)	(0.007)
N	3,600	1,402	2,198
Marginal effects on change in productivity			
Much or little less done	-0.006 ***	-0.008 ***	-0.005 ***
	(0.001)	(0.002)	(0.002)
Same	-=0.002 ***	-0.003***	0.002 **
	(0.0006)	(0.001)	(0.0006)
Much or little more done	0.009 ***	0.011 ***	0.007 ***
	(0.002)	(0.003)	(0.002)
N	3,600	1,402	2,198

* p<0.10, ** p<0.05, *** p<0.01. All regressions contain variables for age, age-squared, educational attainment, ethnicity, employment, living as a couple, number of children in household, health conditions, and whether always working from home. Regressions are weighted using cross-sectional weights.