Assessing the impact of health, qualifications and wages on economic inactivity

Jon Minton

Martin Taulbut

Neil Chalmers

Debs Shipton

# Abstract

**Introduction** The UK has comparatively high rates of working age economic inactivity (EI) due to poor health.

**Methods** This paper uses a novel modelling framework, and data from the UKHLS, to estimate how much of this health-related EI (HREI) may be ‘due to’ i) poor general health directly, or due to two types of socioeconomic driver of health and inactivity: ii) low qualifications; and iii) low wages. We modelled the individual transition probabilities between seven economic (in)activity states for a broadly representative population from the UKHLS observed pre-pandemic, who were between 25 and 64 years of age when interviewed. We projected how many move between states both under observed circumstances and in a scenario where an adverse driver was fully or largely mitigated, and compared how many more or fewer people were in the economically inactive long term sick state in both scenarios. From this we were able to produce estimates of the amount of HREI that may be ‘explained by’ these specific drivers.

**Findings** Considering each driver independently, poor general health may explain up to 30% of HREI, whereas low wages may explain up to UNKNOWN%, and low qualifications may explain up to 17.5% of HREI.

**Discussion** The framework provides a means of estimating the role that such drivers as poor health and low qualifications have on the size of the HREI population. Further work will be required to consider additional drivers, and the effect of multiple drivers in combination. Further work is also required to identify how much known interventions can shift these adverse drivers into HREI in practice.

# Notes/Todos

* ☒ Adapt health code and run
* ☒ Adapt quals code and run
* ☐ Create wages code and run
  + ☒ Look specifically at wages
  + ☐ Adapt hhincome code

## Questions at this stage

* ☐ Should wages be a separate paper (as comparatively complex)?
  + ☐ Can we really do wages as a single indicator without taking household income into account?
  + ☐ What does the wage/age curve mean for inequalities in labour market participation? (By age and by wage level)
  + ☐ What does it mean for participation in training and education?
  + ☐ Should we include an analysis which just gives a higher hourly **living** wage to everyone, rather than the arbitrary 40% threshold currently used?
* ☐ If narrowing to something more health-focused (say), should we look at discrete health states (such as LLTI) as well as continuous health (GHQ-12 MH and PH, as currently?)
* ☐ Do we want to do models/scenarios that use measures of **job insecurity** as well as wages?
  + This is especially important when considering the relationship between low wages and probability of moving to unemployment. Does earning lower wages cause low wage earners to *choose* to leave their job and become unemployed (in search of better paid jobs), *or* do jobs that pay poorly tend to also be insecure work so the probability of becoming unemployed from such jobs is naturally higher?
  + Is this a specific area where a targetted literature review would help? (Quant/Qual/Reviews etc?)

# Key Findings/Contributions

# Introduction

For working-age people, being employed or retired is associated with relatively good health, while being unemployment and long-term sickness) are associated with relatively poor health. Other economic states occupy an intermediate position. There is evidence to say that good quality work (e.g. insecure, high job-strain) is beneficial for health. There is also evidence that unemployment itself increases the risk of poor health (especially poor mental health) and premature mortality, rather than bad health being part of a broader selection effect.

From a policy perspective, there has been long-standing interest in increasing the share of the population in paid employment, and lowering the porportion who are unemployed, long-term sick and looking after home and family. Measures to achieve this have usually focused on moving people out of the detrimental states, rather than preventing them moving into them in the first place. They have tended to focus on the sub-groups of the population in these states and claiming out-of-work benefits.  The policy mix has included public employment support, increased conditionality (including sanctions and changes to benefit eligibility) as well as increasing the financial support for working. In addition, some policies have provided financial help to meet the costs of working (e.g. childcare, Access to Work). While the overall approach has moved some people into work, it also appears to have pushed others into long-term sickness and other forms of economic inactivity.

The ability to account for the differences between reservation wages (lowest wage one will accept for working) and the prevailing market wage effects on labour market supply were studied by Brown, Roberts, and Taylor (2010) who found “*no evidence for the argument that those with health problems will have higher reservation wages*”. This is an important result as Brown, Roberts, and Taylor (2010) describe how poor health therefore effectively weakens labour market participation in contrast with some studies which found that poor health increased the reservation wage. However, it should be noted that sub group analysis were not performed in the aforementioned paper which used the 1991 – 2004 British Household Panel Survey (BHPS).

A further area of interest is the effect of gender on labour market participation.

Verdugo and Allègre (2020) studied 10 countries (including the UK) and how the Great recession of 2007 impacted participation in the labour markets with countries such as UK showing that female participation increased

# Methods

## The data

The data used to fit the models are all observations from wave a to j of the UKHLS, excluding observations where data was missing on any of the predictor and outcome variables.

## The model

The model uses multinomial logistic regression to predict the economic (in)activity state in the next time period (approximately one year) based on the economic activity state in the current time period, the individual’s age, sex, and those specific drivers of interest.

## Foundational Model

The foundational model specification aims to adequately control for the effects that age, current state and sex have on transition probabilities between states. To recap, we know the following:

* That state at time T influences state at time T+1, including that there is path dependence.
* That transition propensities between states vary systematically by sex (in particular regarding the long-term carer state)
* That transitions between states vary by age, but in different ways for different states, and in ways that aren’t linear with age.

The foundational model specification operationalises the above knowledge as follows:

i.e. that next state is predicted on the current state , sex (the term so female is the reference category), the interaction of current state and sex , and a flexible function of age .

The model is implemented using the multinom function of the nnet package as follows

nnet::multinom(  
 next\_status ~ this\_status \* sex + splines::bs(age, 5),  
 ...  
 )

## Exposure Models

Exposure models extend the foundation with one or more additional variables. These variables are the exposures of interest, and for which we want to estimate the influence on economic activity levels and flows.

For a single exposure , the equation simply extends the foundational model specification as follows:

Which is specified in R as follows

nnet::multinom(  
 next\_status ~ this\_status \* sex + splines::bs(age, 5) + Z,  
 ...  
)

For two exposures, and , this simply becomes

nnet::multinom(  
 next\_status ~ this\_status \* sex + splines::bs(age, 5) + Z1 + Z2,  
 ...  
)

and so on.

In some cases (as with estimating the effects of health as an exposure) interaction terms are included between exposure variables as well. The decision about whether to include such interactions is made based on both our understanding of the extent to which factors are likely to interact in practice, and the penalised model fit as assessed using metrics like AIC and BIC.

## The simulation

We simulate three different population groups:

* Populations of various representative working ages, male and female, whom we all assume start off as in employment
* Populations of various representative working ages, male and female, who all start off as unemployed
* A representative population of varying ages, sexes, current statuses, and driver states

# Results

## Descriptive Results

### Simplified graphical transition between three states

* ☒ Show going between employed, unemployed and inactive as three node graph

[Figure 1](#fig-threestateflow) shows the transition probabilities observed throughout the entire vcalid UKHLS dataset between the three states of unemployed, employed, and economically inactive between any two consecutive waves. (This means that the same individual may contribute multiple times, given they may be observed between multiple pairs of waves). With three mutually exclusive states, there are nine possible transitions. Because economic inactivity is so heterogeneous, the models used in practice make use of seven states, leading to 49 transition probabilities.

A node connecting to itself indicates the probability of staying in that state from one wave to the next. In graph theory the nodes are called vertices and the links connecting them are called edges. We can see from the edges connecting a vertex to itself that employment is, overall, the state most likely to ‘flow to’ itself, with a probability of 95.6% of someone employed at one wave being employed at the next wave. The next most ‘sticky’ of the three vertices is economically inactive, with an 80.6% probability of remaining inactive from one wave to the next. The least sticky of the three vertices is Unemployed, where there a 43.4% probability of remaining unemployed between two consecutive waves.

There is a 25.7% probability of transitioning from unemployed to inactive, whereas there is only a 2.7% probability of transitioning from employed to inactive.

Employed Inactive Unemployed  
Employed 156120 4425 2767  
Inactive 3793 25736 2399  
Unemployed 3152 2627 4433

|  |
| --- |
| Figure 1: Transition probabilities between Employment, Unemployment and Inactivity observed between all UKHLS waves |

### Observed transitions between states

Within each wave, people are observed in each of the two economically active states, and each of the five economically inactive states. As the UKHLS are longitudinal, they can be used to calculate the proportion of those observed in each state one wave who then either stay in that state the following wave, or migrate to any of the other six states. [Table 1](#tbl-transitions_overall) shows these proportions as a single table. The rows in the first column indicate the state someone was observed for wave T, and the each of the states on the columns to the right indicate a possible state they could be observed the next wave (wave T+1). The order of the states is the same across rows and columns, meaning that the cells along the top-left to bottom-right diagonal indicate the proportions of those observed to stay in the same state from one wave to the next.

Table 1: Observed transition probabilities between economic (in)activity states between years. Rows indicate state transitioning from

|  | Active | | Inactive | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Employed | Unemployed | Inactive student | Inactive care | Inactive long term sick | Inactive retired | Inactive other |
| **Active** | | | | | | | |
| Employed | 0.956 | 0.017 | 0.003 | 0.009 | 0.004 | 0.009 | 0.002 |
| Unemployed | 0.309 | 0.434 | 0.009 | 0.115 | 0.095 | 0.022 | 0.016 |
| **Inactive** | | | | | | | |
| Inactive student | 0.368 | 0.089 | 0.461 | 0.048 | 0.015 | 0.003 | 0.015 |
| Inactive care | 0.136 | 0.081 | 0.007 | 0.705 | 0.030 | 0.024 | 0.017 |
| Inactive long term sick | 0.044 | 0.087 | 0.002 | 0.043 | 0.774 | 0.043 | 0.006 |
| Inactive retired | 0.071 | 0.016 | 0.001 | 0.034 | 0.042 | 0.827 | 0.008 |
| Inactive other | 0.308 | 0.129 | 0.020 | 0.218 | 0.052 | 0.057 | 0.217 |

Within [Table 1](#tbl-transitions_overall) the diagonal cell values show that some economic (in)activity states are more persistent than others. The high level of heterogeneity between economically inactive states illustrates why it is so important not to collapse these states into a single category.

Some states are more persistent than others. Individuals who are employment, retired, long-term sick or looking after home and family in the current wave have a high probability of still being in that state in the next wave For example, the overall probability of someone who is employed one wave remaining employed the next wave is over 95%, the proportion remaining retired is almost 83%.

Conversely, other states are much less ‘sticky’. Individuals who are unemployed, in full-time education or economically inactive for other reasons have much lower probabilities of remaining in that state.  The probability of someone unemployed remaining unemployed between waves is 43%, which is still higher than the probability of moving to employment (31%). From unemployment, there is also around a one-in-ten chance of moving either to inactive care, or to long-term sickness, but less than a 1% probability of becoming a full-time student in the next wave.For those states other than employment, the conditional probability of moving into employment is worth comparing. We can see that the conditional probability of moving from full time study (third row) to employment (first column) is 37%, which is higher than the 31% conditional probability of moving from unemployment (second row) to employment (first column). In this sense, the state of being a full-time student is *closer* to employment than the state of being unemployed, even though unemployment is considered economic activity whereas full time study is considered economic inactivity.

The transition rates observed vary markedly by sex, as shown in [Table 2](#tbl-transition-bysex).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| (a) Females   |  | Active | | Inactive | | | | | | --- | --- | --- | --- | --- | --- | --- | --- | |  | Employed | Unemployed | Inactive student | Inactive care | Inactive long term sick | Inactive retired | Inactive other | | **Active** | | | | | | | | | Employed | 0.949 | 0.015 | 0.003 | 0.016 | 0.005 | 0.010 | 0.003 | | Unemployed | 0.277 | 0.373 | 0.012 | 0.200 | 0.099 | 0.023 | 0.016 | | **Inactive** | | | | | | | | | Inactive student | 0.362 | 0.081 | 0.457 | 0.067 | 0.015 | 0.004 | 0.015 | | Inactive care | 0.134 | 0.077 | 0.007 | 0.714 | 0.030 | 0.024 | 0.014 | | Inactive long term sick | 0.045 | 0.076 | 0.002 | 0.066 | 0.763 | 0.042 | 0.005 | | Inactive retired | 0.063 | 0.014 | 0.002 | 0.057 | 0.038 | 0.818 | 0.007 | | Inactive other | 0.306 | 0.106 | 0.024 | 0.254 | 0.042 | 0.050 | 0.217 | |  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| (b) Males   |  | Active | | Inactive | | | | | | --- | --- | --- | --- | --- | --- | --- | --- | |  | Employed | Unemployed | Inactive student | Inactive care | Inactive long term sick | Inactive retired | Inactive other | | **Active** | | | | | | | | | Employed | 0.964 | 0.019 | 0.002 | 0.001 | 0.003 | 0.008 | 0.002 | | Unemployed | 0.340 | 0.494 | 0.007 | 0.031 | 0.092 | 0.022 | 0.015 | | **Inactive** | | | | | | | | | Inactive student | 0.382 | 0.107 | 0.469 | 0.011 | 0.016 | 0.002 | 0.014 | | Inactive care | 0.156 | 0.155 | 0.005 | 0.553 | 0.030 | 0.028 | 0.073 | | Inactive long term sick | 0.043 | 0.104 | 0.002 | 0.010 | 0.790 | 0.045 | 0.007 | | Inactive retired | 0.082 | 0.019 | 0.000 | 0.004 | 0.047 | 0.839 | 0.009 | | Inactive other | 0.312 | 0.166 | 0.013 | 0.159 | 0.067 | 0.067 | 0.215 | |  |

Table 2: Transition probabilities by sex

The main differences by sex shown in [Table 2](#tbl-transition-bysex) relates to the full-time care state. Compared with men, women are more likely to remain looking after home and family and to move into looking after home and family from unemployment. They are also more likely to move into inactive care from other forms of economic inactivity. They are less likely to move into work from unemployment or remain unemployed. Rates transition from full-time care to either long-term sickness or employment are similar by sex.

There are also marked differences in transition probabilities by age group, as illustrated in [Table 3](#tbl-transition-byage), which compares transition probabilities between states for persons aged between 25 and 45 years of age inclusive ([Table 3 (a)](#tbl-transition-byage-1)), with those of working age aged over 45 years of age ([Table 3 (b)](#tbl-transition-byage-2))

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| (a) Younger (25-45 years of age)   |  | Active | | Inactive | | | | | | --- | --- | --- | --- | --- | --- | --- | --- | |  | Employed | Unemployed | Inactive student | Inactive care | Inactive long term sick | Inactive retired | Inactive other | | **Active** | | | | | | | | | Employed | 0.962 | 0.017 | 0.004 | 0.012 | 0.003 | 0.000 | 0.002 | | Unemployed | 0.328 | 0.433 | 0.014 | 0.133 | 0.076 | 0.000 | 0.016 | | **Inactive** | | | | | | | | | Inactive student | 0.364 | 0.088 | 0.472 | 0.050 | 0.014 | 0.001 | 0.012 | | Inactive care | 0.153 | 0.081 | 0.009 | 0.723 | 0.021 | 0.000 | 0.013 | | Inactive long term sick | 0.066 | 0.117 | 0.005 | 0.060 | 0.740 | 0.004 | 0.008 | | Inactive retired | 0.118 | 0.029 | NA | 0.029 | 0.382 | 0.441 | NA | | Inactive other | 0.375 | 0.175 | 0.035 | 0.208 | 0.041 | NA | 0.167 | |  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| (b) Older (56 years of age and above)   |  | Active | | Inactive | | | | | | --- | --- | --- | --- | --- | --- | --- | --- | |  | Employed | Unemployed | Inactive student | Inactive care | Inactive long term sick | Inactive retired | Inactive other | | **Active** | | | | | | | | | Employed | 0.948 | 0.017 | 0.001 | 0.006 | 0.005 | 0.020 | 0.003 | | Unemployed | 0.283 | 0.435 | 0.003 | 0.090 | 0.121 | 0.052 | 0.015 | | **Inactive** | | | | | | | | | Inactive student | 0.404 | 0.101 | 0.374 | 0.035 | 0.025 | 0.025 | 0.035 | | Inactive care | 0.099 | 0.081 | 0.002 | 0.668 | 0.050 | 0.074 | 0.027 | | Inactive long term sick | 0.034 | 0.073 | 0.001 | 0.035 | 0.790 | 0.062 | 0.005 | | Inactive retired | 0.071 | 0.016 | 0.001 | 0.034 | 0.040 | 0.829 | 0.008 | | Inactive other | 0.256 | 0.093 | 0.009 | 0.226 | 0.060 | 0.100 | 0.255 | |  |

Table 3: Transition probabilities by broad age group

As might be expected, the rates of transition into retirement are considerably higher in older ages ([Table 3 (b)](#tbl-transition-byage-2)) than younger ages ([Table 3 (a)](#tbl-transition-byage-1)), and the probabilities of someone remaining retired almost twice as high in older than younger ages. Rates of employment are somewhat higher in the younger ages than higher ages, and so are the probabilities of moving from unemployment to employment. Rates of transition into full-time care are somewhat higher in the younger age category.

The two age categories presented above are somewhat arbitrary, and do not capture adequately how the probability of being in each of the states, and moving to other states, varies over the working age life course. As an example of this, figure X shows how the probability of remaining employed, unemployed, a full-time carer, a student, or long-term sick varies over five year intervals.

|  |
| --- |
| Figure 2: Probability of staying in a category by sex and age group |

These stayer probabilities are somewhat artificial for some ages - for example of remaining retired at ages where retirement is unlikely - but do show how the probabilities of remaining in each state vary over the life course, as well as differ by sex. The probabilities of transition between each of these states also changes over sex and age, and so a foundational model which controls for these varying associations in these standard (unmodifiable) demographic variables is important before reasonable estimates of the additional (potentially modifiable) exposures can be produced. The purpose of the foundational model specification is to do this.

## Simulation Model Results

For all models presented here we focus on populations aged 25-64, so that results are comparable. This is to adjust for the differences in participation in education by age group, since younger people may still be gaining formal qualifications.

### Modelling health effects, people aged 25-64

The effect of suboptimal health as an exposure was assessed using SF-12 scores, subdivided into the physical health (PH) and mental health (MH) subdomains, and then standardised over the observed population to have a mean of 0 and standard deviation of 1.

Four different exposure model specifications were considered:

* mod\_mh: MH only
* mod\_ph: PH only
* mod\_ph\_mh: MH and PH as independent effects
* mod\_phmh: MH and PH including an interaction term

Each of these was compared for penalised model fit against the foundational model specification using AIC and BIC, with lower scores preferred. (1 = most preferred; 5 = least preferred)

Table 4: AIC and BIC for different model specifications for including health as an exposure

| model | df | AIC | BIC | aic\_rank | bic\_rank |
| --- | --- | --- | --- | --- | --- |
| mod\_00 | 126 | 162676.1 | 163965.4 | 5 | 5 |
| mod\_ph | 132 | 161731.3 | 163082.0 | 3 | 3 |
| mod\_mh | 132 | 161525.6 | 162876.4 | 2 | 2 |
| mod\_ph\_mh | 138 | 159786.0 | 161198.2 | 1 | 1 |
| mod\_phmh | 144 | 161874.3 | 163347.9 | 4 | 4 |

[Table 4](#tbl-model_fit_ghq12) shows that the specification with independent effects of MH and PH is preferred by both AIC and BIC. Of models including only MH or PH, the model specification for MH is preferred.

Based on this, we will consider the following scenarios with the following models:

* Scenario 1: Mental health only is improved using model mod\_mh
* Scenario 2: Mental health only is improved using model mod\_ph\_mh
* Scenario 3: Physical health only is improved using model mod\_ph\_mh
* Scenario 4: Mental health and physical health are both improved using mod\_ph\_mh

#### Scenarios 1 and 2: Improving mental health only

[Table 5](#tbl-mh_scenarios) shows the predicted effects on the number of people in each economic category of increasing each individual’s mental health by a substantial amount, one standard deviation. The baseline model shows the distribution of the population aged 25-64 before health was improved; the counterfactual shows it after health was improved. The results are based on UKHLS participants for whom relevant information was observed in wave j, the last pre-COVID wave in the dataset. [Table 5 (a)](#tbl-mh_scenarios-1) is based on the model which includes MH as a driver only, whereas [Table 5 (b)](#tbl-mh_scenarios-2) is based on the model in which PH is also included as a driver, but in this scenario is not modified.

(a) Based on MH only model

| State | base | counterfactual | Absolute Change | Relative Change |
| --- | --- | --- | --- | --- |
| Employed | 12497 | 12656 | 159 | 1.3% up |
| Unemployed | 638 | 559 | -79 | 12.4% down |
| Inactive student | 94 | 89 | -5 | 5.3% down |
| Inactive care | 827 | 846 | 19 | 2.3% up |
| Inactive long term sick | 670 | 567 | -103 | 15.4% down |
| Inactive retired | 552 | 568 | 16 | 2.9% up |
| Inactive other | 80 | 74 | -6 | 7.5% down |

(b) Based on MH + PH model

| State | base | counterfactual | Absolute Change | Relative Change |
| --- | --- | --- | --- | --- |
| Employed | 12530 | 12677 | 147 | 1.2% up |
| Unemployed | 619 | 560 | -59 | 9.5% down |
| Inactive student | 88 | 91 | 3 | 3.4% up |
| Inactive care | 857 | 867 | 10 | 1.2% up |
| Inactive long term sick | 681 | 572 | -109 | 16% down |
| Inactive retired | 520 | 530 | 10 | 1.9% up |
| Inactive other | 66 | 61 | -5 | 7.6% down |

Table 5: Estimates of substantially improving Mental Health

The two models whose results are shown in [Table 5](#tbl-mh_scenarios) produce similar but not identical estimates on the effect of improving mental health on the number of people in each state in the following wave. Compared with the baseline scenario in which no one’s MH is changed, in the counterfactual scenario in which MH is substantially there is around a 1.2 to 1.3% increase in the size of the employed population, and the relative size of the unemployed population is reduced by around a tenth. The most substantive relative change in state population size is in the economically inactive, long-term sick state, whose population size is predicted to fall by at least 15.4%, based on the first model, and by 16.0% based on the second model. In absolute terms, the biggest changes would be seen in the employed population (large increase), unemployed (large reduction), and long-term sick (large reduction)

#### Scenario 3: Improving physical health only

[Table 6](#tbl-ph_scenario) shows the estimated impact on state sizes of substantially improving PH. As with the MH intervention scenarios, in the counterfactual scenario PH scores were improved by one standardised unit for all valid wave J observations. The model mod\_ph\_mh was used for this scenario calculation.

Table 6: Estimates of substantially improving Physical Health

| State | base | counterfactual | Absolute Change | Relative Change |
| --- | --- | --- | --- | --- |
| Employed | 12530 | 12812 | 282 | 2.3% up |
| Unemployed | 619 | 541 | -78 | 12.6% down |
| Inactive student | 88 | 92 | 4 | 4.5% up |
| Inactive care | 857 | 831 | -26 | 3% down |
| Inactive long term sick | 681 | 494 | -187 | 27.5% down |
| Inactive retired | 520 | 521 | 1 | 0.2% up |
| Inactive other | 66 | 70 | 4 | 6.1% up |

The scenario indicates that a substantial improvement on physical health could have an even larger effect on the size of the long-term sick population than a similarly sized improvement in mental health. In this scenario, the relative size of the long-term sick population falls by more than a quarter. The relative size of the unemployed population also is estimated to change substantially, falling by 12.6%.

#### Scenario 4: Improving mental and physical health

There is no single health driver/exposure included in the model. Instead there are separate mental health and physical health exposures. However both mental health and physical health have been standardised, meaning we can model a scenario in which ‘health’ has been improved, and the effect of these health improvements is equal across the mental and and physical health subdomains. In order to ensure we are looking at the effect of the type of the driver being modified, rather than the amount of change we are making to these drivers, we need to employ a little trigonometry. If we were to modify both MH and PH by one standard unit, the total amount of change in ‘health’ would be the hypotenuse of a triangle in which both MH and PH are ‘legs’, i.e.  or , which is 1.41 to two decimal places, and so larger than either of the previous exposure reductions being modelled. In order to work out the amount of equal change across both ‘legs’ required for a 1 unit total change across both dimensions, we therefore need to solve , i.e. . This means , so , and therefore . In scenario 4, therefore, both MH and PH are increased by this same amount, which is 0.71 to two decimal places. The results of running this scenario are shown in [Table 7](#tbl-genhealth)

Table 7: Estimates of substantially improving health via equal improvements in mental health and physical health

| State | base | counterfactual | Absolute Change | Relative Change |
| --- | --- | --- | --- | --- |
| Employed | 12530 | 12827 | 297 | 2.4% up |
| Unemployed | 619 | 524 | -95 | 15.3% down |
| Inactive student | 88 | 93 | 5 | 5.7% up |
| Inactive care | 857 | 844 | -13 | 1.5% down |
| Inactive long term sick | 681 | 478 | -203 | 29.8% down |
| Inactive retired | 520 | 528 | 8 | 1.5% up |
| Inactive other | 66 | 65 | -1 | 1.5% down |

[Table 7](#tbl-genhealth) suggests that the effects of intervening both on MH and PH, without changing the total amount of change, may lead to greater shifts than intervening on online one of the two subdimensions. In this scenario, the relative size of the long-term sick population is predicted to fall by almost 30%, and relative size of unemployment to fall by over 15%. The estimated effect on the size of those in the employed state is also estimated to be slightly greater than for either MH or PH only interventions, at 2.4%.

### Modelling effects of qualifications

To model the effect of low or no qualifications on economic inactivity, we use the hiqual\_dv variable from the UKHLS, which standardises a large number of potential categories into a much smaller number of groupings. We then further simplify the qualification categories into:

* No qualifications
* Some qualifications
* Degree or higher

[Table 8](#tbl-quals_j) shows the proportion of wave J UKHLS sample with each level of qualification. In this sample, slightly over half of respondents have degree level qualifications, around 45% have intermediate qualifications, and slightly over 3% have no qualifications listed.

Table 8: Proportion in wave J data with no, some, or degree qualifications

| qual\_group | n | share |
| --- | --- | --- |
| None | 518 | 0.0337240 |
| Some | 6971 | 0.4538411 |
| Degree | 7871 | 0.5124349 |

The counterfactual scenario involves making the following adjustment to the data:

If an individual has no qualifications, assign them ‘some’ qualifications; otherwise leave the qualification listed unchanged.

Because of the relatively small proportion of the sample population this is likely to affect, the effect on the total pool size may be relatively small. However the effect on those who are affected could be substantial.

Both AIC and BIC indicate a preference for the model with qualifications included than without.

#### Illustrative vignettes on the effect of qualifications on remaining in different states

We can explore what effect different qualification levels are predicted to have on transitions by looking at the following vignettes:

* 30 years old and 50 years old
* male and female
* no, some, degree qualifications
* Initially either employed, unemployed, full time carer or long-term sick

[Table 9](#tbl-staybyquals) shows the probability of people who start off either employed, unemployed, long-term sick or a full time carer remaining in that state the next wave by qualification category

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| (a) Employed   | age | sex | None | Some | Degree | | --- | --- | --- | --- | --- | | 30 | male | 0.9409277 | 0.9713657 | 0.9812424 | | 30 | female | 0.9086928 | 0.9509994 | 0.9665159 | | 50 | male | 0.9342390 | 0.9671050 | 0.9762653 | | 50 | female | 0.9085264 | 0.9515921 | 0.9650726 | |  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| (b) Unemployed   | age | sex | None | Some | Degree | | --- | --- | --- | --- | --- | | 30 | male | 0.6820143 | 0.5357910 | 0.4336427 | | 30 | female | 0.4979020 | 0.3956433 | 0.3298599 | | 50 | male | 0.6186675 | 0.4911096 | 0.4007770 | | 50 | female | 0.4730265 | 0.3796713 | 0.3166049 | |  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| (c) Long term sick   | age | sex | None | Some | Degree | | --- | --- | --- | --- | --- | | 30 | male | 0.7529974 | 0.7073101 | 0.6097846 | | 30 | female | 0.7338511 | 0.6842030 | 0.5897351 | | 50 | male | 0.8605863 | 0.8247745 | 0.7386771 | | 50 | female | 0.8581425 | 0.8190813 | 0.7310506 | |  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| (d) Full time carer   | age | sex | None | Some | Degree | | --- | --- | --- | --- | --- | | 30 | male | 0.6895095 | 0.6507908 | 0.5914976 | | 30 | female | 0.8142791 | 0.7723590 | 0.7201150 | | 50 | male | 0.5760518 | 0.5340457 | 0.4651449 | | 50 | female | 0.7219586 | 0.6741193 | 0.6072774 | |  |

Table 9: Probability of staying in category by qualification level

We can see from [Table 9](#tbl-staybyquals) that higher qualification categories are associated with increased probabilities of remaining employed, and reduced probabilities of remaining unemployed or remaining long-term sick, but also reduced probabilities of remaining a full time carer. But we can also see that the extent to which these absolute changes occur is highly dependent on other demographic characteristics.

#### Simulation with wave J data

[Table 10](#tbl-nosomequal_counter) shows the estimated effect of a counterfactual where those with no qualifications are instead assigned ‘some’ qualifications. As mentioned only a small proportion of the sample frame have no qualifications so the majority of the observations are unchanged between the baseline and counterfactual scenarios.

Table 10: Estimates of effect of ‘giving’ those with no qualifications some qualifications

| State | base | counterfactual | Absolute Change | Relative Change |
| --- | --- | --- | --- | --- |
| Employed | 12582 | 12610 | 28 | 0.2% up |
| Unemployed | 594 | 578 | -16 | 2.7% down |
| Inactive student | 90 | 91 | 1 | 1.1% up |
| Inactive care | 826 | 819 | -7 | 0.8% down |
| Inactive long term sick | 640 | 628 | -12 | 1.9% down |
| Inactive retired | 550 | 556 | 6 | 1.1% up |
| Inactive other | 78 | 78 | 0 | 0% down |

Even given the small number of observations changed, [Table 10](#tbl-nosomequal_counter) projects an almost 3% fall in the size of the unemployed pool, and a 2% fall in the size of the inactive sick pool.

A more extreme counterfactual to consider is a scenario in which everyone has the same economic engagement tendencies of those with degrees. This is shown in [Table 11](#tbl-alldegree_counter).

Table 11: Estimates of effect of ‘giving’ everyone a degree

| State | base | counterfactual | Absolute Change | Relative Change |
| --- | --- | --- | --- | --- |
| Employed | 12582 | 12764 | 182 | 1.4% up |
| Unemployed | 594 | 494 | -100 | 16.8% down |
| Inactive student | 90 | 100 | 10 | 11.1% up |
| Inactive care | 826 | 750 | -76 | 9.2% down |
| Inactive long term sick | 640 | 528 | -112 | 17.5% down |
| Inactive retired | 550 | 636 | 86 | 15.6% up |
| Inactive other | 78 | 89 | 11 | 14.1% up |

[Table 11](#tbl-alldegree_counter) shows that, in this more extreme counterfactual, a larger proportionate reduction in the size of the Unemployed pool is expected, along with the pool of those in the economically inactive long-term sick pool. However, increases in the proportion who become inactive due to being a full time student, and due to retirement, is also projected to result.

### Modelling effects of higher or lower wages

The previous drivers considered are those which apply to persons in any of the economic (in)activity categories modelled. By contrast, it is only reasonable to consider wages as a driver for those in receipt of wages, i.e. those in employment.

**NOTE: THERE ARE TOO MANY UNCERTAINTIES/RESEARCHER DEGREES OF FREEDOM REGARDING WAGE EFFECT MODELLING TO BE SENSIBLE TO PROCEED, WE THINK, IN THIS FIRST PAPER. INSTEAD, A NOTEBOOK WITH WAGE ANALYSIS WILL BE SHARED ALONGSIDE SOME QUESTION ABOUT HOW BEST TO PROCEED IN THIS AREA**

#### Some interrim findings

* Lower wages *are* associated with increased probability of leaving employment
* These differences do not appear to substantively differ by gender
* But they do appear to differ by age, with younger adults more sensitive to differences in relative wages than older adults.

#### Descriptive

* Want to get some kind of event history summary statistics
  + e.g. number of people in panel who have been employed 1 wave, 2 waves, 3 waves etc

# Discussion

## Strengths

* Very generalisable and adaptable framework for potentially considering many drivers both individually and in combination.
* Makes use of a large and complex dataset which was initially broadly representative of the UK population.
* UKHLS has both strengths in following individuals and households over time. But on the other hand there is selective retention from wave to wave, so the study sample may become less representative of the general population over time.
  + Particular strength is the longitudinal nature of the study. This was essential for allowing transitions at individual level to be modelled, as a first stage to producing estimates of the effects of drivers on the overall populations

## Comparison with other findings

* We haven’t looked yet. **A priority to focus on?**

## Limitations/Caveats

* We aren’t capturing uncertainty (i) where there are few observations (ii) representativeness of the sample (including sub-groups) (iii) multiplication of uncertainty - more variables, more biased data (iv) computations stop at 100 runs…we could do more for a more precise model.

## Implications

### Research

* As more variables are included in a model, there is less complete data available to populate the models. This may bias the results. We could look at methods such as multiple imputation to attempt to reduce this source of bias.
* We can attempt to make the results more representative of specific within-UK populations, such as those more representative of Scotland, Northern Ireland, or Wales. Different UK nations differ in terms of their demographic and socioeconomic characteristics and so the same exposures are likely to have different effects in different parts of the UK.
* We can also use the modelling framework to consider the effect that ageing populations alone will likely have on the proportion of the working age population who are economically active and inactive in different ways.
* We can also use the framework to project the effect that modest but sustained interventions on drivers may have over multiple years.
* Will link geographic drivers to provide broader upstream context to movements into and out of different economic (in)activity states.

### Practice

* These estimates attempt to quantify the amount of HREI ‘explained by’ specific drivers, based on a PAF-style framing of the questions. This is equivalent to asking what the effect of fully mitigating any driver would be on the population as a whole. However the effectiveness of any real-world intervention is likely to be less, as no adverse driver/exposure is likely to be fully mitigated. As such these estimates likely give an upper bound to the effect of any real-world intervention.

# References

Brown, Sarah, Jennifer Roberts, and Karl Taylor. 2010. “Reservation Wages, Labour Market Participation and Health.” *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 173 (3): 501–29. <http://www.jstor.org/stable/40666273>.

Verdugo, Gregory, and Guillaume Allègre. 2020. “Labour Force Participation and Job Polarization: Evidence from Europe during the Great Recession.” IZA Discussion Papers 13425. Institute of Labor Economics (IZA). <https://ideas.repec.org/p/iza/izadps/dp13425.html>.

# Appendices

## Appendix 1

## Appendix 2