Using UKMOD to estimate the changes in income, taxes, and benefits from entering work: Modelling protocol

The purpose of this document is to outline in detail the steps to take in using UKMOD to model changes in taxes, benefits, and incomes of individuals with health conditions when transitioning into or out of employment.

**Step 1: Obtain UKMOD software and data**

To obtain the open source model download the EUROMOD software, download the UKMOD model, and apply for access to the Family Resources Survey (FRS) dataset needed for the simulation [from the UKMOD website](https://www.microsimulation.ac.uk/ukmod/access/).

To apply for access to the data, you will need to prove you have access to the [FRS data](https://beta.ukdataservice.ac.uk/datacatalogue/series/series?id=200017) via the UK data service.

Once access has been granted to the FRS dataset(s) required, place the text files in the *Inputs* folder in UKMOD.

**Step 2: Use UKMOD to simulate income, tax payments, and benefit receipt**

Steps:

* Open EUROMOD and select the UKMOD model project folder.
* On the Countries tab in the top left, select UK.
* The columns which appear on screen (*UK\_2016*, *UK\_2017* etc.) represent the tax/benefit policy for each year.
  + These can be edited to create alternative policy scenarios to compare to the baseline (*we don’t need to do this in this case, we are using the baseline*).
* In the top left corner, click *Run UKMOD*.
* Check the box for each policy scenario(s) you want to run. Next to each checkbox, use the dropdown to select the input dataset.
  + The selected dataset must be present in the *Inputs* folder. UKMOD will automatically apply uprating factors to adjust the input data to the price year of the scenario selected.
* The result will be a dataset in the Outputs folder of the model, **uk\_XXXX\_std.txt** with the individual level results of the simulation.

**Step 3: Filtering the data to identify subgroups**

The data need to be filtered to identify subgroups based on health condition and employment status. To do so requires obtaining the raw FRS data from the UK Data Service and identifying the subgroups of interest - The UKMOD data do not contain variables for health condition and time spent in work.

The individual identifiers for the relevant observations can be obtained from the full raw FRS data, merged to the UKMOD output dataset, and used to filter the output dataset to the observations of interest.

**Health Conditions:**

Health conditions are identified from a set of binary variables in the **adult** dataset of the FRS:

* *DISD01*: Difficulty with vision
* *DISD02*: Difficulty with hearing
* *DISD03*: Difficulty with mobility
* *DISD04*: Difficulty with dexterity
* *DISD05*: Difficulty with learning
* *DISD06*: Difficulty with memory
* *DISD07*: Difficulty with mental health
* *DISD08*: Difficulty with stamina or breathing
* *DISD09*: Difficulty with social or behavioural issues
* *DISD010*: Other

**Employment status**

Employment status - *in-work (IW)* or *out-of-work (OOW)* can be identified in the UKMOD output data using the variable *les*, which is coded as follows:

1. Employer or self-employed (IN WORK) **(OUT OF WORK)**
2. Employee (IN WORK) **(OUT OF WORK)**
3. Pensioner
4. Unemployed **(OUT OF WORK)**
5. Student
6. Inactive **(OUT OF WORK)**
7. Sick or disabled **(OUT OF WORK)**
8. Family worker

This variable can be used to compare the income/tax/benefits profiles of individuals who are OOW with those who are IW to estimate the impact of moving from OOW to IW or vice-versa.

The IW group can be further refined to be more representative of individuals who have recently transitioned into IW from OOW:

* restrict to those who have been in their current job for less than a year
* restrict to those who, prior to their current job, were not in work.

These variables can be constructed from variables available in the **job** data:

* *WRKPREV*: What doing immediately before the current job
  1. In paid employment
  2. Self-employed
  3. Not in paid employment
* *WORKMTH*: Month started present job
* *WORKYR*: Year started present job

**Step 4: Match the FRS to UKMOD output data**

The identifiers in the UKMOD data are generated from the Family Resources Survey. To match from the data downloaded from UKDS to the UKMOD the identifiers are:

* **person**: original FRS individual identifier. Renamed **idorigperson** in UKMOD data.
* **sernum**: original FRS household identifier. Renamed **idorighh** in UKMOD data.
* **benunit**: original FRS benefit unit identifier. Renamed **idorigbenunit** in UKMOD data.

**Step 5: Calculate mean outcomes for each subgroup from the UKMOD output data**

With UKMOD output data matched to the additional subgroup identifier variables from the raw FRS, calculate subgroup means of the key outcome variables from the UKMOD simulation.

* **ils\_ben** - total benefit income. It is the sum of **ils\_pen** (pensions), **ils\_benmt** (means-tested benefits), and **ils\_ben\_nnt** (non-means tested benefits).
* **ils\_tax** - direct taxtion (income tax + plus council tax).
* **ils\_sicdy** - employee/self-employed national insurance contributions.
* **ils\_sicer** - employer national insurance contributions.
* **ils\_earns** - labour market earnings.
* **ils\_dispy** - disposable income.

Weighted averages of these variables can be calculated using the *dwt* weight variable - the cross-sectional weights from the FRS data which grosses aggregate values to the population level.