Comparison between model predicted and observed distributions, wave j

Jon Minton

Martin Taulbut

# Aim

The aim of this document is to see how the distribution of states, for a broadly representative sample at wave i, as projected to wave j, compares to what’s observed in wave j.

We will do this for two different model specifications:

* Foundational model: age, sex, current state
* Health model: Foundational model with health scores

We know what adding the health variables, including interaction between mh and ph, improves the penalised model fit using the AIC and BIC metrics. However we have not so far looked at how much this improvement leads to improved estimates of the distribution of persons in each state.

# Set up

Creating the models to compare:

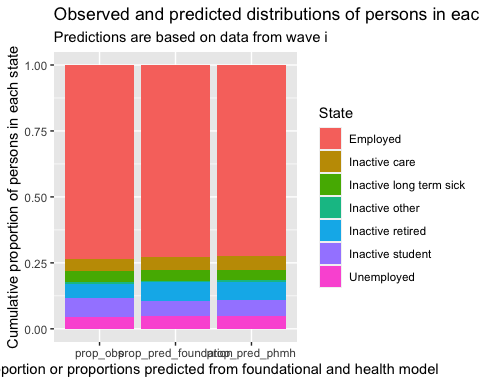
# weights: 238 (198 variable)  
initial value 524979.315476   
iter 10 value 192658.313860  
iter 20 value 176857.990038  
iter 30 value 169692.163769  
iter 40 value 151958.454105  
iter 50 value 144648.651059  
iter 60 value 138954.343056  
iter 70 value 134503.216911  
iter 80 value 131857.626953  
iter 90 value 129284.963603  
iter 100 value 126220.997935  
iter 110 value 125067.570758  
iter 120 value 124662.135636  
iter 130 value 124525.456987  
iter 140 value 124508.428226  
iter 150 value 124506.845713  
iter 160 value 124506.610549  
iter 170 value 124506.582142  
final value 124506.576940   
converged

# weights: 259 (216 variable)  
initial value 524979.315476   
iter 10 value 171010.034623  
iter 20 value 160995.384580  
iter 30 value 157371.134178  
iter 40 value 154178.315979  
iter 50 value 138207.269153  
iter 60 value 132229.522606  
iter 70 value 128515.204070  
iter 80 value 124893.595049  
iter 90 value 123381.433099  
iter 100 value 122567.261067  
iter 110 value 122124.701172  
iter 120 value 121793.846192  
iter 130 value 121640.204552  
iter 140 value 121588.507824  
iter 150 value 121566.646540  
iter 160 value 121566.233865  
iter 170 value 121566.206741  
iter 180 value 121566.187480  
final value 121566.184846   
converged

Now to get the predicted values for wave j give the data in wave i

# A tibble: 7 × 4  
 this\_status observed\_n pred\_foundation pred\_phmh  
 <chr> <int> <dbl> <dbl>  
1 Employed 14917 15767. 15728.  
2 Inactive care 959 1100. 1099.  
3 Inactive long term sick 813 856. 881.  
4 Inactive other 168 134. 135.  
5 Inactive retired 1064 1522. 1518.  
6 Inactive student 1492 1277. 1276.  
7 Unemployed 881 1038. 1056.

Graphically, the observed and predicted distributions for people in states are as follows:



We can see that overall the two models produce very similar distributions of persons in each state. The proportion who are students appears to be slightly underpredicted by the models, and the proportion who are retired slightly overpredicted. However, so long as the foundational model is used as the foundation of all other models, such biases should be consistent across all models, and so should still be usable to predict the kinds of influence that single (or multiple) additional drivers have on propensities to move between states, and resultant predicted/projected distributions.

Now add column proportions for each of the three columns

| this\_status | prop\_obs | prop\_pred\_foundation | prop\_pred\_phmh | abs\_diff\_foundation | abs\_diff\_phmh | rel\_diff\_foundation | rel\_diff\_phmh |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Employed | 0.7350448 | 0.7267895 | 0.7249915 | -0.0082554 | -0.0100533 | -0.0112311 | -0.0136772 |
| Inactive care | 0.0472553 | 0.0507049 | 0.0506523 | 0.0034496 | 0.0033970 | 0.0729989 | 0.0718859 |
| Inactive long term sick | 0.0400611 | 0.0394791 | 0.0406199 | -0.0005820 | 0.0005588 | -0.0145275 | 0.0139491 |
| Inactive other | 0.0082783 | 0.0061665 | 0.0062255 | -0.0021118 | -0.0020529 | -0.2551005 | -0.2479798 |
| Inactive retired | 0.0524293 | 0.0701348 | 0.0699929 | 0.0177055 | 0.0175636 | 0.3377022 | 0.3349968 |
| Inactive student | 0.0735193 | 0.0588819 | 0.0588300 | -0.0146373 | -0.0146893 | -0.1990950 | -0.1998016 |
| Unemployed | 0.0434118 | 0.0478432 | 0.0486878 | 0.0044314 | 0.0052760 | 0.1020781 | 0.1215333 |

In the above we have the following columns:

* prop\_obs: The observed proportion of persons in each state in wave j
* prop\_pred\_foundation: The proportion of persons in each state in wave j, as predicted by the foundational model and data from wave i
* prop\_pred\_phmh: The proportion of persons in each state in wave j, as predicted by the health model and data from wave i
* abs\_diff\_foundation: The absolute difference between the observed and predicted proportions of persons in each state, for the foundational model
* abs\_diff\_phmh: The absolute difference between the observed and predicted proportions of persons in each state, for the health model
* rel\_diff\_foundation: The relative difference between the observed and predicted proportions of persons in each state, for the foundational model
* rel\_diff\_phmh: The relative difference between the observed and predicted proportions of persons in each state, for the health model

Overall, it appears that the foundational model alone does a very good job of producing predicted distributions for wave j, which are close to those actually observed. For all but two states (retired and student), the absolute percentage point differences between what’s predicted and what’s observed is less than 1% point; for the retired and student states, the absolute percentage point differences are 1.5% and 1.7% respectively.

In terms of the additional improvement from adding health variables, these are more modest than expected, but are associated with improvements for the inactive long-term sick state, in particular.

An important caveat is that the observed wave j data are slightly different to the wave i data used by both the foundational and health model to predict wave j. This is because some people in wave i will have dropped out by wave j, and some people may be in wave j who were not in wave i. Despite this, the comparison between model predicted and observed distributions is a useful and reassuring check to ensure that the model produces aggregate results which are broadly consistent with what’s observed.