**Prior Justification**

We designed the priors to be vague. To do this, we considered parameter values that would be reasonable, and then chose the priors so that an interval with 95% probability could be constructed that included both all the values in the reasonable range as well as some outside the range.

***Multivariate latent class model***

The priors for the latent class model were as follows:



TTO utilities ranged between -1 and 1 by design. Therefore regression coefficients can reasonably be expected to fall below 1 in absolute value. The prior is even more diffuse than this, with a 95% probability that any given regression coefficient falls below 1.64 in absolute value.

Our pilot work indicated that regression coefficients for the DCE can be expected to fall between 0 and -1.5 (or 0.1 and -1.5) for the intercept; our prior is more diffuse than this, with a 95% that the regression coefficients are larger than -1.64.

The N(0,1) prior on TTO, DCE represents a 95% probability that the odds of belonging to class 2 vs 1 lies between 0.15 and 7, and similarly for class 3, allowing for large variation in the size of the latent classes.

Given the bounds on the TTO scores, it is reasonable to expect that both uTTO and e are less than 1/2; a vague prior would assign some probability to the event that they are larger than 0.5. The Gamma priors yield a 95% probability that the standard deviations are smaller than 4.4, and a 78% probability that they are larger than 0.5.

The N(0,1) distribution on  yields a 95% probability that the ratio of odds ratios of belonging to class c vs 1 for individual i vs j, for the TTO vs. the DCE data lies between 0.04 and 20.4, thus capturing no association (ROR=1) negative association (ROR<1) and positive association (ROR>1), and also including very strong and negative and positive association.

***Multivariate mixed model***



The priors for the regression coefficients  and , and for the precision parameter , are as in the latent class model, and vague in the mixed model for the same reasons as in the latent class mixed model.

The random effects can be expected to be less than 1 in absolute value, and hence their standard deviation can be expected to be less than 0.5. The Wishart prior on the precision matrix  places a 95% probability on the event that they are less than 5.4. Similarly, the Wishart prior on  places a 95% probability on the event that the correlation between random effects for the same dimension is between -0.997 and 0.997.

**Sensitivity to priors**

We assessed the sensitivity of the results to the priors by re-running the analysis with widened priors, as follows:

For the multivariate latent class model:



For the multivariate mixed model:



In the above, the Normal distributions are parameterized in terms of their mean and variance.

The revised Gamma prior for the precision  results in a prior 95% CrI for  of 0.0046 to 1.68×1077. This credible interval is wide enough that it would be hard to imagine the true value of  lying outside it.

We leave the Wishart distribution unchanged as the shape parameter (in this case 2) cannot be smaller than the dimension of the matrix, i.e. we have chosen the least informative prior possible.

The results are in Appendix Tables 1-3, and give similar results to the original priors