**Appendix B**

In this appendix we explain how to use our estimates and software to predict the EQ-5D when only the SF-12 is available.

**1) Using our estimates**: In the following example, we use estimates from the Mixture 2 model to make predictions based on classification (CEC). Table 2 shows parameters in the EQ-5D scale but in this example we use parameters in the reversed scale (1-EQ-5D) and we also use estimates with more decimal places than in Table 2. We calculate predicted values for an individual with PCS=48, MCS=30, and age= 50. Because we centered our covariates (see paper for details) these values correspond to -2, -20, and -15, respectively, in the centered scale.

First, we calculate the expected value for each Tobit component. In the Tobit model, the estimated parameters correspond to the latent variable, which is assumed to distribute normal, while our interest is on the censored predictions. To obtain censored predictions, we first calculate predictions for the latent variables:

µ1 = .2308311 -.0020312\*MCS -.0045566\*PCS

µ1 = .2308311 -.0020312\*(-20) -.0045566\*(-2)

µ1 = 0.280568.

µ2 = .6449207 -.0046078 \*MCS -.0095772 \*PCS

µ2 = .6449207 -.0046078 \*(-20) -.0095772 \*(-2)

µ2 = 0.756231.

The estimated standard deviations of component 1 and 2 are σ1 = .0547989 and σ2 = .1521847. The values µ1, µ2, σ1, σ2 enter into the calculation of the censored predictions as parameters of the cumulative standard normal distribution and the standard normal density (34). We recommend using Stata or another statistical package for this step:

gen u1c = normal(0.280568/.0547989)\*( 0.280568+.0547989\*normalden(0.280568/.0547989)/normal(0.280568/.0547989))

gen u2c = normal(0.756231/.1521847)\*( 0.756231+.1521847\*normalden(0.756231/.1521847)/normal(0.756231/.1521847))

Because in this particular example predictions corresponding to PCS=48, MCS=30 and age= 50 were not censored, u1c = µ1 = 0.280568 and u2c = µ2 = 0.756231.

Second, we calculate the probability of belonging to each class in the multinomial scale and then in the probability scale. In the multinomial scale:

m1 = 1.563787 -.1483147\*MCS -.2152055\*PCS + .0263548\*age

m1 = 1.563787 -.1483147\*(-20) -.2152055\*(-2) + .0263548\*(-15)

m1 = 4.56517.

m2 = -1.954959 -.2490964\*MCS -.3208357 \*PCS + .0212149 \*age

m2 = -1.954959 -.2490964\*(-20) -.3208357 \*(-2) + .0212149 \*(-15)

m2 = 3.350417.

In the probability scale:

p1 = exp(4.56517) / (1+exp(4.56517)+ exp(3.350417)) = 0.764999

p2 = exp(3.350417) / (1+exp(4.56517)+ exp(3.350417)) = 0.227039

p3 = 1 - 0.764999 - 0.227039 = 0.007962.

Because p1 > p2 > p3, this individual is classified as belonging to the first Tobit component. Therefore, the predicted value for this individual is u1c = 0.280568. In the EQ-5D scale, 1-0.280568 = .660196.

**2) Using MEPS data:** When possible, we recommend using our Stata command zicen along with MEPS data to estimate models and make predictions. MEPS datasets are publicly available and easy to use because of their extensive documentation[[1]](#footnote-1). For this example, we assume the variables sfpc, sfmc, and age represent the PCS, MCS, and respondent’s age, respectively. The variable eq5dr is the EQ-5D preference index in the reversed scale. The variable validation is an indicator variable equal to 1 if an observation belongs to the validation sample and 0 if the observation belongs to the estimation sample.

The following code estimates a mixture model with covariates in both the mean Tobit components and the mixture probabilities:

zicen eq5dr sfpc sfmc if validation==0, cl(3) prob(agec sfpc sfmc) difficult

To obtain WA predictions after estimating the model:

predict wapreds

By default, predictions are calculated for all the observations in the dataset, regardless of whether they were used to estimate the model. Predictions are stored in the variable wapreds.

To obtain CEC predictions:

zicenec cecpreds, predict

 Predictions are stored in the cecpreds variable.

Type help zicen in Stata for more predict and zicenec options.

**2) Combining MEPS with other datasets:** The easiest way to use our models for predicting the EQ-5D in other datasets is to append the dataset without the EQ-5D to the MEPS dataset (see help on Stata’s append command) ensuring that in both datasets the relevant variables are stored in the same format. Then create an indicator variable that identifies each data source and estimate models restricting observations to the MEPS dataset. After estimating the models, use the post-estimation commands predict and zicenec as in 2) to make predictions for all observations.

1. They can be downloaded from http://meps.ahrq.gov/mepsweb/data\_stats/download\_data\_files.jsp [↑](#footnote-ref-1)