Final Prediction rules for Data sets 1 and 2

Data set 1

Data set 1 has 36 items and 9 subscales. Sample size is 240. The rule with the lowest estimated prediction error MSE_{pr} is PCovR item rules.

Hyperparameters of PCovR item rules

There are two hyperparameters in PCovR that we need to set: 1) the number of components, and 2) a weighting parameter which can run from .1, to .99. Weights above .5 means that we emphasize the explanation of the total variance in the items.

As you can see in the figures below, the frequently chosen solution is 6 components and a weighting parameter of .99. These two values I will use to generate the final prediction rule for the CERQ.

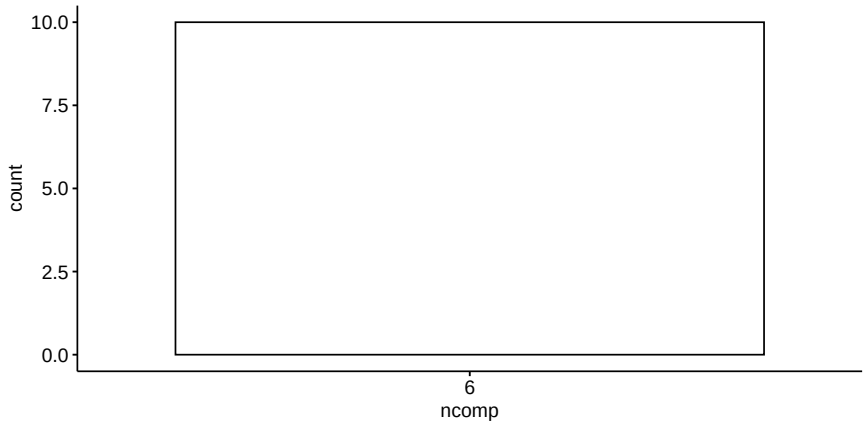

Figure 1. Histogram of the number of components that were most frequently chosen in each fold (over repetitions).

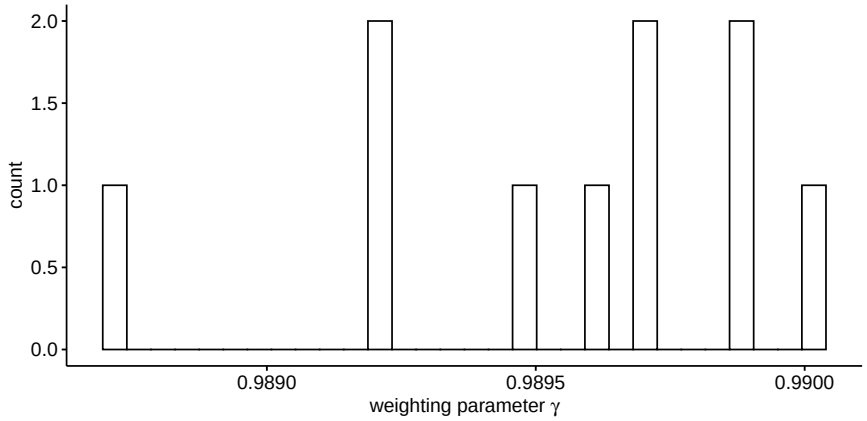

Figure 2. Histogram of the weighting parameter γ for each fold averaged over repetitions.

Note that in PCovR we can perform a rotation on the components. The rotation changes the distribution of the variance explained between components, which also leads to a change of weights of the coefficients to predict Y. However, the overall predictive performance of Y does not change.

In the following we display a PCovR solution with a varimax rotation. The items have been standardized beforehand.

Final rule: PCovR on items with varimax rotation

Note that the colors represent different subscales. For your reference the subscales are:

- selfbl: self blame
- acc: acceptance
- rum: rumination
- posref: positive refocusing
- refpln: refocus on planning
- posreap: positive reappraisal
- putper: putting into perspective
- cat: catastrophizing
- othbl: blaming others

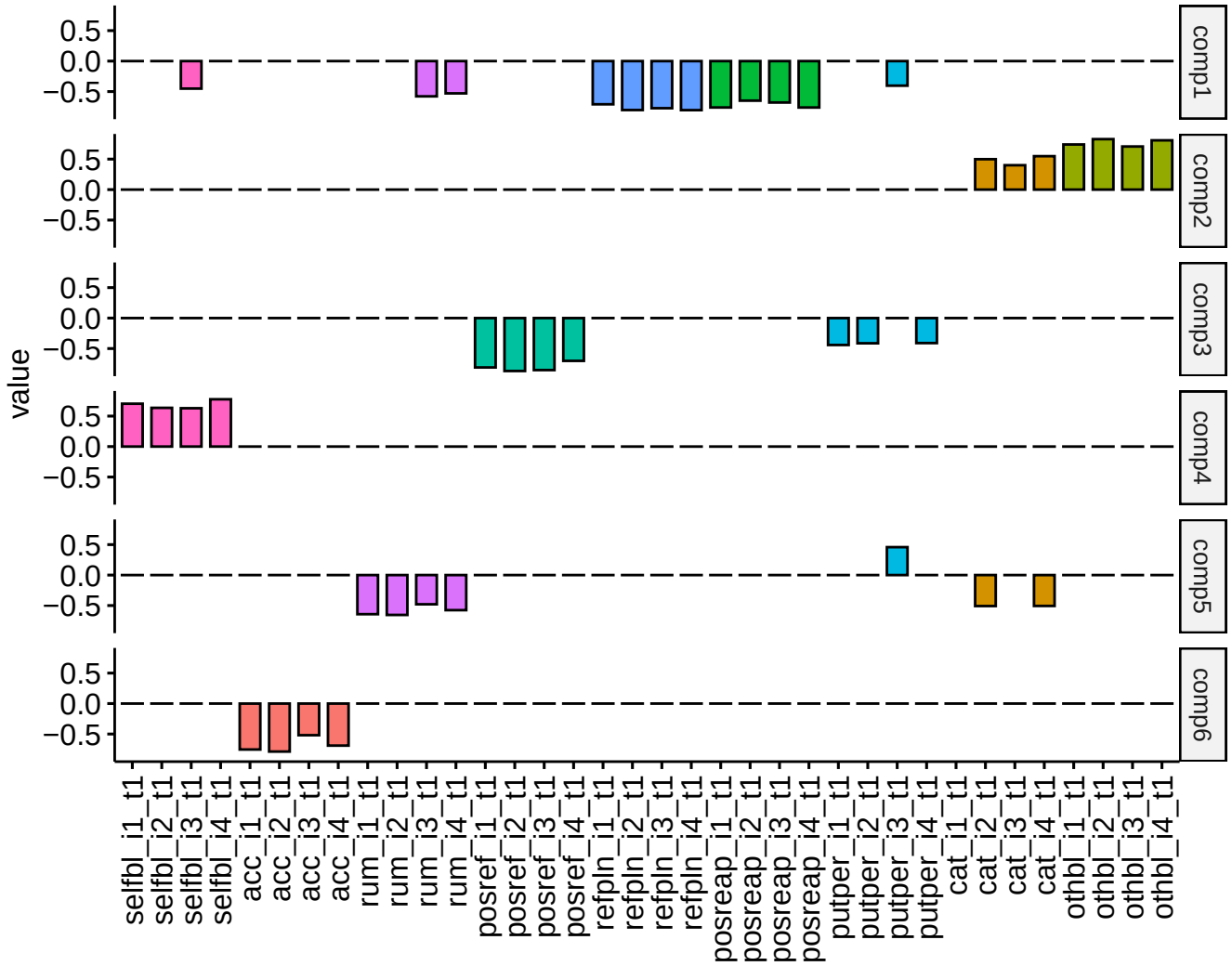

Figure 3. Loadings with absolute value below .4 is set to 0

The weight of the components can be found in Table 1.

Table 1. Standardized coefficients of the component scores from PCovR based on varimax rotation.

	logYdepr_TOTAL
component1	0.144
component2	0.307
component3	-0.043
component4	0.286
component5	-0.341
component6	-0.078

It seems that the components are:

- Component 1: positive coping strategies such as refocus on planning and positive reappraisal
- Component 2: negative cognitive coping strategies such as catastrophizing and blaming others
- Component 3: positive refocusing
- Component 4: self blame
- Component 5: rumination
- Component 6: acceptance

Loadings below .4 in absolute value are removed

	component1	component2	component3	component4	component5	component6
selfbl_i1_t1				0.703		
selfbl_i2_t1				0.635		
selfbl_i3_t1	-0.451			0.629		
selfbl_i4_t1				0.776		
acc_i1_t1						-0.754
acc_i2_t1						-0.788
acc_i3_t1						-0.519
acc_i4_t1						-0.690
rum_i1_t1					-0.642	
rum_i2_t1					-0.654	
rum_i3_t1	-0.578				-0.479	
rum_i4_t1	-0.530				-0.576	
posref_i1_t1			-0.810			
posref_i2_t1			-0.868			
posref_i3_t1			-0.853			
posref_i4_t1			-0.702			
refpln_i1_t1	-0.709					
refpln_i2_t1	-0.803					
refpln_i3_t1	-0.773					
refpln_i4_t1	-0.805					
posreap_i1_t1	-0.761					
posreap_i2_t1	-0.649					
posreap_i3_t1	-0.680					
posreap_i4_t1	-0.762					
putper_i1_t1			-0.442			
putper_i2_t1			-0.413			
putper_i3_t1	-0.403				0.458	
putper_i4_t1			-0.410			
cat_i1_t1						
cat_i2_t1		0.499			-0.508	
cat_i3_t1		0.400				
cat_i4_t1		0.548			-0.506	
othbl_i1_t1		0.740				
othbl_i2_t1		0.828				
othbl_i3_t1		0.707				
othbl_i4_t1		0.809				

Data set 2

For this data there were 161 items that can be summarized to 27 subscales. Twenty six subscales consist of six items each and one subscale consist of five items. Sample size is 92. For the purpose of demonstration we simply chose the rule that had the lowest estimated prediction error, which

was elastic net subscale rules. This means that the original subscales should be retained.

Hyperparameters of Elastic net subscales

As you can see the frequently chosen solution for elastic net subscale rules is when α is 0. This means all the subscales in the PAPI are retained. In order to choose λ we averaged the obtained λ values over folds and repetitions.

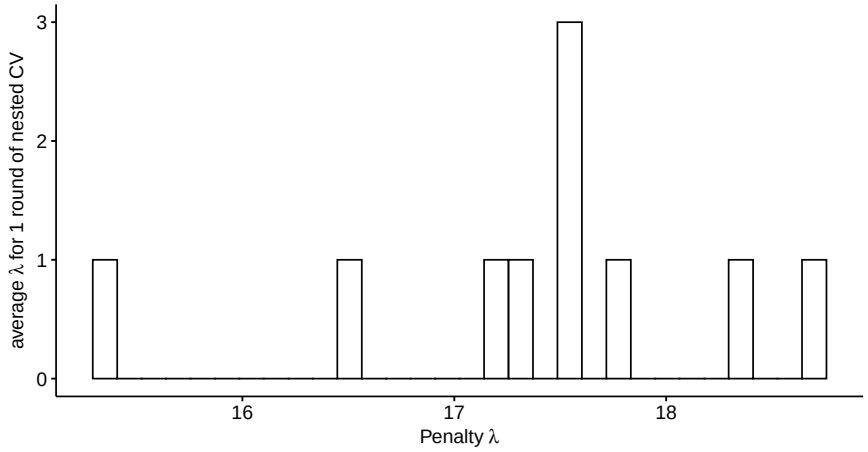

Table 2. Frequency table of the mode of the α values (over 100 repetitions) selected for Elastic net Subscales type across folds

Var1	Freq
0	10

Final rule: Elastic net subscales

Chosen values of the hyperparameters:

- $\alpha = 0$
- $\lambda = 19.558$

Table 3. Coefficient table (raw) from a subscale rule using a ridge penalty

	Scale Names	Value
(Intercept)		50.279
A	Need to achieve	0.104
B	Need to belong to groups	0.022
C	Need to be organised	0.034
D	Attention to detail	0.031
E	Emotional restraint	-0.064
F	Need to be upwardly supportive	0.113
G	Work focus	-0.034
H	Planner	0.035
I	Ease in decision making	0.015
J	Optimism	0.041
K	Need to be direct	0.050
L	Leadership role	0.052
M	Inspirational motivator	0.032
N	Need to finish a task	0.040
O	Need to relate closely to individuals	-0.012
P	Need to influence	0.103
Q	Need to connect	0.072
R	Conceptual thinker	-0.002
S	Social harmoniser	0.040
SD	Social desirability	-0.130
T	Work tempo	0.030
U	Resilience	0.032
V	Persistence	0.004
W	Need for rules and guidelines	0.043
X	Need to be noticed	0.072
Y	Core composure	0.007
Z	Need for change	0.021