**Appendix S.6. User Manual for the MOMCS Program.**

**1. Requirements to run the program.**

The program runs under Windows 10 and Windows 11. To run the program a recent 64 bits of Cygwin must be installed on the computer. To download Cygwin, go to the webpage of Cygwin: https://www.cygwin.com

After installing Cygwin update the installation (i.e., go back to the Cygwin webpage) by adding the libgfortran5 package to your installation. If the latter package is not in your installation, the program does not run!

**2. Installing the program.**

Go to the web site (unmentioned for masked review) and download the program (i.e., the momcs.exe file). Copy the program to a directory that Cygwin has access to.

**3. Running the program.**

Prepare the input file for the selection situation. Details on this are given below. Copy the input file to the same directory that holds the program.

Run Cygwin from your PC. This opens a command window under Cygwin. If necessary change to the directory where both the program and the input file are stored. Assuming that you did not change the name of the program (i.e. momcs.exe) and that the name of your input file is momcs.i, type the following command at the prompt

./momcs < momcs.i

The program starts running. Completion may take seconds to a few minutes, but also hours in high density 5-objective situations. The output of the program is on a single file with the name given in the input file. The latter file contains a list with all the PO systems that is ready for use in R scripts to perform further analyses (e.g., filtering preferred PO solutions, drawing the PO front, etc.).

**4. The input file.**

Here is an example input file.

9 4 10 0

2 4 1 3 15 1

1 2 3 9

1 2 4

6

1.00 0.34 0.57 0.17 0.13 0.20 0.12 0.25 0.45

0.34 1.00 0.11 0.13 0.12 0.09 0.28 0.55 0.10

0.57 0.11 1.00 0.11 0.58 0.28 0.12 0.19 0.16

0.17 0.13 0.11 1.00 0.11 0.18 0.16 0.01 0.03

0.13 0.12 0.58 0.11 1.00 0.03 0.24 0.37 0.36

0.20 0.09 0.28 0.18 0.03 1.00 0.11 0.24 0.34

0.12 0.28 0.12 0.16 0.24 0.11 1.00 0.59 0.07

0.25 0.55 0.19 0.01 0.37 0.24 0.59 1.00 -0.06

0.45 0.10 0.16 0.03 0.36 0.34 0.07 -0.06 1.00

.60 .22 .03 .15

.00 .00 .00 .00 .00 .00 .00 .00 .00

.08 .23 .05 -.49 -.20 -.54 .03 .42 -.83

-.01 .21 .16 -.24 .10 -.38 .17 .28 -.72

-.31 .31 -.06 .04 -.35 -.11 .02 -.05 .06

2 2

1 4

2 3

0.0 0.15

0.6 0.3

bof5

We now comment each of the input lines.

Line 1: 9 4 10 0

The line contains 4 values, specifying the **nvt, ngt, nre** and **ish** input parameter values, with **nvt** (integer): the total number of predictors and criteria; **ngt** (integer): the total number of groups (including the majority group); **nre** (integer): the number of times all nonlinear programs are solved, each time starting from a different random start solution; and **ish** (integer) with value of either 0, meaning full hybrid, NBI and subspace solutions, or value 1, meaning only subspace procedures. Recommended values for **nre** are between 5 and 20. Upper limit values for **nvt** and **ngt** are 10 and 7.

Line 2: 2 4 1 3 15 1

The line contains 6 values, specifying the **nf, np, nt, ng, ns,** and **nd** parameter values, with **nf** (integer): the number of selection stages; **np** (integer): the number of actually used predictors; **nt** (integer): the number of quality goals in the actual selection; **ng** (integer): the total number of groups (including the majority group) in *the actual selection situation*; **ns** (integer); parameter controlling the density of the discrete representation of the full PO front; and **nd** (integer) ): **nd** specifies the chosen metric for the quality and diversity objectives, if **nd** equals 1 then quality corresponds to the expected criterion performance of the selected applicants and diversity corresponds to the AIR, if **nd** equals 2 then quality corresponds to the validity of the selection composite and diversity corresponds to the AIR, if **nd** equals 3 then quality corresponds to the validity of the selection composite and diversity corresponds to the effect size of the selection composite—**nd** equal to 2 or 3 is only permitted if **nf** equals 1. The total number of objectives, **nobj**, equals **nt + ng – 1**. If the **ish** parameter (see above the specification of Line 1) is set equal to 0, the upper limit values for **ns** are 50 (if **nobj** = 3), 20 (if **nobj** = 4) and 10 (if **nobj** = 5). If **is** = 1, then **ns** must not be greater than 1000. The program does not check these limits and may abort when they are not met.

Line 3: 1 2 3 9

The line has as many numbers as the value of **np** and each number indicates the row (and column) of the correlation matrix **pct** (see below) that corresponds to one of the actually used predictors. So, in this case, the actually used predictors 1, 2 and 3 correspond to the first three rows (columns) of the matrix **pct**, whereas predictor 4 corresponds to the 9th row (column) of **pct.**

Line 4: 1 2 4

The line has as many numbers as the value of **ng**, and the numbers indicate the rows of the effect size matrix **ef** (see below) that correspond to the groups included in the selection. The program assumes that the first group is the majority group and that the group is always included in the actual selection. The numbers 2 and 4 in the line correspond to the actual minority groups one and two (i.e., minority group one is the group with effect sizes detailed in row 2 of the **ef** matrix, whereas minority group 2 is the group with effect sizes detailed in row 4 of the **ef** matrix.

Lin 5: ..6

The line has as many numbers as the value of **nt**, with the numbers indicating the rows (columns) of the correlation matrix **pct** (see below) that correspond to the quality criteria of the actual selection

Line 6.1-6.**nvt**

1.00 0.34 0.57 0.17 0.13 0.20 0.12 0.25 0.45

0.34 1.00 0.11 0.13 0.12 0.09 0.28 0.55 0.10

0.57 0.11 1.00 0.11 0.58 0.28 0.12 0.19 0.16

0.17 0.13 0.11 1.00 0.11 0.18 0.16 0.01 0.03

0.13 0.12 0.58 0.11 1.00 0.03 0.24 0.37 0.36

0.20 0.09 0.28 0.18 0.03 1.00 0.11 0.24 0.34

0.12 0.28 0.12 0.16 0.24 0.11 1.00 0.59 0.07

0.25 0.55 0.19 0.01 0.37 0.24 0.59 1.00 -0.06

0.45 0.10 0.16 0.03 0.36 0.34 0.07 -0.06 1.00

In total **nvt** lines that list the elements (decimal point numbers) of the correlation matrix **pct**

Line 7: .60 .22 .03 .15

The line provides the proportion (decimal point numbers) of the **ngt** groups in the total applicant group. The proportions must sum to 1. The program does not check this! If **ng** is smaller than **ngt**, the program automatically adapts the given proportions.

Line 8.1-8.**ngt**

.00 .00 .00 .00 .00 .00 .00 .00 .00

.08 .23 .05 -.49 -.20 -.54 .03 .42 -.83

-.01 .21 .16 -.24 .10 -.38 .17 .28 -.72

-.31 .31 -.06 .04 -.35 -.11 .02 -.05 .06

In total **ngt** lines that list the effect size values (decimal point numbers) of the **ngt** groups (the rows) for the **nvt** (the columns) predictors and criteria. Together the **ngt** lines define the effect size matrix **ef**. The program assumes that the first group is the majority group such that all elements in the first row of the **ef** matrix are all equal to zero.

Line 9: 2 2

The line contains **nf** (integer) numbers, indicating the number of predictors used in the first, the second a.s.o stage until the number used in the final **nf**th stage.

Line 10.1-10.**nf**

1 4

2 3

In total **nf** lines, the elements (integer numbers) in the first line indicate the predictors used in the first stage, the elements in the second line indicate the predictors used in the second stage, a.s.o. Be careful to note that the indication refers to the **ng** numbers listed in Line 3 of the input. So, in this example, the numbers 1 and 4 in the line 10.1 indicate that the predictors used in the first stage correspond to the variable in the first and the nineth row (column) of the **pct** correlation; whereas the predictors used in the second stage correspond to the second and the third row (column) of the **pct** matrix.

Line 11: 0.0 0.15

The line indicates the lower bound (decimal point number) on the admissible predictor weights and the final selection rate (decimal point number). The program enforces that the sum of the predictor weights across the predictors used in the stage is equal 1 for all stages.

Line 12.1-12.(**nf**-1)

0.6 0.3

In total (**nf** -1) lines, the first line indicating the upper and lower bound (decimal point numbers) on the retention rate after the first selection stage, the second line indicating the upper and lower bound on the retention rate after the second selection stage, a.s.o. **However note that the lines 12.1-12.(nf-1) must be omitted if nf equals 1**. Also, take care that the upper bound is at least equal to the lower bound for each intermediate stage and that the upper bound of the retention rate in stage (x+1) is at most equal to the lower bound on the retention rate in the preceding stage x. Again, the program does not check this!

Line 13: **namout**

The line specifies the name of the file (at most 8 characters) holding the list of computed PO-systems.

**5. Using the output of the program.**

The program results in a file with name as specified in the input file. The format of the file is such that it can be read as a data frame in the R programming environment. Thus, the first line specifies the header line and the following lines each specify a solution. Each of these lines contains 4 + nt + (ng-1) + 4 + ng + nf + np elements:

**num**: the number of the system;

**inn**: a value of 1 if the system is PO else a value of 0

**typ**: identifies the solution set the system is part of, with value 1 if the system is obtained by using the modified SBG procedure, a value of 2 if the system is obtained by using the NBI method, a value of 5 if the system is obtained by the subspace procedure using average quality and minimum diversity as aggregated quality and diversity goals, and a value of 6 if the system is obtained by the subspace procedure using average quality and minimum equal diversity (across the minority groups) as aggregated quality and diversity goals;

**comm**: can be ignored;

**QUA1-QUA(nt)**: the value of the **nt** quality trade-off components of the system;

**DIV1-DIV(ng-1)**: the value of the **ng-1** diversity trade-off components of the system;

**AQUA**: average quality trade-off value of the system

**ADIV**: average diversity trade-off value of the system

**ASEG**: no longer used, is set to 0

**DIMI**: minimum diversity trade-off value of the system

SRG1-SRG(**ng**): selection rates for the ng groups

SRS1-SRS(**nf**): retention (selection) rates for the nf stages

PWE1-PWE(**np**): weights of the np predictors.

**Note that the value of the trade-off components is reported in the metric as chosen in Line 2 of the input file for the quality and diversity objectives.** In the present example Line 2 specifies using the expected performance on the quality criterion of the selected applicants as metric for gauging quality, and using the AIR metric for gauging the two diversity objectives. So, the three values for the first PO system (i.e., 0.501, 0.369 and 1.041 indicate that the expected performance of the selected candidates on the quality criterion equals 0.501 and that the AIR for the first and the second minority group equals 0.369 and 1.041, respectively. The expected performance value is standardized with respect to the majority group metric.

The second output file can be read into R as a data frame using the command “read.table(scan(“namefile”),header=T,sep=””)” to perform subsequent analyses.