891   **Appendix A: SSD based on dichotomy search algorithm**

892   In this appendix, an improved algorithm that can effectively speed up the calculation

893   process based on the dichotomy search is introduced. As described in the basic algorithm

894   in the section "The Basic Algorithm Used for Sample Size Determination", a large amount

895   of iterations are conducted for the calculation process from Steps 2-5, which brings great

896   burdens on the computation and costs more time to reach the conditions in Step 5. The

897   presented search-based process can sharply reduce the computing burden by reducing the

898   number of iterations. The basic idea of the proposed algorithm is to gradually adjust the

899   sample size using a dichotomy algorithm until $p_s \geq \eta$ and $p_v \geq \eta$ hold. The Steps 4-5

900   described in the basic algorithm in the section "The Basic Algorithm Used for Sample Size

901   Determination" are replaced by the following steps.

902   *Step 4: Define the lower and upper bound of sample size N for the dichotomy*

903   *search method.*

904   To use the dichotomy search method, the interval of the sample size needs to be

905   determined. Let LB and UB denote the lower and upper bound of the optimal sample size

906   N, where the LB cannot be smaller than 10. In order to narrow the distance between

907   lower and upper bound, the following steps are conducted.

908   • Compute p_s and p_v using Steps 2-4 from the basic algorithm based on $N = 100$.

909   •   (i) If $p_s \geq \eta$ and $p_v \geq \eta$, then set $N = \frac{N}{2}$, and repeat Steps 2-4 and (i) until $p_s < \eta$

910   or $p_v < \eta$. Then set $LB = N$, $UB = 2N$.

911   (ii) If $p_s < \eta$ or $p_v < \eta$, then set $N = 2N$, and repeat Steps 2-4 and (ii) until $p_s \geq \eta$

912   and $p_v \geq \eta$. Then set $LB = \frac{N}{2}$, $UB = N$.

913   *Step 5: Compute the optimal sample size*

914   (I) Based on the LB and UB determined by above, let $N_{\text{mid}} = \frac{LB+UB}{2}$.

915   (II) Compute p_s and p_v with $N = N_{\text{mid}}$ using Steps 2-4.

916   (III) If $p_s \geq \eta$ and $p_v \geq \eta$, $UB = N_{\text{mid}}$; else, $LB = N_{\text{mid}}$.

917   (IV) Update the value of N_{mid} with $N_{\text{mid}} = \frac{LB+UB}{2}$.

918   (V) Return to Step II with $N = N_{\text{mid}}$ and repeat Steps II-V until $N_{\text{mid}} = LB + 1$

919   satisfied. Then $N = N_{\text{mid}}$.

920   **Appendix B: How to calculate the regression coefficients based on the**

921   **coefficient of determination and the ratio among the regression coefficients**

922 In this appendix, the process of calculating the regression coefficients is described if the

923 input ingredients coefficient of determination R^2 and the ratio among the regression

924 coefficients are given instead of the regression coefficients are given directly. This matter

925 has been mentioned in Section "The Basic Algorithm Used for Sample Size Determination".

926 Based on the Model in Equation 1, the regression coefficients can be computed as follows.

927   1. Variance can be calculated on both sides of Equation 1:

$$VAR[y_i] = VAR[\sum_{k=1}^{K} \beta_k x_{ik}] + VAR(\epsilon_i). \tag{11}$$

928   2. Divide by VAR$[y_i]$ on both sides of Equation 11,

$$1 = R^2 + \frac{VAR(\epsilon_i)}{VAR[y_i]}. \tag{12}$$

929   3. As presented in the section "The Basic Algorithm Used for Sample Size

930   Determination", VAR$[x_{ik}]$=1 and VAR$[y_i]$=1. Combining with Equation 11 and 12,

931   the formula for coefficient of determination R^2 can be rewritten as

$$R^2 = VAR[\sum_{k=1}^{K} \beta_k x_{ik}] = \sum_{k=1}^{K} \beta_k^2 + 2 \sum_{k<k\prime} \beta_k \beta_{k\prime} \rho_{kk\prime}, \tag{13}$$

where $\rho_{kk'}$ denotes the correlation between predictor variables x_{ik} and $x_{ik'}$, which is the element in the correlation matrix Σ.

The ratio $\beta_1 : \beta_2 : \cdots : \beta_K$ can be ascertained from pilot study, published results from a similar study, and can be estimated based on the expert's advice and the prior knowledge of the field. If hypotheses H_a and H_1 are considered, and the ratio $\beta_1 : \beta_2 : \cdots : \beta_K = 1 : 1 : \cdots : 1$, then $\beta_1 = \beta_2 = \cdots = \beta_K$. For hypothesis $H_0 : \beta_1 = \beta_2 = \cdots = \beta_K$ itself, the relation of the regression coefficients is also $\beta_1 = \beta_2 = \cdots = \beta_K$. By substituting R^2, Σ, $\beta_1 = \beta_2 = \cdots = \beta_K$ into Equation 13, $\beta_1, \beta_2, \cdots, \beta_K$ can be derived. If hypotheses H_2 is considered, and the ratio $\beta_1 : \beta_2 : \cdots : \beta_K = r_1 : r_2 : \cdots : r_K$, then $\beta_1 = r_1/r_K\beta_K$, $\beta_2 = r_2/r_K\beta_K$, $\cdots$, $\beta_{K-1} = r_{K-1}/r_K\beta_K$. By substituting R^2, Σ, $\beta_1 = r_1/r_K\beta_K$, $\beta_2 = r_2/r_K\beta_K$, $\cdots$, $\beta_{K-1} = r_{K-1}/r_K\beta_K$ into Equation 13, β_K can be derived. Subsequently, β_1, $\beta_2, \cdots, \beta_{K-1}$ can be obtained. It should be noted that the default signs of the regression coefficients are positive for all the hypotheses, unless they are designated to be negative in the hypotheses.

It will now be explained how the population regression coefficients for the complement of H_1 will be determined. Changing the sign of the regression coefficients once is called a violation. The number of violations is determined by the number of the signs of regression coefficients changed. The complement hypothesis of H_1 can be divided into K categories based on the number of violations. That is, if K predictors are considered, there will be K categories, namely one deviation from H_1, two violations from H_1, $\cdots$, K violations from H_1. To facilitate the reader's understanding of the proposed approach, two examples are provided with two and five predictors. Firstly, this paper discusses the simplest situation where only two predictors are needed. In this situation, the hypothesis H_1 can be expressed as $\beta_1 > 0 \& \beta_2 > 0$. The complement of H_1 with two predictors includes the following three cases:

$H_{c1} : \beta_1 < 0 \& \beta_2 > 0$.

$H_{c2} : \beta_1 > 0 \& \beta_2 < 0.$

$H_{c3} : \beta_1 < 0 \& \beta_2 < 0.$

In the above three cases, the hypotheses H_{c1} and H_{c2} have one thing in common. They have one violation from H_1. The order could also be H_{c2}, H_{c1}, and H_{c3}, since the Bayes factor $BF_{c1,1}$ for H_{c1} versus H_1 if data are simulated from populations H_{c1} and $BF_{c2,1}$ for H_{c2} versus H_1 if data are simulated from populations H_{c2} are almost the same. That is, there is no preference for one of these two hypotheses. Different from these two hypotheses, the hypothesis H_{c3} has two violations. The more numbers of violations in the complement of H_1, the easier it is to distinguish the complement hypothesis from the hypothesis H_1. Therefore, the Bayes factor $BF_{c3,1}$ for H_{c3} versus H_1 if data are simulated from population H_{c3} is larger than the $BF_{c1,1}$ and $BF_{c2,1}$ if data are simulated from populations H_{c1} and H_{c2}, respectively. To determine the regression coefficients for the complement of H_1, a representative hypothesis has to be selected. In this paper, the hypothesis corresponding to the median of the number of hypotheses ordered using the number of violations is selected as the representative hypothesis of the complement hypothesis of H_1, in this case, H_{c2}. The measure of median is used because it refers to the most central value in the ascending Bayes factors which are ordered using the number of violations.

Based on the ratio $\beta_1 : \beta_2 = 1:1$ for H_1, the relation of the regression coefficients under the complement hypothesis H_{1c} is $\beta_2 = -\beta_1$, where $\beta_1 > 0$. By substituting R^2, ρ, and $\beta_2 = -\beta_1$ into Equation 13, the regression coefficients can be calculated.

To further clarify the proposed method, a more complex scenario with five predictors is discussed. The hypothesis H_1 can be expressed as $\beta_1 > 0 \& \beta_2 > 0 \& \beta_3 > 0 \& \beta_4 > 0 \& \beta_5 > 0$. To find the hypothesis representative of the complement of H_1, all the possible hypotheses in the complement are ordered using the number of violations. First, only one violation is considered. The number of hypotheses in this case will be $\binom{5}{1}$. In other words, five hypotheses should be considered if one violation happens:

984 $H_{c1}: \beta_1 < 0 \& \beta_2 > 0 \& \beta_3 > 0 \& \beta_4 > 0 \& \beta_5 > 0.$

985 $H_{c2}: \beta_1 > 0 \& \beta_2 < 0 \& \beta_3 > 0 \& \beta_4 > 0 \& \beta_5 > 0.$

986 $H_{c3}: \beta_1 > 0 \& \beta_2 > 0 \& \beta_3 < 0 \& \beta_4 > 0 \& \beta_5 > 0.$

987 $H_{c4}: \beta_1 > 0 \& \beta_2 > 0 \& \beta_3 > 0 \& \beta_4 < 0 \& \beta_5 > 0.$

988 $H_{c5}: \beta_1 > 0 \& \beta_2 > 0 \& \beta_3 > 0 \& \beta_4 > 0 \& \beta_5 < 0.$

989 It should be noted that the order is arbitrary, and any permutation would also be

990 acceptable. The order is irrelevant because all hypothesis containing one violation will lead

991 to about the same Bayes factors $BF_{ci,1}$ for H_{ci} versus H_1 if data are simulated from

992 populations H_{ci} $(i = 1, \cdots, 5)$, respectively. Similarly, in the following, permutations are

993 also arbitrary for the same number of violations.

994 If there are two violations in a hypothesis, the total number of hypotheses will be $\binom{5}{2}$, that

995 is ten. We name these hypotheses as H_{c6}, H_{c7},...and H_{c15}, which are given as:

996 $H_{c6}: \beta_1 < 0 \& \beta_2 < 0 \& \beta_3 > 0 \& \beta_4 > 0 \& \beta_5 > 0.$

997 $H_{c7}: \beta_1 < 0 \& \beta_2 > 0 \& \beta_3 < 0 \& \beta_4 > 0 \& \beta_5 > 0.$

998 $H_{c8}: \beta_1 < 0 \& \beta_2 > 0 \& \beta_3 > 0 \& \beta_4 < 0 \& \beta_5 > 0.$

999 $H_{c9}: \beta_1 < 0 \& \beta_2 > 0 \& \beta_3 > 0 \& \beta_4 > 0 \& \beta_5 < 0.$

1000 $H_{c10}: \beta_1 > 0 \& \beta_2 < 0 \& \beta_3 < 0 \& \beta_4 > 0 \& \beta_5 > 0.$

1001 $H_{c11}: \beta_1 > 0 \& \beta_2 < 0 \& \beta_3 > 0 \& \beta_4 < 0 \& \beta_5 > 0.$

1002 $H_{c12}: \beta_1 > 0 \& \beta_2 < 0 \& \beta_3 > 0 \& \beta_4 > 0 \& \beta_5 < 0.$

1003 $H_{c13}: \beta_1 > 0 \& \beta_2 > 0 \& \beta_3 < 0 \& \beta_4 < 0 \& \beta_5 > 0.$

1004 $H_{c14}: \beta_1 > 0 \& \beta_2 > 0 \& \beta_3 < 0 \& \beta_4 > 0 \& \beta_5 < 0.$

1005 $H_{c15}: \beta_1 > 0 \& \beta_2 > 0 \& \beta_3 > 0 \& \beta_4 < 0 \& \beta_5 < 0.$

1006 Similarly, there are $\binom{5}{3}$ (ten) hypotheses with three violations, which are shown as:

1007 $H_{c16}: \beta_1 < 0 \& \beta_2 < 0 \& \beta_3 < 0 \& \beta_4 > 0 \& \beta_5 > 0.$

1008 $H_{c17}: \beta_1 < 0 \& \beta_2 < 0 \& \beta_3 > 0 \& \beta_4 < 0 \& \beta_5 > 0.$

1009 $H_{c18}: \beta_1 < 0 \& \beta_2 < 0 \& \beta_3 > 0 \& \beta_4 > 0 \& \beta_5 < 0.$

1010 $H_{c19}: \beta_1 < 0 \& \beta_2 > 0 \& \beta_3 < 0 \& \beta_4 < 0 \& \beta_5 > 0.$

1011 $H_{c20} : \beta_1 < 0 \& \beta_2 > 0 \& \beta_3 > 0 \& \beta_4 > 0 \& \beta_5 < 0.$

1012 $H_{c21} : \beta_1 > 0 \& \beta_2 < 0 \& \beta_3 < 0 \& \beta_4 < 0 \& \beta_5 > 0.$

1013 $H_{c22} : \beta_1 > 0 \& \beta_2 < 0 \& \beta_3 < 0 \& \beta_4 > 0 \& \beta_5 < 0.$

1014 $H_{c23} : \beta_1 > 0 \& \beta_2 < 0 \& \beta_3 > 0 \& \beta_4 < 0 \& \beta_5 < 0.$

1015 $H_{c24} : \beta_1 > 0 \& \beta_2 < 0 \& \beta_3 > 0 \& \beta_4 < 0 \& \beta_5 < 0.$

1016 $H_{c25} : \beta_1 > 0 \& \beta_2 > 0 \& \beta_3 < 0 \& \beta_4 < 0 \& \beta_5 < 0.$

1017 There are $\binom{5}{4}$ (five) hypotheses with four violations, which are displayed as follows.

1018 $H_{c26} : \beta_1 < 0 \& \beta_2 < 0 \& \beta_3 < 0 \& \beta_4 < 0 \& \beta_5 > 0.$

1019 $H_{c27} : \beta_1 < 0 \& \beta_2 < 0 \& \beta_3 < 0 \& \beta_4 > 0 \& \beta_5 < 0.$

1020 $H_{c28} : \beta_1 < 0 \& \beta_2 < 0 \& \beta_3 > 0 \& \beta_4 < 0 \& \beta_5 < 0.$

1021 $H_{c29} : \beta_1 < 0 \& \beta_2 > 0 \& \beta_3 < 0 \& \beta_4 < 0 \& \beta_5 < 0.$

1022 $H_{c30} : \beta_1 > 0 \& \beta_2 < 0 \& \beta_3 < 0 \& \beta_4 < 0 \& \beta_5 < 0.$

1023 Finally, the hypotheses with five violations only have $\binom{5}{5}$ (one) case, which is given as

1024 $H_{c31} : \beta_1 < 0 \& \beta_2 < 0 \& \beta_3 < 0 \& \beta_4 < 0 \& \beta_5 < 0.$

1025

1026 As mentioned in the last example, the Bayes factor $\mathrm{BF}_{1c,1}$ for H_{1c} versus H_1 increases with

1027 the number of violations. There are totally $\binom{5}{1} + \binom{5}{2} + \binom{5}{3} + \binom{5}{4} + \binom{5}{5} = 2^5 - 1 = 31$

1028 hypothesis for the complement of H_1. After all the hypotheses are presented, the

1029 representative hypothesis H_{c16} is selected, which is the hypothesis corresponding the

1030 median of the hypotheses ordered using the number of violations. Based on the ratio β_1:

1031 β_2: β_3: β_4: β_5=1:1:1:1:1 for H_1, the relation of the regression coefficients under the

1032 complement hypothesis H_{1c} is $\beta_1 = \beta_2 = \beta_3 = -\beta_4 = -\beta_5$, where β_4 and β_5 are larger than

1033 zero. By substituting R^2, ρ, and $\beta_1 = \beta_2 = \beta_3 = -\beta_4 = -\beta_5$ into Equation 13, the

1034 regression coefficients can be calculated.

1035 In summary, we can conclude that the complement hypotheses of H_1 include

1036 $\binom{k}{1} + \binom{k}{2} + ... + \binom{k}{k} = 2^k - 1$ cases. The hypothesis corresponding to the median number of

1037 $\{1, 2, \cdots, 2^k - 1\}$ (i.e., the hypothesis H_{2^k-1}) is selected as representative of the

1038 complement hypothesis of H_1. If the ratio $\beta_1 : \beta_2 : \cdots : \beta_K = 1 : 1 : \cdots : 1$, then

1039 $\beta_1 = \beta_2 = \cdots \beta_{K-1} = \beta_K$. By substituting R^2, $\boldsymbol{\Sigma}$, $\beta_1 = \beta_2 = \cdots \beta_{K-1} = \beta_K$ into Equation

1040 13, β_K can be derived. Subsequently, β_1, $\beta_2, \cdots$, β_{K-1} can be obtained. The default signs

1041 of the regression coefficients are positive for all the hypotheses at first. If the signs of the

1042 regression coefficients are negative in the representative hypothesis, they will be finally

1043 designated to be negative.

1044 Another issue that needs to be addressed is how to calculate the regression coefficients for

1045 the complement of H_2. Firstly, with two predictors H_2 is given as $\beta_1 > \beta_2$. There is only

1046 one hypothesis for the complement of H_2, namely H_{c1}: $\beta_1 < \beta_2$. There is no doubt that H_{c1}

1047 can be regarded as the representative hypothesis. Based on Supplementary Material, Table

1048 1, if H_2 is true, the ratio β_1: β_2=2:1 for H_2. The relation of the regression coefficients

1049 under the complement hypothesis H_{2c} is β_1: β_2=1:2, where β_1 and β_2 are larger than zero.

1050 By substituting R^2, ρ, and β_1: β_2=1:2 into Equation 13, the regression coefficients can be

1051 calculated.

1052 In following examples, swapping the regression coefficients of adjacent positions once is

1053 called a violation. The number of violations is determined by the number of the regression

1054 coefficients of adjacent positions swapped. This process can be described as follows. For

1055 the convenience of description, a specific example is used to illustrated. For example,

1056 $H_2 : \beta_1 > \beta_2 > \beta_3$ is considered, β_1 and β_2 are at adjacent positions, and β_2 and β_3 are also

1057 at adjacent positions. After swapping adjacent positions β_1 and β_2, a new order

1058 $H_{c1} : \beta_2 > \beta_1 > \beta_3$ can be obtained. In the new order, β_1 and β_3 are at adjacent positions.

1059 Swap them, and a new order can be obtained again, which is $H_{c2} : \beta_2 > \beta_3 > \beta_1$. By

1060 swapping adjacent positions β_2 and β_3, the order $H_{c3} : \beta_3 > \beta_2 > \beta_1$ can be obtained.

1061 During this process, the adjacent positions are swapped for three times. Therefore, there

1062 are three violations from $H_2 : \beta_1 > \beta_2 > \beta_3$ to H_{c3}. The complement hypothesis of H_2 can

1063 be divided into $\binom{K}{2}$ categories based on the number of violations. For three or more

1064 predictors, there is more than one hypothesis for the complement of H_2. Therefore, all the

1065 possible hypotheses should be considered and a representative one should be selected. For

1066 three predictors, there are three categories for the complement hypothesis of H_2, namely

1067 one violation from H_2, two violations from H_2, and three violations from H_2. There are

1068 two hypotheses with one violation:

1069 $H_{c1} : \beta_2 > \beta_1 > \beta_3$.

1070 $H_{c2} : \beta_1 > \beta_3 > \beta_2$.

1071 There are two hypotheses containing two violations:

1072 $H_{c3} : \beta_2 > \beta_3 > \beta_1$.

1073 $H_{c4} : \beta_3 > \beta_1 > \beta_2$.

1074 There is only one hypothesis with three violations:

1075 $H_{c5} : \beta_3 > \beta_2 > \beta_1$.

1076 As the Bayes factor $\text{BF}_{2c,2}$ for H_{2c} versus H_2 becomes larger with an increasing number of

1077 violations, the hypothesis corresponding the median of the number of hypotheses ordered

1078 using the number of violations ($H_{c3} : \beta_2 > \beta_3 > \beta_1$) is selected as the hypothesis

1079 representing the complement of H_2. Based on Supplementary Material, Table 1, if H_2 is

1080 true, the ratio $\beta_1 : \beta_2 : \beta_3 = 3:2:1$ for H_2. The relation of the regression coefficients under the

1081 complement hypothesis H_c is $\beta_1 : \beta_2 : \beta_3 = 1:3:2$ (reordered using $H_{c3} : \beta_2 > \beta_3 > \beta_1$), where

1082 β_1, β_2, and β_3 are larger than zero. By substituting R^2, ρ, and $\beta_1 : \beta_2 : \beta_3 = 1:3:2$ into

1083 Equation 13, the regression coefficients can be calculated.

1084 In order to better summarize the rule of the selection of the representative hypothesis, the

1085 situation with four predictors $H_2 : \beta_1 > \beta_2 > \beta_3 > \beta_4$ is further discussed. The hypotheses

1086 with one, two, three and four violations are:

1087 One violation:

1088 $H_{c1} : \beta_2 > \beta_1 > \beta_3 > \beta_4$.

1089 $H_{c2} : \beta_1 > \beta_3 > \beta_2 > \beta_4$.

1090 $H_{c3} : \beta_1 > \beta_2 > \beta_4 > \beta_3.$

1091 two violations:

1092 $H_{c4} : \beta_2 > \beta_3 > \beta_1 > \beta_4.$

1093 $H_{c5} : \beta_2 > \beta_1 > \beta_4 > \beta_3.$

1094 $H_{c6} : \beta_3 > \beta_1 > \beta_2 > \beta_4.$

1095 $H_{c7} : \beta_1 > \beta_3 > \beta_4 > \beta_2.$

1096 $H_{c8} : \beta_1 > \beta_4 > \beta_2 > \beta_3.$

1097 three violations:

1098 $H_{c9} : \beta_3 > \beta_2 > \beta_1 > \beta_4.$

1099 $H_{c10} : \beta_2 > \beta_3 > \beta_4 > \beta_1.$

1100 $H_{c11} : \beta_3 > \beta_1 > \beta_4 > \beta_2.$

1101 $H_{c12} : \beta_1 > \beta_4 > \beta_3 > \beta_2.$

1102 $H_{c13} : \beta_4 > \beta_1 > \beta_2 > \beta_3.$

1103 four violations:

1104 $H_{c14} : \beta_3 > \beta_2 > \beta_4 > \beta_1.$

1105 $H_{c15} : \beta_2 > \beta_4 > \beta_3 > \beta_1.$

1106 $H_{c16} : \beta_3 > \beta_4 > \beta_1 > \beta_2.$

1107 $H_{c17} : \beta_4 > \beta_1 > \beta_3 > \beta_2.$

1108 $H_{c18} : \beta_4 > \beta_2 > \beta_1 > \beta_3.$

1109 five violations:

1110 $H_{c19} : \beta_3 > \beta_4 > \beta_1 > \beta_2.$

1111 $H_{c20} : \beta_4 > \beta_2 > \beta_3 > \beta_1.$

1112 $H_{c21} : \beta_2 > \beta_4 > \beta_1 > \beta_3.$

1113 $H_{c22} : \beta_4 > \beta_3 > \beta_1 > \beta_2.$

1114 six violations:

1115 $H_{c23} : \beta_4 > \beta_3 > \beta_2 > \beta_1.$

1116 Overall, there are 23 hypotheses for the complement of H_2. Similarly, the hypothesis

corresponding to the median of the number of hypotheses ordered using the number of

violations ($H_{c12} : \beta_1 > \beta_4 > \beta_3 > \beta_2$) is recommended as the representative hypothesis of

the complement hypothesis of H_2. Based on Table 1, if H_2 is true, the ratio β_1: β_2: β_3:

β_4=4:3:2:1 for H_2. The relation of the regression coefficients under the complement

hypothesis H_{2c} is β_1: β_2: β_3: β_4=4:1:2:3 (reordered using $H_{c12}: \beta_1 > \beta_4 > \beta_3 > \beta_2$), where

β_1, β_2, β_3 and β_4 are larger than zero. By substituting R^2, ρ, and β_1: β_2: β_3: β_4=4:1:2:3

into Equation 13, the regression coefficients can be calculated.

By summarizing the current examples, the total number of hypotheses in the complement

of H_2 is $K! - 1$. The hypothesis corresponding to the median of the number of hypotheses

ordered using the number of violations (i.e., the hypothesis $H_{K!/2}$) can be selected as the

representative hypothesis. If the ratio $\beta_1 : \beta_2 : \cdots : \beta_K = r_1 : r_2 : \cdots : r_K$ for H_2, the ratio

of the complement hypothesis of H_2 would be obtained based on the order of hypothesis

$H_{K!/2}$. By substituting R^2, $\boldsymbol{\Sigma}$, and the ratio of the complement hypothesis into Equation

13, β_K can be derived. Subsequently, β_1, β_2,$\cdots$, β_{K-1} can be obtained.

Some researchers may recommend placing the regression coefficients under the complement

of H_1/H_2 on the boundary of H_1/H_2 (set all regression coefficients equal to 0). Although

the boundary value does not belong to H_1 or H_2, it is the closest value to H_1 or H_2.

However, the value of the Bayes factors $\mathrm{BF}_{1c,1}$ for H_{1c} versus H_1 or $\mathrm{BF}_{2c,2}$ for H_{2c} versus

H_2 would be always around 1 no matter how large the sample size is. That is, neither

hypothesis is preferred over the other. Therefore, SSD cannot be performed if the

regression coefficients for the complement are based on the boundary of the parameter

space of H_1 or H_2.

Table 1

Chosen population values for the regression coefficients in the multivariate linear model with two predictors.

	$\rho = 0$		$\rho = 0.2$		$\rho = 0.5$	
	β_1	β_2	β_1	β_2	β_1	β_2
H_0: $\beta_1 = \beta_2 = 0$	0	0	0	0	0	0
H_0: $\beta_1 = \beta_2$	0.255	0.255	0.233	0.233	0.208	0.208
H_a	0.255	0.255	0.233	0.233	0.208	0.208
H_1: $\beta_1 > 0$ & $\beta_2 > 0$	0.255	0.255	0.233	0.233	0.208	0.208
H_2: $\beta_1 > \beta_2$	0.322	0.161	0.299	0.150	0.272	0.136
H_{1c}	-0.255	0.255	-0.233	0.233	-0.208	0.208
H_{2c}	0.161	0.322	0.150	0.299	0.136	0.272

Note: For H_1 and H_a, the ratio of the regression $\beta_1 : \beta_2 = 1 : 1$ is used. For hypothesis H_2, the ratio of the regression $\beta_1 : \beta_2 = 2 : 1$ is used.

Table 2

Chosen population values for the regression coefficients in the multivariate linear model with three predictors.

	$\rho = 0$			$\rho = 0.2$			$\rho = 0.5$		
	β_1	β_2	β_3	β_1	β_2	β_3	β_1	β_2	β_3
H_0: $\beta_1 = \beta_2 = \beta_3 = 0$	0	0	0	0	0	0	0	0	0
H_0: $\beta_1 = \beta_2 = \beta_3$	0.208	0.208	0.208	0.176	0.176	0.176	0.147	0.147	0.147
H_a	0.208	0.208	0.208	0.176	0.176	0.176	0.147	0.147	0.147
H_1: $\beta_1 > 0$ & $\beta_2 > 0$ & $\beta_3 > 0$	0.208	0.208	0.208	0.176	0.176	0.176	0.147	0.147	0.147
H_2: $\beta_1 > \beta_2 > \beta_3$	0.289	0.193	0.096	0.252	0.168	0.084	0.216	0.144	0.072
H_{1c}	-0.208	-0.208	0.208	-0.176	-0.176	0.176	-0.147	-0.147	0.147
H_{2c}	0.096	0.289	0.193	0.084	0.252	0.168	0.072	0.216	0.144

For H_1 and H_a, the ratio of the regression $\beta_1 : \beta_2 : \beta_3 = 1 : 1 : 1$ is used. For hypothesis H_2, the ratio of the regression $\beta_1 : \beta_2 : \beta_3 = 3 : 2 : 1$ is used.

Table 3

Chosen population values for the regression coefficients in the multivariate linear model with four predictors.

	$\rho = 0$				$\rho = 0.2$				$\rho = 0.5$			
	β_1	β_2	β_3	β_4	β_1	β_2	β_3	β_4	β_1	β_2	β_3	β_4
H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$	0	0	0	0	0	0	0	0	0	0	0	0
H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4$	0.180	0.180	0.180	0.180	0.142	0.142	0.142	0.142	0.114	0.114	0.114	0.114
H_a	0.180	0.180	0.180	0.180	0.142	0.142	0.142	0.142	0.114	0.114	0.114	0.114
H_1: $\beta_1 > 0$ & $\beta_2 > 0$ & $\beta_3 > 0$ & $\beta_4 > 0$	0.180	0.180	0.180	0.180	0.142	0.142	0.142	0.142	0.114	0.114	0.114	0.114
H_2: $\beta_1 > \beta_2 > \beta_3 > \beta_4$	0.263	0.197	0.132	0.066	0.217	0.163	0.109	0.054	0.179	0.134	0.089	0.045
H_{1c}	-0.180	-0.180	0.180	0.180	-0.142	-0.142	0.142	0.142	-0.114	-0.114	0.114	0.114
H_{2c}	0.132	0.263	0.066	0.197	0.109	0.217	0.054	0.163	0.089	0.179	0.045	0.134

For H_1 and H_a, the ratio of the regression $\beta_1 : \beta_2 : \beta_3 : \beta_4 = 1 : 1 : 1 : 1$ is used. For hypothesis H_2, the ratio of the regression $\beta_1 : \beta_2 : \beta_3 : \beta_4 = 4 : 3 : 2 : 1$ is used.

Table 4

The required sample size and the corresponding probability that the Bayes factor is larger than 3 when the $\mathrm{BF}_{thresh} = 3$, $\eta = 0.8$, $R^2 = 0.13$, *ratio* $\beta_1 : \beta_2 = 1 : 1$ *for* H_a *and* H_1, *ratio* $\beta_1 : \beta_2 = 2 : 1$ *for* H_2 *and the number of predictors is 2.*

		$\rho = 0$			$\rho = 0.2$			$\rho = 0.5$		
		$b = J/N$	$b = 2J/N$	$b = 3J/N$	$b = J/N$	$b = 2J/N$	$b = 3J/N$	$b = J/N$	$b = 2J/N$	$b = 3J/N$
H_0: $\beta_1 = \beta_2 = 0$	p_0	0.948	0.880	0.801	0.948	0.881	0.801	0.948	0.881	0.801
	N	121 (122)	104 (105)	95 (96)	121 (122)	105 (104)	93 (93)	121 (122)	105 (106)	93 (93)
H_a	p_a	0.804	0.800	0.804	0.804	0.805	0.802	0.804	0.805	0.802
H_0: $\beta_1 = \beta_2$	p_0	0.962	0.944	0.931	0.973	0.955	0.944	0.980	0.971	0.961
	N	887 (889)	816 (815)	775 (773)	1326 (1328)	1221 (1220)	1176 (1176)	2721 (2725)	2527 (2525)	2426 (2426)
H_2: $\beta_1 > \beta_2$	p_2	0.802	0.804	0.800	0.806	0.805	0.801	0.804	0.805	0.801
H_0: $\beta_1 = \beta_2 = 0$	p_0	0.939	0.862	0.805	0.940	0.862	0.804	0.938	0.860	0.803
	N	90 (92)	74 (72)	71 (70)	88 (88)	71 (73)	71 (70)	86 (88)	70 (71)	69 (69)
H_1: $\beta_1 > 0$ & $\beta_2 > 0$	p_1	0.802	0.811	0.837	0.808	0.805	0.843	0.804	0.805	0.842
H_1: $\beta_1 > 0$ & $\beta_2 > 0$	p_1		0.950			0.968			0.987	
	N		60 (60)			79 (77)			134 (133)	
H_{1c}: not H_1	p_{1c}		0.802			0.800			0.804	
H_2: $\beta_1 > \beta_2$	p_2		0.804			0.812			0.801	
	N		163 (162)			235 (233)			438 (435)	
H_{2c}: not H_2	p_{2c}		0.808			0.800			0.800	

The correlation between the predictors is denoted as ρ, where ρ is chosen as 0, 0.2 or 0.5, respectively in this table. The prior distributions with different b are used to perform a sensitivity analysis. It should be noted that b is not relevant for the evaluation of H_1 and H_2 versus their respective complements. The symbol p_i denotes the probability that the Bayes factor supports H_i when H_i is true, where $i = 0, 1, 2, a, 1c, 2c$. The sample size values in the parentheses is calculated with the set.seed 1234. Comparing the values without and within parenthesis shows that using T=10,000 renders stable sample size estimates.

Table 5

The required sample size and the corresponding probability that the Bayes factor is larger than 3 when the $\mathrm{BF}_{thresh} = 3$, $\eta = 0.8$, $R^2 = 0.13$, ratio $\beta_1 : \beta_2 : \beta_3 = 1 : 1 : 1$ for H_a and H_1, ratio $\beta_1 : \beta_2 : \beta_3 = 3 : 2 : 1$ for H_2 and the number of the predictors is 3.

		$\rho = 0$			$\rho = 0.2$			$\rho = 0.5$		
		$b = J/N$	$b = 2J/N$	$b = 3J/N$	$b = J/N$	$b = 2J/N$	$b = 3J/N$	$b = J/N$	$b = 2J/N$	$b = 3J/N$
H_0: $\beta_1 = \beta_2 = \beta_3 = 0$	p_0	0.974	0.909	0.831	0.973	0.918	0.840	0.973	0.918	0.840
	N	148	119	104	146	120	105	146	120	105
H_a	p_a	0.809	0.806	0.805	0.803	0.804	0.806	0.803	0.804	0.806
H_0: $\beta_1 = \beta_2 = \beta_3$	p_0	0.993	0.983	0.976	0.995	0.992	0.986	0.998	0.996	0.993
	N	776	675	624	1350	1210	1124	3301	2981	2794
H_2: $\beta_1 > \beta_2 > \beta_3$	p_2	0.803	0.806	0.803	0.802	0.804	0.801	0.802	0.804	0.800
H_0: $\beta_1 = \beta_2 = \beta_3 = 0$	p_0	0.964	0.879	0.811	0.964	0.887	0.809	0.968	0.899	0.806
	N	100	71	66	98	71	62	94	69	59
H_1: $\beta_1 > 0$ & $\beta_2 > 0$&$\beta_3 > 0$	p_1	0.802	0.802	0.833	0.807	0.802	0.833	0.802	0.801	0.825
H_1: $\beta_1 > 0$ & $\beta_2 > 0$ & $\beta_3 > 0$	p_1		0.801			0.809			0.809	
	N		35			40			47	
H_{1c}: not H_1	p_{1c}		0.833			0.827			0.814	
H_2: $\beta_1 > \beta_2 > \beta_3$	p_2		0.801			0.801			0.808	
	N		254			410			882	
H_{2c}: not H_2	p_{2c}		0.909			0.909			0.901	

The correlation between the predictors is denoted as ρ, where ρ is chosen as 0, 0.2 or 0.5, respectively in this table. The prior distributions with different b are used to perform a sensitivity analysis. It should be noted that b is not relevant for the evaluation of H_1 and H_2 versus their respective complements. The symbol p_i denotes the probability that the Bayes factor supports H_i when H_i is true, where $i = 0, 1, 2, a, 1c, 2c$.

Table 6

The required sample size and the corresponding probability that the Bayes factor is larger than 3 when the $BF_{thresh} = 3$, $\eta = 0.8$, $R^2 = 0.13$, ratio $\beta_1 : \beta_2 : \beta_3 : \beta_4 = 1 : 1 : 1 : 1$ for H_a and H_1, ratio $\beta_1 : \beta_2 : \beta_3 : \beta_4 = 4 : 3 : 2 : 1$ for H_2 and the number of the predictors is 4.

			$\rho = 0$			$\rho = 0.2$			$\rho = 0.5$	
		$b = J/N$	$b = 2J/N$	$b = 3J/N$	$b = J/N$	$b = 2J/N$	$b = 3J/N$	$b = J/N$	$b = 2J/N$	$b = 3J/N$
H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$	p_0	0.986	0.934	0.840	0.984	0.935	0.840	0.984	0.935	0.840
	N	173	134	109	172	135	109	172	135	109
H_a	p_a	0.813	0.803	0.800	0.808	0.805	0.800	0.808	0.805	0.800
H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4$	p_0	0.998	0.993	0.989	0.999	0.998	0.996	0.999	0.999	0.998
	N	762	652	583	1580	1373	1257	4298	3901	3580
H_2: $\beta_1 > \beta_2 > \beta_3 > \beta_4$	p_2	0.800	0.803	0.805	0.802	0.800	0.800	0.801	0.807	0.810
H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$	p_0	0.982	0.896	0.800	0.980	0.905	0.805	0.981	0.906	0. 809
	N	109	72	65	103	70	61	102	67	56
H_1: $\beta_1 > 0$ & $\beta_2 > 0$ & $\beta_3 > 0$ & $\beta_4 > 0$	p_1	0.800	0.804	0.859	0.807	0.806	0.855	0.806	0.803	0.844
H_1: $\beta_1 > 0$ & $\beta_2 > 0$ & $\beta_3 > 0$ & $\beta_4 > 0$	p_1		0.892			0.948			0.986	
	N		56			83			167	
H_{1c}: not H_1	p_{1c}		0.805			0.802			0.803	
H_2: $\beta_1 > \beta_2 > \beta_3 > \beta_4$	p_2		0.803			0.802			0.802	
	N		272			488			1148	
H_{2c}: not H_2	p_{2c}		0.860			0.854			0.857	

The correlation between the predictors is denoted as ρ, where ρ is chosen as 0, 0.2 or 0.5, respectively in this table. The prior distributions with different b are used to perform a sensitivity analysis. It should be noted that b is not relevant for the evaluation of H_1 and H_2 versus their respective complements. The symbol p_i denotes the probability that the Bayes factor supports H_i when H_i is true, where $i = 0, 1, 2, a, 1c, 2c$.

Table 7

Comparison between classical sample size determined using power=0.8, $\alpha = 0.05$, $f^2 = 0.15$ and the Bayesian sample size determined using $\eta = 0.8$, $BF_{thresh} = 3$, $R^2 = 0.13$, the ratio between each pair of coefficients is 1:1.

$K = 2$	H_0: $\beta_1 = \beta_2 = 0$ vs H_a	classical	68		
		Bayesian	$b = J/N$	$b = 2J/N$	$b = 3J/N$
			121	104	95
	H_0: $\beta_1 = \beta_2 = 0$ vs H_1: $\beta_1 > 0$ & $\beta_2 > 0$	Bayesian	90	74	71
$K = 3$	H_0: $\beta_1 = \beta_2 = \beta_3 = 0$ vs H_a	classical	77		
		Bayesian	$b = J/N$	$b = 2J/N$	$b = 3J/N$
			148	119	104
	H_0: $\beta_1 = \beta_2 = \beta_3 = 0$ vs H_1: $\beta_1 > 0$ & $\beta_2 > 0$ & $\beta_3 > 0$	Bayesian	100	71	66
$K = 4$	H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ vs H_a	classical	85		
		Bayesian	$b = J/N$	$b = 2J/N$	$b = 3J/N$
			173	134	109
	H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ vs H_1: $\beta_1 > 0$ & $\beta_2 > 0$ & $\beta_3 > 0$ & $\beta_4 > 0$	Bayesian	109	72	65