

Supplementary Materials to “An Introduction to Bayesian Hypothesis Testing for Management Research”

In these supplementary materials, we explain how to conduct Bayesian hypothesis testing for linear regression using both the statistical programming language R and the on-line applet developed by Jeff Rouder (<http://pcl.missouri.edu/bf-reg>). Importantly, researchers can fit their statistical model using a standard tool of their choice (i.e., Mplus, SPSS, R, etc.) and use the output from these programs to obtain a Bayes factors for H1 (i.e. there is a linear relationship between the independent and dependent variables) to H0 (i.e. there is no relationship between independent and dependent variables).

APPENDIX A

Bayesian Hypothesis Testing with the R Programming Language

Below, we provide functions implemented in the R programming language that can be used to compute a Bayes factor from the R^2 given the number of covariates and sample size (function “R2.to.bf”). In order to use these functions you need to have the R programming language (version 3.0 or higher) installed on your computer (<http://cran.r-project.org/>).

Here we provide a package with R-scripts and a data set. The package is available at <http://tinyurl.com/onlineSupplement> and contains the following files:

- *data.csv* – simulated data based on Dierdorff et al. (2012) study, saved in comma-delimited format;
- *covariateAnalysis.R* – an R-script that reproduces the results shown in Table 3 of the main text;
- *modelComparison.R* – an R-script that reproduces the results shown in Table 4 of the main text;
- *R2tobf.R* – an R function that generates all possible Bayes factors for N models, given the data sample, number of parameters in each model, and their respective R^2 .

Conducting the Analysis in R

To replicate the results of covariate analysis (results are presented in Table 3 of the main text), run the file “covariateAnalysis.R” in R. To replicate the results of model comparison (results are presented in Table 4 of the main text), run the file “modelComparison.R”.

Getting started

Before running either file, first unpack all the files in the online zip-folder listed above into one folder on your computer. Next, set the working directory in R to that folder, either by typing a command in the R-console:

```
> setwd("path to your file")
```

or by choosing “Misc → Change Working Directory” in the Menu of R.

Run the R scripts

To run the R file either type in the console:

```
> source("covariateAnalysis.R")
```

or

```
> source("modelComparison.R")
```

or choose in the R-menu “File → Source File”. In the next two paragraphs, we provide a step-by-step explanation of R scripts, “covariateAnalysis.R” and “modelComparison.R”.

Replicating Covariate Analysis from the Simulated Data Based on Dierdorff et al. (2012) Using R

In order to replicate the results presented in Table 3 of the main text please execute the following steps:

(1) Load data set “data.csv” and use all columns apart from participant number:

```
> originalData <- read.table("data.csv", sep=";", header=T)
> dat <- originalData[,2:8]
```

(2) Define both the full regression model and Models 2-5 (testing hypotheses 2-5):

```
> modelf = V6 ~ (V1 + V2 + V3 + V4 + V5 + V7) + (V2 + V3 + V4 + V5):V1
> model2 = V6 ~ (V1 + V2 + V3 + V4 + V5 + V7) + (V3 + V4 + V5):V1
> model3 = V6 ~ (V1 + V2 + V3 + V4 + V5 + V7) + (V2 + V4 + V5):V1
> model5 = V6 ~ (V1 + V2 + V3 + V4 + V5 + V7) + (V2 + V3 + V5):V1
> model4 = V6 ~ (V1 + V2 + V3 + V4 + V5 + V7) + (V2 + V3 + V4):V1
```

(3) Fit each regression model and print the result to the console, for example for the full model type:

```
> fit = lm(modelf, data = dat)
```

```
> print(summary(fit))
```

Write down R^2 for each model. Alternatively, you can use different software to fit the regression models.

(4) Once you have obtained the R^2 values, the sample size, and the number of parameters for each model, you can compute Bayes factors for each model over the null model (BF_{10}), as well as every model n over every other model m (BF_{nm}). Save the relevant information as variables in R. Save the sample size into a variable labelled “n”:

```
> n <- 198
```

Save the number of predictors of all 5 models into a vector labeled “nPred”:

```
> nPred <- c(10, 9, 9, 9, 9)
```

Save the R^2 values in a vector labeled “r2”:

```
> r2 <- c(.7041, .7022, .6980, .6493, .6953)
```

Next, call a function labeled “R2.to.bf()” which computes all possible Bayes factors:

```
> R2.to.bf(n, nPred, r2)
```

Replicating Model Comparison from the Simulated Data Based on Dierdorff et al. (2012) Using R

In order to replicate the results presented in Table 4 of the main text, follow the same steps as in the previous section. In Step 2, define the full regression model and Models 1-14 (testing hypotheses 2-5 alone and in different combinations):

```
> modelf = V6 ~ (V1 + V2 + V3 + V4 + V5 + V7) + (V2 + V3 + V4 + V5):V1
```

```

> model11 = V6 ~ (V1 + V2 + V3 + V4 + V5 + V7) + (V2 + V3 + V5):V1
> model12 = V6 ~ (V1 + V2 + V3 + V4 + V5 + V7) + (V2 + V3 + V4):V1
> model13 = V6 ~ (V1 + V2 + V3 + V4 + V5 + V7) + (V2 + V4 + V5):V1
> model14 = V6 ~ (V1 + V2 + V3 + V4 + V5 + V7) + (V3 + V4 + V5):V1
> model15 = V6 ~ (V1 + V2 + V3 + V4 + V5 + V7) + (V2 + V3):V1
> model16 = V6 ~ (V1 + V2 + V3 + V4 + V5 + V7) + (V2 + V5):V1
> model17 = V6 ~ (V1 + V2 + V3 + V4 + V5 + V7) + (V2 + V4):V1
> model18 = V6 ~ (V1 + V2 + V3 + V4 + V5 + V7) + (V3 + V5):V1
> model19 = V6 ~ (V1 + V2 + V3 + V4 + V5 + V7) + (V3 + V4):V1
> model110 = V6 ~ (V1 + V2 + V3 + V4 + V5 + V7) + (V5 + V4):V1
> model111 = V6 ~ (V1 + V2 + V3 + V4 + V5 + V7) + (V2):V1
> model112 = V6 ~ (V1 + V2 + V3 + V4 + V5 + V7) + (V3):V1
> model113 = V6 ~ (V1 + V2 + V3 + V4 + V5 + V7) + (V5):V1
> model114 = V6 ~ (V1 + V2 + V3 + V4 + V5 + V7) + (V4):V1

```

In Step 5, compute Bayes factors of each model over the null model (BF_{10}), as well as every model n over every other model m (BF_{nm}). Save the sample size in a variable labelled “n”:

```
> n <- 198
```

Save the number of predictors of all 15 models in a vector labelled “nPred”:

```
> nPred <- c(10, 9, 9, 9, 9, 8, 8, 8, 8, 8, 8, 7, 7, 7, 7)
```

Save the R^2 values in a vector labelled “r2”:

```
> r2 <- c(.7041, .6493, .6953, .6980, .7022, .6460, .6367, .6892, .6477, .6939,
.6950, .6336, .6446, .6337, .6886)
```

Next, call the function “R2.to.bf” to compute the Bayes factors:

```
> R2.to.bf(n, nPred, r2)
```

Note that the main difference between the covariate analysis and model comparison methods is the definition of regression models in Step 2, which later results in different values of variables “nPred”, and “r2”.

APPENDIX B

Replicating Analysis from the Simulated Data Based on the Dierdorff et al. (2012)

study

Using Rouder's Web-Applet

Here, we show how to compute Bayes factors for the simulated data based on Dierdorff et al. (2012). To replicate results presented in Table 3, you first need to fit 5 regression models – a full model containing all predictors (covariates), and four models containing all predictors but one, which is the predictor referring to the hypothesis of interest. For formal definition of the regression model, see Step 2 of Appendix A and Table 3 of the main text. The analysis can be done with software of your choice. You can also skip the model fitting and use R^2 values from Table 3 of the main text.

Next, go to the web-page <http://pcl.missouri.edu/bf-reg> and use the sample size, number of parameters, and R^2 values in order to compute a Bayes factor of any model n over the null model (BF_{n0}) containing no covariates and therefore explaining zero variance in the data (i.e. $R^2 = 0$). Figure B1 shows a screenshot of the web-page before and after computing the Bayes factor. In the covariate analysis (see Table 3 of the main text), there is one full model and four models testing hypotheses 2-5. Therefore, we obtain BF_{70} , BF_{20} , BF_{30} , BF_{40} and BF_{50} .

As shown in the upper panel of Figure B1, for the full model, we enter the sample size ($N = 198$), number of covariates ($N_{\text{covariates}} = 10$), and unadjusted R^2 ($R^2 = .7041$). After clicking “Submit”, as in the bottom of Figure B2, we obtain a Bayes factor of the full model over the null model ($BF_{70} = 5.238160 \times 10^{40}$). We repeat the procedure to obtain

Bayes factor of Model 2 over the null model ($N = 198$, $N_{\text{covariates}} = 9$, $R^2 = .7022$, $\text{BF}_{20} = 2.382260 \times 10^{41}$).

Further, to compare the full model over Model 2, we need to compute BF_{f2} . This can be done manually, such that $\text{BF}_{f2} = \text{BF}_{f0} / \text{BF}_{20}$. Therefore, $\text{BF}_{f2} = 5.238160 \times 10^{40} / 2.382260 \times 10^{41} \approx 4.55$. This leads to the conclusion that the observed data are about 4.5 times more likely under Model 2 than under the full model. The same way, we compute BF_{f3} , BF_{f5} , and BF_{f4} .

This procedure can be applied to any set of nested linear regression models, given that the sample size, the number of covariates, and the unadjusted R^2 are known. To replicate the results of the model comparison method outlined in the manuscript, use the R^2 values provided in Table 4 of the main text.

Figure B1

Screenshot of Jeff Rouder's Web-Applet for JZS Bayes Factor Regression Analysis (<http://pcl.missouri.edu/bf-reg>). Top Panel: Applet Input, Requiring Specification of the Number of Covariates, the Number of Observations, and the Model's R^2 Value. Bottom Panel: Applet Output, Showing the Bayes Factor for the Model at Hand Versus the Null Model Without any Covariates.

Publications ▾ Bayes Factor Calculators ▾ Code/Data/S

A

Perception and Cognition Lab
Department of Psychological Sciences, University of Missouri

Bayes Factor for Regression

Sample Size:

Number of Covariates:

Unadjusted R-squared :

B

Regression

Input is:
Sample Size N = 198
R-squared = 0.7041
Number of Covariates = 10
Output are odds of Model over Null Model:
JZS Bayes Factor = 5.23816e+40
(Accuracy in Gaussian Quadrature Integration is 5.092471e+36)