organisatiosiloModelConstruction

This is an R Markdown Notebook. When you execute code within the notebook, the results appear beneath the code.

This notebook has been used to create a Linear Mixed Effect Model for the organisation silo analysis.

```
# Package names
packages <- c("carData", "car", "Matrix", "lme4", "LMERConvenienceFunctions")

# Install packages not yet installed
installed_packages <- packages %in% rownames(installed.packages())
if (any(installed_packages == FALSE)) {
  install.packages(packages[!installed_packages])
}

# Packages loading
invisible(lapply(packages, library, character.only = TRUE))
```

```
## Warning: il pacchetto 'carData' è stato creato con R versione 4.2.3
```

```
## Warning: il pacchetto 'car' è stato creato con R versione 4.2.3
```

```
## Warning: il pacchetto 'Matrix' è stato creato con R versione 4.2.3
```

```
## Warning: il pacchetto 'lme4' è stato creato con R versione 4.2.3
```

```
## Warning: il pacchetto 'LMERConvenienceFunctions' è stato creato con R versione
## 4.2.3
```

Datasets load

```
# Reading input hofstede data.
data <- read.csv("./organisation_silo_metrics_hofstede.csv", sep = ";", header = TRUE, stringsAsFactors=
# Reading input trompenaars data.
dataT <- read.csv("./organisation_silo_metrics_trompenaars.csv", sep = ";", header = TRUE, stringsAsFac
# Reading input globe data.
dataG <- read.csv("./organisation_silo_metrics_globe.csv", sep = ";", header = TRUE, stringsAsFactors=F

# Excluding some columns from hofstede data
working_data <- na.omit(data)
# Excluding some columns from trompenaars data
working_dataT <- na.omit(dataT)
# Excluding some columns from globe data
working_dataG <- na.omit(dataG)
```

Linear Mixed Model using lmer function on all the variables for Hofstede

```
#----------------------------
#ALL THE VARIABLES

# Applying a Linear Mixed Model using the lmer function
organisation <- lmer(working_data$organisation~log(working_data$totalCommitters)+log(working_data$totalc
             +working_data$projectAge+working_data$turnover+working_data$blauGender
             +working_data$tenureMedian+working_data$tenureDiversity+log(working_data$teamSize)
             +working_data$stCongruence+working_data$truckFactor+working_data$female
             +working_data$expertise+working_data$centrality+working_data$CV_1
             +working_data$CV_2+working_data$CV_3
             +working_data$CV_4+working_data$CV_5+working_data$CV_6
             +(1 | working_data$window_idx ), REML=FALSE)
```

```
## boundary (singular) fit: see help('isSingular')
```

```
# Remove outlier
#romr.fnc(black, working_data, trim = 2.5)

# Applying vif <5
print(vif(organisation))
```

```
## log(working_data$totalCommitters)    log(working_data$totalcommits)
##                             3.331866                             3.149351
##            working_data$projectAge              working_data$turnover
##                             1.445216                             1.422047
##            working_data$blauGender           working_data$tenureMedian
##                             2.689824                             1.113645
##       working_data$tenureDiversity          log(working_data$teamSize)
##                             1.063933                             2.659345
##          working_data$stCongruence            working_data$truckFactor
##                             1.063036                             1.091875
##                working_data$female              working_data$expertise
##                             1.074503                             1.104687
##            working_data$centrality                   working_data$CV_1
##                             1.162484                             4.984718
##                   working_data$CV_2                   working_data$CV_3
##                             6.259349                             3.302360
##                   working_data$CV_4                   working_data$CV_5
##                             8.488179                             4.322881
##                   working_data$CV_6
##                             7.499225
```

```
# Applying a Linear Mixed Model using the lmer function, after vif - NO REMOVAL

# print result
print(summary(organisation))
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: working_data$organisation ~ log(working_data$totalCommitters) +
##     log(working_data$totalcommits) + working_data$projectAge +
##     working_data$turnover + working_data$blauGender + working_data$tenureMedian +
##     working_data$tenureDiversity + log(working_data$teamSize) +
```

```
##     working_data$stCongruence + working_data$truckFactor + working_data$female +
##     working_data$expertise + working_data$centrality + working_data$CV_1 +
##     working_data$CV_2 + working_data$CV_3 + working_data$CV_4 +
##     working_data$CV_5 + working_data$CV_6 + (1 | working_data$window_idx)
##
##      AIC      BIC   logLik deviance df.resid
##    765.6    837.3   -360.8    721.6      170
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.93475 -0.75308 -0.00831  0.70689  2.06126
##
## Random effects:
##  Groups                   Name        Variance Std.Dev.
##  working_data$window_idx (Intercept) 0.00     0.000
##  Residual                             2.51     1.584
## Number of obs: 192, groups:  working_data$window_idx, 24
##
## Fixed effects:
##                                  Estimate Std. Error t value
## (Intercept)                      1.395977   1.541676   0.905
## log(working_data$totalCommitters)  0.332618   0.177967   1.869
## log(working_data$totalcommits)    -0.101356   0.134933  -0.751
## working_data$projectAge           -0.027446   0.028382  -0.967
## working_data$turnover             -1.207661   0.581334  -2.077
## working_data$blauGender            1.437807   2.040480   0.705
## working_data$tenureMedian         -0.099330   0.066761  -1.488
## working_data$tenureDiversity       0.021192   0.046804   0.453
## log(working_data$teamSize)         0.354546   0.176055   2.014
## working_data$stCongruence          0.373505   0.346870   1.077
## working_data$truckFactor           0.039678   0.087027   0.456
## working_data$female                0.042810   0.020935   2.045
## working_data$expertise            -0.392552   0.381443  -1.029
## working_data$centrality            0.006655   0.276379   0.024
## working_data$CV_1                  1.033520   2.511175   0.412
## working_data$CV_2                  0.573831   3.040383   0.189
## working_data$CV_3                 -1.959870   2.229393  -0.879
## working_data$CV_4                 -4.062589   3.997663  -1.016
## working_data$CV_5                  1.236888   1.950336   0.634
## working_data$CV_6                 -1.903487   3.351464  -0.568
##
## Correlation matrix not shown by default, as p = 20 > 12.
## Use print(summary(organisation), correlation=TRUE)  or
##     vcov(summary(organisation))        if you need it
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```r
# Applying anova
Anova(organisation)
```

```
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: working_data$organisation
##                                  Chisq Df Pr(>Chisq)
```

```
## log(working_data$totalCommitters) 3.4931  1    0.06162 .
## log(working_data$totalcommits)    0.5642  1    0.45256
## working_data$projectAge           0.9351  1    0.33353
## working_data$turnover             4.3156  1    0.03777 *
## working_data$blauGender           0.4965  1    0.48103
## working_data$tenureMedian         2.2137  1    0.13679
## working_data$tenureDiversity      0.2050  1    0.65071
## log(working_data$teamSize)        4.0555  1    0.04403 *
## working_data$stCongruence         1.1595  1    0.28158
## working_data$truckFactor          0.2079  1    0.64844
## working_data$female               4.1816  1    0.04087 *
## working_data$expertise            1.0591  1    0.30342
## working_data$centrality           0.0006  1    0.98079
## working_data$CV_1                 0.1694  1    0.68066
## working_data$CV_2                 0.0356  1    0.85030
## working_data$CV_3                 0.7728  1    0.37934
## working_data$CV_4                 1.0327  1    0.30951
## working_data$CV_5                 0.4022  1    0.52596
## working_data$CV_6                 0.3226  1    0.57006
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Save in a txt file
sink("hofstede/output_organisation_hofstede_all_variables.txt")
print(summary(organisation))
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: working_data$organisation ~ log(working_data$totalCommitters) +
##     log(working_data$totalcommits) + working_data$projectAge +
##     working_data$turnover + working_data$blauGender + working_data$tenureMedian +
##     working_data$tenureDiversity + log(working_data$teamSize) +
##     working_data$stCongruence + working_data$truckFactor + working_data$female +
##     working_data$expertise + working_data$centrality + working_data$CV_1 +
##     working_data$CV_2 + working_data$CV_3 + working_data$CV_4 +
##     working_data$CV_5 + working_data$CV_6 + (1 | working_data$window_idx)
##
##      AIC      BIC   logLik deviance df.resid
##    765.6    837.3   -360.8    721.6      170
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.93475 -0.75308 -0.00831  0.70689  2.06126
##
## Random effects:
##  Groups                   Name        Variance Std.Dev.
##  working_data$window_idx (Intercept) 0.00     0.000
##  Residual                             2.51     1.584
## Number of obs: 192, groups:  working_data$window_idx, 24
##
## Fixed effects:
##                                  Estimate Std. Error t value
## (Intercept)                      1.395977   1.541676   0.905
## log(working_data$totalCommitters)  0.332618   0.177967   1.869
## log(working_data$totalcommits)    -0.101356   0.134933  -0.751
## working_data$projectAge           -0.027446   0.028382  -0.967
```

```
## working_data$turnover             -1.207661   0.581334  -2.077
## working_data$blauGender            1.437807   2.040480   0.705
## working_data$tenureMedian         -0.099330   0.066761  -1.488
## working_data$tenureDiversity       0.021192   0.046804   0.453
## log(working_data$teamSize)         0.354546   0.176055   2.014
## working_data$stCongruence          0.373505   0.346870   1.077
## working_data$truckFactor           0.039678   0.087027   0.456
## working_data$female                0.042810   0.020935   2.045
## working_data$expertise            -0.392552   0.381443  -1.029
## working_data$centrality            0.006655   0.276379   0.024
## working_data$CV_1                  1.033520   2.511175   0.412
## working_data$CV_2                  0.573831   3.040383   0.189
## working_data$CV_3                 -1.959870   2.229393  -0.879
## working_data$CV_4                 -4.062589   3.997663  -1.016
## working_data$CV_5                  1.236888   1.950336   0.634
## working_data$CV_6                 -1.903487   3.351464  -0.568

##
## Correlation matrix not shown by default, as p = 20 > 12.
## Use print(summary(organisation), correlation=TRUE)  or
##     vcov(summary(organisation))        if you need it

## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```
Anova(organisation)
```

```
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: working_data$organisation
##                                  Chisq Df Pr(>Chisq)
## log(working_data$totalCommitters) 3.4931  1    0.06162 .
## log(working_data$totalcommits)    0.5642  1    0.45256
## working_data$projectAge           0.9351  1    0.33353
## working_data$turnover             4.3156  1    0.03777 *
## working_data$blauGender           0.4965  1    0.48103
## working_data$tenureMedian         2.2137  1    0.13679
## working_data$tenureDiversity      0.2050  1    0.65071
## log(working_data$teamSize)        4.0555  1    0.04403 *
## working_data$stCongruence         1.1595  1    0.28158
## working_data$truckFactor          0.2079  1    0.64844
## working_data$female               4.1816  1    0.04087 *
## working_data$expertise            1.0591  1    0.30342
## working_data$centrality           0.0006  1    0.98079
## working_data$CV_1                 0.1694  1    0.68066
## working_data$CV_2                 0.0356  1    0.85030
## working_data$CV_3                 0.7728  1    0.37934
## working_data$CV_4                 1.0327  1    0.30951
## working_data$CV_5                 0.4022  1    0.52596
## working_data$CV_6                 0.3226  1    0.57006
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
sink()
```

Linear Mixed Model using lmer function on all the variables for trompenaars

```
#----------------------------
#ALL THE VARIABLES

# Applying a Linear Mixed Model using the lmer function
organisationT <- lmer(working_dataT$organisation~log(working_data$totalCommitters)+log(working_dataT$to
             +working_data$projectAge+working_dataT$turnover+working_dataT$blauGender
             +working_dataT$tenureMedian+working_dataT$tenureDiversity+log(working_dataT$teamSize)
             +working_dataT$stCongruence+working_dataT$truckFactor+working_dataT$female
             +working_dataT$expertise+working_dataT$centrality+working_dataT$CV_1
             +working_dataT$CV_2+working_dataT$CV_3
             +working_dataT$CV_4+working_dataT$CV_5+working_dataT$CV_6
             +working_dataT$CV_7+working_dataT$CV_8
             +(1 | working_dataT$window_idx ), REML=FALSE)
```

```
## boundary (singular) fit: see help('isSingular')
```

```
# Remove outlier
#romr.fnc(blackT, working_dataT, trim = 2.5)

# Applying vif <5
print(vif(organisationT))
```

```
## log(working_data$totalCommitters)   log(working_dataT$totalcommits)
##                          3.093368                          3.120539
##           working_data$projectAge           working_dataT$turnover
##                          1.494272                          1.485956
##          working_dataT$blauGender       working_dataT$tenureMedian
##                          2.576141                          1.095823
##     working_dataT$tenureDiversity      log(working_dataT$teamSize)
##                          1.083495                          2.332263
##        working_dataT$stCongruence        working_dataT$truckFactor
##                          1.053654                          1.090676
##              working_dataT$female          working_dataT$expertise
##                          1.106136                          1.156034
##          working_dataT$centrality               working_dataT$CV_1
##                          1.168007                         14.859392
##                working_dataT$CV_2               working_dataT$CV_3
##                         11.823005                          6.325515
##                working_dataT$CV_4               working_dataT$CV_5
##                         11.188490                         21.379993
##                working_dataT$CV_6               working_dataT$CV_7
##                          3.040881                          6.905493
##                working_dataT$CV_8
##                          4.444750
```

```
# Applying a Linear Mixed Model using the lmer function, after vif - NO REMOVAL

# print result
```

```r
print(summary(organisationT))
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: working_dataT$organisation ~ log(working_data$totalCommitters) +
##     log(working_dataT$totalcommits) + working_data$projectAge +
##     working_dataT$turnover + working_dataT$blauGender + working_dataT$tenureMedian +
##     working_dataT$tenureDiversity + log(working_dataT$teamSize) +
##     working_dataT$stCongruence + working_dataT$truckFactor +
##     working_dataT$female + working_dataT$expertise + working_dataT$centrality +
##     working_dataT$CV_1 + working_dataT$CV_2 + working_dataT$CV_3 +
##     working_dataT$CV_4 + working_dataT$CV_5 + working_dataT$CV_6 +
##     working_dataT$CV_7 + working_dataT$CV_8 + (1 | working_dataT$window_idx)
##
##      AIC      BIC   logLik deviance df.resid
##    768.2    846.4   -360.1    720.2      168
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.03099 -0.76384 -0.06739  0.75043  2.06033
##
## Random effects:
##  Groups                    Name        Variance Std.Dev.
##  working_dataT$window_idx (Intercept) 0.000    0.000
##  Residual                             2.493    1.579
## Number of obs: 192, groups:  working_dataT$window_idx, 24
##
## Fixed effects:
##                                  Estimate Std. Error t value
## (Intercept)                       1.06933    1.53801   0.695
## log(working_data$totalCommitters)  0.34392    0.17087   2.013
## log(working_dataT$totalcommits)   -0.07603    0.13384  -0.568
## working_data$projectAge           -0.01442    0.02876  -0.501
## working_dataT$turnover            -1.30050    0.59215  -2.196
## working_dataT$blauGender           1.70782    1.98984   0.858
## working_dataT$tenureMedian        -0.10970    0.06599  -1.662
## working_dataT$tenureDiversity      0.01354    0.04707   0.288
## log(working_dataT$teamSize)        0.23340    0.16429   1.421
## working_dataT$stCongruence         0.35592    0.34412   1.034
## working_dataT$truckFactor          0.03276    0.08667   0.378
## working_dataT$female               0.03722    0.02117   1.758
## working_dataT$expertise           -0.28610    0.38883  -0.736
## working_dataT$centrality           0.06381    0.27606   0.231
## working_dataT$CV_1                -4.75624    4.15030  -1.146
## working_dataT$CV_2                 1.52702    3.04888   0.501
## working_dataT$CV_3                 0.74848    2.49283   0.300
## working_dataT$CV_4                -5.83161    3.30305  -1.766
## working_dataT$CV_5                 3.72863    4.04585   0.922
## working_dataT$CV_6                 1.19236    1.08454   1.099
## working_dataT$CV_7                 3.19365    2.80149   1.140
## working_dataT$CV_8                -2.89280    2.04903  -1.412
##
## Correlation matrix not shown by default, as p = 22 > 12.
## Use print(summary(organisationT), correlation=TRUE)  or
```

```
##     vcov(summary(organisationT))        if you need it
```

```
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```
# Applying anova
Anova(organisationT)
```

```
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: working_dataT$organisation
##                                    Chisq Df Pr(>Chisq)
## log(working_data$totalCommitters) 4.0511  1    0.04414 *
## log(working_dataT$totalcommits)   0.3227  1    0.57000
## working_data$projectAge           0.2514  1    0.61611
## working_dataT$turnover            4.8234  1    0.02808 *
## working_dataT$blauGender          0.7366  1    0.39074
## working_dataT$tenureMedian        2.7634  1    0.09644 .
## working_dataT$tenureDiversity     0.0827  1    0.77363
## log(working_dataT$teamSize)       2.0182  1    0.15542
## working_dataT$stCongruence        1.0698  1    0.30100
## working_dataT$truckFactor         0.1428  1    0.70548
## working_dataT$female              3.0922  1    0.07867 .
## working_dataT$expertise           0.5414  1    0.46185
## working_dataT$centrality          0.0534  1    0.81720
## working_dataT$CV_1                1.3133  1    0.25180
## working_dataT$CV_2                0.2508  1    0.61648
## working_dataT$CV_3                0.0902  1    0.76398
## working_dataT$CV_4                3.1171  1    0.07748 .
## working_dataT$CV_5                0.8493  1    0.35674
## working_dataT$CV_6                1.2087  1    0.27158
## working_dataT$CV_7                1.2996  1    0.25429
## working_dataT$CV_8                1.9931  1    0.15801
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Save in a txt file
sink("trompe/output_organisation_trompenaars_all_variables.txt")
print(summary(organisationT))
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: working_dataT$organisation ~ log(working_data$totalCommitters) +
##     log(working_dataT$totalcommits) + working_data$projectAge +
##     working_dataT$turnover + working_dataT$blauGender + working_dataT$tenureMedian +
##     working_dataT$tenureDiversity + log(working_dataT$teamSize) +
##     working_dataT$stCongruence + working_dataT$truckFactor +
##     working_dataT$female + working_dataT$expertise + working_dataT$centrality +
##     working_dataT$CV_1 + working_dataT$CV_2 + working_dataT$CV_3 +
##     working_dataT$CV_4 + working_dataT$CV_5 + working_dataT$CV_6 +
##     working_dataT$CV_7 + working_dataT$CV_8 + (1 | working_dataT$window_idx)
##
##      AIC      BIC   logLik deviance df.resid
##    768.2    846.4   -360.1    720.2      168
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
```

```
## -2.03099 -0.76384 -0.06739  0.75043  2.06033
##
## Random effects:
##  Groups                    Name        Variance Std.Dev.
##  working_dataT$window_idx (Intercept) 0.000    0.000
##  Residual                             2.493    1.579
## Number of obs: 192, groups:  working_dataT$window_idx, 24
##
## Fixed effects:
##                                   Estimate Std. Error t value
## (Intercept)                        1.06933    1.53801   0.695
## log(working_data$totalCommitters)  0.34392    0.17087   2.013
## log(working_dataT$totalcommits)   -0.07603    0.13384  -0.568
## working_data$projectAge           -0.01442    0.02876  -0.501
## working_dataT$turnover            -1.30050    0.59215  -2.196
## working_dataT$blauGender           1.70782    1.98984   0.858
## working_dataT$tenureMedian        -0.10970    0.06599  -1.662
## working_dataT$tenureDiversity      0.01354    0.04707   0.288
## log(working_dataT$teamSize)        0.23340    0.16429   1.421
## working_dataT$stCongruence         0.35592    0.34412   1.034
## working_dataT$truckFactor          0.03276    0.08667   0.378
## working_dataT$female               0.03722    0.02117   1.758
## working_dataT$expertise           -0.28610    0.38883  -0.736
## working_dataT$centrality           0.06381    0.27606   0.231
## working_dataT$CV_1                -4.75624    4.15030  -1.146
## working_dataT$CV_2                 1.52702    3.04888   0.501
## working_dataT$CV_3                 0.74848    2.49283   0.300
## working_dataT$CV_4                -5.83161    3.30305  -1.766
## working_dataT$CV_5                 3.72863    4.04585   0.922
## working_dataT$CV_6                 1.19236    1.08454   1.099
## working_dataT$CV_7                 3.19365    2.80149   1.140
## working_dataT$CV_8                -2.89280    2.04903  -1.412
##
## Correlation matrix not shown by default, as p = 22 > 12.
## Use print(summary(organisationT), correlation=TRUE)  or
##     vcov(summary(organisationT))        if you need it
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```
Anova(organisationT)
```

```
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: working_dataT$organisation
##                                    Chisq Df Pr(>Chisq)
## log(working_data$totalCommitters) 4.0511  1    0.04414 *
## log(working_dataT$totalcommits)   0.3227  1    0.57000
## working_data$projectAge           0.2514  1    0.61611
## working_dataT$turnover            4.8234  1    0.02808 *
## working_dataT$blauGender          0.7366  1    0.39074
## working_dataT$tenureMedian        2.7634  1    0.09644 .
## working_dataT$tenureDiversity     0.0827  1    0.77363
## log(working_dataT$teamSize)       2.0182  1    0.15542
## working_dataT$stCongruence        1.0698  1    0.30100
```

```
## working_dataT$truckFactor         0.1428  1    0.70548
## working_dataT$female              3.0922  1    0.07867 .
## working_dataT$expertise           0.5414  1    0.46185
## working_dataT$centrality          0.0534  1    0.81720
## working_dataT$CV_1                1.3133  1    0.25180
## working_dataT$CV_2                0.2508  1    0.61648
## working_dataT$CV_3                0.0902  1    0.76398
## working_dataT$CV_4                3.1171  1    0.07748 .
## working_dataT$CV_5                0.8493  1    0.35674
## working_dataT$CV_6                1.2087  1    0.27158
## working_dataT$CV_7                1.2996  1    0.25429
## working_dataT$CV_8                1.9931  1    0.15801
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
sink()
```

Linear Mixed Model using lmer function on all the variables for Globe

```
#-----------------------------
#ALL THE VARIABLES

# Applying a Linear Mixed Model using the lmer function
organisationG <- lmer(working_dataG$organisation~log(working_dataG$totalCommitters)+log(working_dataG$t
             +working_dataG$projectAge+working_dataG$turnover+working_dataG$blauGender
             +working_dataG$tenureMedian+working_dataG$tenureDiversity+log(working_dataG$teamSize)
             +working_dataG$stCongruence+working_dataG$truckFactor+working_dataG$female
             +working_dataG$expertise+working_dataG$centrality+working_dataG$CV_1
             +working_dataG$CV_2+working_dataG$CV_3
             +working_dataG$CV_4+working_dataG$CV_5+working_dataG$CV_6+working_dataG$CV_7
             +working_dataG$CV_8+working_dataG$CV_9
             +(1 | working_dataG$window_idx ), REML=FALSE)
```

```
## boundary (singular) fit: see help('isSingular')
```

```
# Remove outlier
#romr.fnc(blackG, working_dataG, trim = 2.5)

# Applying vif <5
print(vif(organisationG))
```

```
## log(working_dataG$totalCommitters)    log(working_dataG$totalcommits)
##                           3.394274                           3.448057
##           working_dataG$projectAge            working_dataG$turnover
##                           1.463089                           1.589672
##           working_dataG$blauGender        working_dataG$tenureMedian
##                           2.555189                           1.113615
##      working_dataG$tenureDiversity       log(working_dataG$teamSize)
##                           1.086646                           2.528986
##         working_dataG$stCongruence         working_dataG$truckFactor
##                           1.050260                           1.115467
##               working_dataG$female           working_dataG$expertise
##                           1.067570                           1.203557
```

```
##           working_dataG$centrality                 working_dataG$CV_1
##                           1.153879                          19.030437
##                 working_dataG$CV_2                 working_dataG$CV_3
##                           7.471510                           5.770990
##                 working_dataG$CV_4                 working_dataG$CV_5
##                          10.426962                           5.160977
##                 working_dataG$CV_6                 working_dataG$CV_7
##                          15.120440                           6.088040
##                 working_dataG$CV_8                 working_dataG$CV_9
##                          17.379809                           8.348261
```

```r
# Applying a Linear Mixed Model using the lmer function, after vif - NO REMOVAL

# print result
print(summary(organisationG))
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: working_dataG$organisation ~ log(working_dataG$totalCommitters) +
##     log(working_dataG$totalcommits) + working_dataG$projectAge +
##     working_dataG$turnover + working_dataG$blauGender + working_dataG$tenureMedian +
##     working_dataG$tenureDiversity + log(working_dataG$teamSize) +
##     working_dataG$stCongruence + working_dataG$truckFactor +
##     working_dataG$female + working_dataG$expertise + working_dataG$centrality +
##     working_dataG$CV_1 + working_dataG$CV_2 + working_dataG$CV_3 +
##     working_dataG$CV_4 + working_dataG$CV_5 + working_dataG$CV_6 +
##     working_dataG$CV_7 + working_dataG$CV_8 + working_dataG$CV_9 +
##     (1 | working_dataG$window_idx)
##
##      AIC      BIC   logLik deviance df.resid
##    769.1    850.5   -359.5    719.1      167
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.92619 -0.78614 -0.06918  0.76229  2.16341
##
## Random effects:
##  Groups                    Name        Variance Std.Dev.
##  working_dataG$window_idx (Intercept) 0.000    0.000
##  Residual                             2.478    1.574
## Number of obs: 192, groups:  working_dataG$window_idx, 24
##
## Fixed effects:
##                                     Estimate Std. Error t value
## (Intercept)                         2.545158   1.491705   1.706
## log(working_dataG$totalCommitters)  0.420619   0.178454   2.357
## log(working_dataG$totalcommits)    -0.176115   0.140266  -1.256
## working_dataG$projectAge           -0.030251   0.028370  -1.066
## working_dataG$turnover             -1.639682   0.610633  -2.685
## working_dataG$blauGender            1.328480   1.975786   0.672
## working_dataG$tenureMedian         -0.125299   0.066325  -1.889
## working_dataG$tenureDiversity       0.001283   0.046992   0.027
## log(working_dataG$teamSize)         0.290309   0.170566   1.702
## working_dataG$stCongruence          0.334951   0.342530   0.978
## working_dataG$truckFactor           0.034951   0.087388   0.400
## working_dataG$female                0.039667   0.020731   1.913
```

```
## working_dataG$expertise             -0.344030   0.395550  -0.870
## working_dataG$centrality             0.028069   0.273559   0.103
## working_dataG$CV_1                  -7.507475  14.071473  -0.534
## working_dataG$CV_2                  -9.952546  10.619782  -0.937
## working_dataG$CV_3                  -8.910067  12.958495  -0.688
## working_dataG$CV_4                  15.770592  14.494298   1.088
## working_dataG$CV_5                  -2.693425  10.854846  -0.248
## working_dataG$CV_6                  12.220115  22.018531   0.555
## working_dataG$CV_7                  -1.534864   7.344964  -0.209
## working_dataG$CV_8                  12.259021  20.763052   0.590
## working_dataG$CV_9                 -25.090925  15.491484  -1.620

##
## Correlation matrix not shown by default, as p = 23 > 12.
## Use print(summary(organisationG), correlation=TRUE)  or
##     vcov(summary(organisationG))        if you need it

## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```
# Applying anova
Anova(organisationG)
```

```
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: working_dataG$organisation
##                                 Chisq Df Pr(>Chisq)
## log(working_dataG$totalCommitters) 5.5555  1   0.018423 *
## log(working_dataG$totalcommits)    1.5765  1   0.209270
## working_dataG$projectAge           1.1369  1   0.286301
## working_dataG$turnover             7.2104  1   0.007248 **
## working_dataG$blauGender           0.4521  1   0.501342
## working_dataG$tenureMedian         3.5690  1   0.058868 .
## working_dataG$tenureDiversity      0.0007  1   0.978220
## log(working_dataG$teamSize)        2.8969  1   0.088750 .
## working_dataG$stCongruence         0.9562  1   0.328137
## working_dataG$truckFactor          0.1600  1   0.689191
## working_dataG$female               3.6611  1   0.055697 .
## working_dataG$expertise            0.7565  1   0.384436
## working_dataG$centrality           0.0105  1   0.918275
## working_dataG$CV_1                 0.2846  1   0.593671
## working_dataG$CV_2                 0.8783  1   0.348671
## working_dataG$CV_3                 0.4728  1   0.491714
## working_dataG$CV_4                 1.1839  1   0.276571
## working_dataG$CV_5                 0.0616  1   0.804033
## working_dataG$CV_6                 0.3080  1   0.578900
## working_dataG$CV_7                 0.0437  1   0.834473
## working_dataG$CV_8                 0.3486  1   0.554906
## working_dataG$CV_9                 2.6233  1   0.105306
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Save in a txt file
sink("globe/output_organisation_globe_all_variables.txt")
print(summary(organisationG))
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
```

```
## Formula: working_dataG$organisation ~ log(working_dataG$totalCommitters) +
##     log(working_dataG$totalcommits) + working_dataG$projectAge +
##     working_dataG$turnover + working_dataG$blauGender + working_dataG$tenureMedian +
##     working_dataG$tenureDiversity + log(working_dataG$teamSize) +
##     working_dataG$stCongruence + working_dataG$truckFactor +
##     working_dataG$female + working_dataG$expertise + working_dataG$centrality +
##     working_dataG$CV_1 + working_dataG$CV_2 + working_dataG$CV_3 +
##     working_dataG$CV_4 + working_dataG$CV_5 + working_dataG$CV_6 +
##     working_dataG$CV_7 + working_dataG$CV_8 + working_dataG$CV_9 +
##     (1 | working_dataG$window_idx)
##
##      AIC      BIC   logLik deviance df.resid
##    769.1    850.5   -359.5    719.1      167
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.92619 -0.78614 -0.06918  0.76229  2.16341
##
## Random effects:
##  Groups                   Name        Variance Std.Dev.
##  working_dataG$window_idx (Intercept) 0.000    0.000
##  Residual                             2.478    1.574
## Number of obs: 192, groups:  working_dataG$window_idx, 24
##
## Fixed effects:
##                                    Estimate Std. Error t value
## (Intercept)                        2.545158   1.491705   1.706
## log(working_dataG$totalCommitters) 0.420619   0.178454   2.357
## log(working_dataG$totalcommits)    -0.176115   0.140266  -1.256
## working_dataG$projectAge           -0.030251   0.028370  -1.066
## working_dataG$turnover             -1.639682   0.610633  -2.685
## working_dataG$blauGender            1.328480   1.975786   0.672
## working_dataG$tenureMedian         -0.125299   0.066325  -1.889
## working_dataG$tenureDiversity       0.001283   0.046992   0.027
## log(working_dataG$teamSize)         0.290309   0.170566   1.702
## working_dataG$stCongruence          0.334951   0.342530   0.978
## working_dataG$truckFactor           0.034951   0.087388   0.400
## working_dataG$female                0.039667   0.020731   1.913
## working_dataG$expertise            -0.344030   0.395550  -0.870
## working_dataG$centrality            0.028069   0.273559   0.103
## working_dataG$CV_1                 -7.507475  14.071473  -0.534
## working_dataG$CV_2                 -9.952546  10.619782  -0.937
## working_dataG$CV_3                 -8.910067  12.958495  -0.688
## working_dataG$CV_4                 15.770592  14.494298   1.088
## working_dataG$CV_5                 -2.693425  10.854846  -0.248
## working_dataG$CV_6                 12.220115  22.018531   0.555
## working_dataG$CV_7                 -1.534864   7.344964  -0.209
## working_dataG$CV_8                 12.259021  20.763052   0.590
## working_dataG$CV_9                -25.090925  15.491484  -1.620
##
## Correlation matrix not shown by default, as p = 23 > 12.
## Use print(summary(organisationG), correlation=TRUE)  or
##     vcov(summary(organisationG))        if you need it
```

```
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```
Anova(organisationG)
```

```
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: working_dataG$organisation
##                                    Chisq Df Pr(>Chisq)
## log(working_dataG$totalCommitters) 5.5555  1   0.018423 *
## log(working_dataG$totalcommits)    1.5765  1   0.209270
## working_dataG$projectAge           1.1369  1   0.286301
## working_dataG$turnover             7.2104  1   0.007248 **
## working_dataG$blauGender           0.4521  1   0.501342
## working_dataG$tenureMedian         3.5690  1   0.058868 .
## working_dataG$tenureDiversity      0.0007  1   0.978220
## log(working_dataG$teamSize)        2.8969  1   0.088750 .
## working_dataG$stCongruence         0.9562  1   0.328137
## working_dataG$truckFactor          0.1600  1   0.689191
## working_dataG$female               3.6611  1   0.055697 .
## working_dataG$expertise            0.7565  1   0.384436
## working_dataG$centrality           0.0105  1   0.918275
## working_dataG$CV_1                 0.2846  1   0.593671
## working_dataG$CV_2                 0.8783  1   0.348671
## working_dataG$CV_3                 0.4728  1   0.491714
## working_dataG$CV_4                 1.1839  1   0.276571
## working_dataG$CV_5                 0.0616  1   0.804033
## working_dataG$CV_6                 0.3080  1   0.578900
## working_dataG$CV_7                 0.0437  1   0.834473
## working_dataG$CV_8                 0.3486  1   0.554906
## working_dataG$CV_9                 2.6233  1   0.105306
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
sink()
```

Linear Mixed Model using lmer function on all the confounding variables hofstede

```
#-----------------------------
#ALL THE CONFOUNDING VARIABLES

# Applying a Linear Mixed Model using the lmer function
organisation <- lmer(working_data$organisation~log(working_data$totalCommitters)+log(working_data$totalc
             +working_data$projectAge+working_data$turnover
             +working_data$tenureMedian+working_data$tenureDiversity+log(working_data$teamSize)
             +working_data$stCongruence+working_data$centrality+working_data$truckFactor
             +working_data$expertise+working_data$female+working_data$blauGender
             +(1 | working_data$window_idx ), REML=FALSE)
```

```
## boundary (singular) fit: see help('isSingular')
```

```
# Remove outlier
#romr.fnc(organisation, working_data, trim = 2.5)
```

```r
# Applying vif <5
print(vif(organisation))
```

```
## log(working_data$totalCommitters)    log(working_data$totalcommits)
##                          2.592895                          2.336089
##           working_data$projectAge             working_data$turnover
##                          1.341070                          1.248622
##         working_data$tenureMedian      working_data$tenureDiversity
##                          1.065323                          1.041936
##        log(working_data$teamSize)         working_data$stCongruence
##                          1.833618                          1.033152
##           working_data$centrality          working_data$truckFactor
##                          1.107221                          1.068792
##            working_data$expertise               working_data$female
##                          1.078462                          1.040100
##           working_data$blauGender
##                          2.022374
```

```r
# Applying a Linear Mixed Model using the lmer function, after vif, NO REMOVALS
organisation <- lmer(working_data$organisation~log(working_data$totalCommitters)+log(working_data$totalc
            +working_data$projectAge+working_data$turnover
            +working_data$tenureMedian+working_data$centrality+working_data$tenureDiversity
            +working_data$stCongruence+working_data$truckFactor
            +working_data$expertise+working_data$female+working_data$blauGender
            +(1 | working_data$window_idx ), REML=FALSE)
```

```
## boundary (singular) fit: see help('isSingular')
```

```r
# print result
print(summary(organisation))
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: working_data$organisation ~ log(working_data$totalCommitters) +
##     log(working_data$totalcommits) + working_data$projectAge +
##     working_data$turnover + working_data$tenureMedian + working_data$centrality +
##     working_data$tenureDiversity + working_data$stCongruence +
##     working_data$truckFactor + working_data$expertise + working_data$female +
##     working_data$blauGender + (1 | working_data$window_idx)
##
##      AIC      BIC   logLik deviance df.resid
##    758.9    807.8   -364.4    728.9      177
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.93131 -0.80195 -0.06154  0.81380  2.04131
##
## Random effects:
##  Groups                   Name        Variance Std.Dev.
##  working_data$window_idx (Intercept) 0.000    0.000
##  Residual                             2.608    1.615
## Number of obs: 192, groups:  working_data$window_idx, 24
##
## Fixed effects:
##                                 Estimate Std. Error t value
## (Intercept)                      1.76031    1.18296   1.488
```

```
## log(working_data$totalCommitters)  0.40372    0.15834   2.550
## log(working_data$totalcommits)    -0.14431    0.11819  -1.221
## working_data$projectAge           -0.02443    0.02768  -0.883
## working_data$turnover             -1.38135    0.55466  -2.490
## working_data$tenureMedian         -0.09215    0.06655  -1.385
## working_data$centrality            0.02872    0.27424   0.105
## working_data$tenureDiversity       0.01451    0.04720   0.308
## working_data$stCongruence          0.38007    0.34847   1.091
## working_data$truckFactor           0.04206    0.08775   0.479
## working_data$expertise            -0.37246    0.38274  -0.973
## working_data$female                0.03992    0.02095   1.906
## working_data$blauGender            1.94591    1.61696   1.203
```

```
##
## Correlation matrix not shown by default, as p = 13 > 12.
## Use print(summary(organisation), correlation=TRUE)  or
##     vcov(summary(organisation))        if you need it
```

```
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```
# Applying anova
Anova(organisation)
```

```
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: working_data$organisation
##                                   Chisq Df Pr(>Chisq)
## log(working_data$totalCommitters) 6.5014  1    0.01078 *
## log(working_data$totalcommits)    1.4908  1    0.22210
## working_data$projectAge           0.7789  1    0.37747
## working_data$turnover             6.2023  1    0.01276 *
## working_data$tenureMedian         1.9174  1    0.16614
## working_data$centrality           0.0110  1    0.91660
## working_data$tenureDiversity      0.0946  1    0.75846
## working_data$stCongruence         1.1896  1    0.27542
## working_data$truckFactor          0.2298  1    0.63167
## working_data$expertise            0.9470  1    0.33049
## working_data$female               3.6331  1    0.05664 .
## working_data$blauGender           1.4482  1    0.22881
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Save in a txt file
sink("hofstede/output_organisation_hofstede_confounding_variables.txt")
print(summary(organisation))
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: working_data$organisation ~ log(working_data$totalCommitters) +
##     log(working_data$totalcommits) + working_data$projectAge +
##     working_data$turnover + working_data$tenureMedian + working_data$centrality +
##     working_data$tenureDiversity + working_data$stCongruence +
##     working_data$truckFactor + working_data$expertise + working_data$female +
##     working_data$blauGender + (1 | working_data$window_idx)
##
##      AIC      BIC   logLik deviance df.resid
##    758.9    807.8   -364.4    728.9      177
```

```
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.93131 -0.80195 -0.06154  0.81380  2.04131
##
## Random effects:
##  Groups                   Name        Variance Std.Dev.
##  working_data$window_idx (Intercept) 0.000    0.000
##  Residual                             2.608    1.615
## Number of obs: 192, groups:  working_data$window_idx, 24
##
## Fixed effects:
##                                  Estimate Std. Error t value
## (Intercept)                       1.76031    1.18296   1.488
## log(working_data$totalCommitters)  0.40372    0.15834   2.550
## log(working_data$totalcommits)    -0.14431    0.11819  -1.221
## working_data$projectAge           -0.02443    0.02768  -0.883
## working_data$turnover             -1.38135    0.55466  -2.490
## working_data$tenureMedian         -0.09215    0.06655  -1.385
## working_data$centrality            0.02872    0.27424   0.105
## working_data$tenureDiversity       0.01451    0.04720   0.308
## working_data$stCongruence          0.38007    0.34847   1.091
## working_data$truckFactor           0.04206    0.08775   0.479
## working_data$expertise            -0.37246    0.38274  -0.973
## working_data$female                0.03992    0.02095   1.906
## working_data$blauGender            1.94591    1.61696   1.203
##
## Correlation matrix not shown by default, as p = 13 > 12.
## Use print(summary(organisation), correlation=TRUE)  or
##     vcov(summary(organisation))        if you need it
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```
Anova(organisation)
```

```
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: working_data$organisation
##                                   Chisq Df Pr(>Chisq)
## log(working_data$totalCommitters) 6.5014  1    0.01078 *
## log(working_data$totalcommits)    1.4908  1    0.22210
## working_data$projectAge           0.7789  1    0.37747
## working_data$turnover             6.2023  1    0.01276 *
## working_data$tenureMedian         1.9174  1    0.16614
## working_data$centrality           0.0110  1    0.91660
## working_data$tenureDiversity      0.0946  1    0.75846
## working_data$stCongruence         1.1896  1    0.27542
## working_data$truckFactor          0.2298  1    0.63167
## working_data$expertise            0.9470  1    0.33049
## working_data$female               3.6331  1    0.05664 .
## working_data$blauGender           1.4482  1    0.22881
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
sink()
```

Linear Mixed Model using lmer function on all the confounding variables trompenaars

```
#----------------------------
#ALL THE CONFOUNDING VARIABLES

# Applying a Linear Mixed Model using the lmer function
organisationT <- lmer(working_dataT$organisation~log(working_dataT$totalCommitters)+log(working_dataT$t
             +working_dataT$projectAge+working_dataT$turnover
             +working_dataT$tenureMedian+working_dataT$tenureDiversity+log(working_dataT$teamSize)
             +working_dataT$stCongruence+working_dataT$centrality+working_dataT$truckFactor
             +working_dataT$expertise+working_dataT$female+working_dataT$blauGender
             +(1 | working_dataT$window_idx ), REML=FALSE)
```

```
## boundary (singular) fit: see help('isSingular')
```

```
# Remove outlier
#romr.fnc(blackT, working_dataT, trim = 2.5)

# Applying vif <5
print(vif(organisationT))
```

```
## log(working_dataT$totalCommitters)    log(working_dataT$totalcommits)
##                           2.592895                           2.336089
##           working_dataT$projectAge            working_dataT$turnover
##                           1.341070                           1.248622
##         working_dataT$tenureMedian      working_dataT$tenureDiversity
##                           1.065323                           1.041936
##        log(working_dataT$teamSize)         working_dataT$stCongruence
##                           1.833618                           1.033152
##           working_dataT$centrality          working_dataT$truckFactor
##                           1.107221                           1.068792
##            working_dataT$expertise               working_dataT$female
##                           1.078462                           1.040100
##           working_dataT$blauGender
##                           2.022374
```

```
# Applying a Linear Mixed Model using the lmer function, after vif, NO REMOVALS
organisationT <- lmer(working_dataT$organisation~log(working_dataT$totalCommitters)+log(working_dataT$t
             +working_dataT$projectAge+working_dataT$turnover
             +working_dataT$tenureMedian+working_dataT$centrality+working_dataT$tenureDiversity
             +working_dataT$stCongruence+working_dataT$truckFactor
             +working_dataT$expertise+working_dataT$female+working_dataT$blauGender
             +(1 | working_dataT$window_idx ), REML=FALSE)
```

```
## boundary (singular) fit: see help('isSingular')
```

```
# print result
print(summary(organisationT))
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: working_dataT$organisation ~ log(working_dataT$totalCommitters) +
```

```
##     log(working_dataT$totalcommits) + working_dataT$projectAge +
##     working_dataT$turnover + working_dataT$tenureMedian + working_dataT$centrality +
##     working_dataT$tenureDiversity + working_dataT$stCongruence +
##     working_dataT$truckFactor + working_dataT$expertise + working_dataT$female +
##     working_dataT$blauGender + (1 | working_dataT$window_idx)
##
##      AIC      BIC   logLik deviance df.resid
##    758.9    807.8   -364.4    728.9      177
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.93131 -0.80195 -0.06154  0.81380  2.04131
##
## Random effects:
##  Groups                    Name        Variance Std.Dev.
##  working_dataT$window_idx (Intercept) 0.000    0.000
##  Residual                             2.608    1.615
## Number of obs: 192, groups:  working_dataT$window_idx, 24
##
## Fixed effects:
##                                    Estimate Std. Error t value
## (Intercept)                         1.76031    1.18296   1.488
## log(working_dataT$totalCommitters)  0.40372    0.15834   2.550
## log(working_dataT$totalcommits)    -0.14431    0.11819  -1.221
## working_dataT$projectAge           -0.02443    0.02768  -0.883
## working_dataT$turnover             -1.38135    0.55466  -2.490
## working_dataT$tenureMedian         -0.09215    0.06655  -1.385
## working_dataT$centrality            0.02872    0.27424   0.105
## working_dataT$tenureDiversity       0.01451    0.04720   0.308
## working_dataT$stCongruence          0.38007    0.34847   1.091
## working_dataT$truckFactor           0.04206    0.08775   0.479
## working_dataT$expertise            -0.37246    0.38274  -0.973
## working_dataT$female                0.03992    0.02095   1.906
## working_dataT$blauGender            1.94591    1.61696   1.203
##
## Correlation matrix not shown by default, as p = 13 > 12.
## Use print(summary(organisationT), correlation=TRUE)  or
##     vcov(summary(organisationT))        if you need it
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```r
# Applying anova
Anova(organisationT)
```

```
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: working_dataT$organisation
##                                    Chisq Df Pr(>Chisq)
## log(working_dataT$totalCommitters) 6.5014  1    0.01078 *
## log(working_dataT$totalcommits)    1.4908  1    0.22210
## working_dataT$projectAge           0.7789  1    0.37747
## working_dataT$turnover             6.2023  1    0.01276 *
## working_dataT$tenureMedian         1.9174  1    0.16614
## working_dataT$centrality           0.0110  1    0.91660
```

```
## working_dataT$tenureDiversity      0.0946  1    0.75846
## working_dataT$stCongruence         1.1896  1    0.27542
## working_dataT$truckFactor          0.2298  1    0.63167
## working_dataT$expertise            0.9470  1    0.33049
## working_dataT$female               3.6331  1    0.05664 .
## working_dataT$blauGender           1.4482  1    0.22881
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Save in a txt file
sink("trompe/output_organisation_trompenaars_confounding_variables.txt")
print(summary(organisationT))
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: working_dataT$organisation ~ log(working_dataT$totalCommitters) +
##     log(working_dataT$totalcommits) + working_dataT$projectAge +
##     working_dataT$turnover + working_dataT$tenureMedian + working_dataT$centrality +
##     working_dataT$tenureDiversity + working_dataT$stCongruence +
##     working_dataT$truckFactor + working_dataT$expertise + working_dataT$female +
##     working_dataT$blauGender + (1 | working_dataT$window_idx)
##
##      AIC      BIC   logLik deviance df.resid
##    758.9    807.8   -364.4    728.9      177
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.93131 -0.80195 -0.06154  0.81380  2.04131
##
## Random effects:
##  Groups                    Name        Variance Std.Dev.
##  working_dataT$window_idx (Intercept) 0.000    0.000
##  Residual                             2.608    1.615
## Number of obs: 192, groups:  working_dataT$window_idx, 24
##
## Fixed effects:
##                                   Estimate Std. Error t value
## (Intercept)                        1.76031    1.18296   1.488
## log(working_dataT$totalCommitters)  0.40372    0.15834   2.550
## log(working_dataT$totalcommits)    -0.14431    0.11819  -1.221
## working_dataT$projectAge           -0.02443    0.02768  -0.883
## working_dataT$turnover             -1.38135    0.55466  -2.490
## working_dataT$tenureMedian         -0.09215    0.06655  -1.385
## working_dataT$centrality            0.02872    0.27424   0.105
## working_dataT$tenureDiversity       0.01451    0.04720   0.308
## working_dataT$stCongruence          0.38007    0.34847   1.091
## working_dataT$truckFactor           0.04206    0.08775   0.479
## working_dataT$expertise            -0.37246    0.38274  -0.973
## working_dataT$female                0.03992    0.02095   1.906
## working_dataT$blauGender            1.94591    1.61696   1.203
##
## Correlation matrix not shown by default, as p = 13 > 12.
## Use print(summary(organisationT), correlation=TRUE)  or
##     vcov(summary(organisationT))        if you need it
## optimizer (nloptwrap) convergence code: 0 (OK)
```

```
## boundary (singular) fit: see help('isSingular')
```

```
Anova(organisationT)
```

```
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: working_dataT$organisation
##                                    Chisq Df Pr(>Chisq)
## log(working_dataT$totalCommitters) 6.5014  1    0.01078 *
## log(working_dataT$totalcommits)    1.4908  1    0.22210
## working_dataT$projectAge           0.7789  1    0.37747
## working_dataT$turnover             6.2023  1    0.01276 *
## working_dataT$tenureMedian         1.9174  1    0.16614
## working_dataT$centrality           0.0110  1    0.91660
## working_dataT$tenureDiversity      0.0946  1    0.75846
## working_dataT$stCongruence         1.1896  1    0.27542
## working_dataT$truckFactor          0.2298  1    0.63167
## working_dataT$expertise            0.9470  1    0.33049
## working_dataT$female               3.6331  1    0.05664 .
## working_dataT$blauGender           1.4482  1    0.22881
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
sink()
```

Linear Mixed Model using lmer function on all the confounding variables globe

```
#----------------------------
#ALL THE CONFOUNDING VARIABLES

# Applying a Linear Mixed Model using the lmer function
organisationT <- lmer(working_dataG$organisation~log(working_dataG$totalCommitters)+log(working_dataG$t
             +working_dataG$projectAge+working_dataG$turnover
             +working_dataG$tenureMedian+working_dataG$tenureDiversity+log(working_dataG$teamSize)
             +working_dataG$stCongruence+working_dataG$centrality+working_dataG$truckFactor
             +working_dataG$expertise+working_dataG$female+working_dataG$blauGender
             +(1 | working_dataG$window_idx ), REML=FALSE)
```

```
## boundary (singular) fit: see help('isSingular')
```

```
# Remove outlier
#romr.fnc(blackT, working_dataG, trim = 2.5)

# Applying vif <5
print(vif(organisationT))
```

```
## log(working_dataG$totalCommitters)    log(working_dataG$totalcommits)
##                           2.592895                           2.336089
##           working_dataG$projectAge             working_dataG$turnover
##                           1.341070                           1.248622
##         working_dataG$tenureMedian      working_dataG$tenureDiversity
##                           1.065323                           1.041936
##        log(working_dataG$teamSize)         working_dataG$stCongruence
##                           1.833618                           1.033152
```

```
##           working_dataG$centrality          working_dataG$truckFactor
##                           1.107221                           1.068792
##            working_dataG$expertise               working_dataG$female
##                           1.078462                           1.040100
##           working_dataG$blauGender
##                           2.022374
```

```
# Applying a Linear Mixed Model using the lmer function, after vif, NO REMOVALS
organisationT <- lmer(working_dataG$organisation~log(working_dataG$totalCommitters)+log(working_dataG$t
            +working_dataG$projectAge+working_dataG$turnover
            +working_dataG$tenureMedian+working_dataG$centrality+working_dataG$tenureDiversity
            +working_dataG$stCongruence+working_dataG$truckFactor
            +working_dataG$expertise+working_dataG$female+working_dataG$blauGender
            +(1 | working_dataG$window_idx ), REML=FALSE)
```

```
## boundary (singular) fit: see help('isSingular')
```

```
# print result
print(summary(organisationT))
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: working_dataG$organisation ~ log(working_dataG$totalCommitters) +
##     log(working_dataG$totalcommits) + working_dataG$projectAge +
##     working_dataG$turnover + working_dataG$tenureMedian + working_dataG$centrality +
##     working_dataG$tenureDiversity + working_dataG$stCongruence +
##     working_dataG$truckFactor + working_dataG$expertise + working_dataG$female +
##     working_dataG$blauGender + (1 | working_dataG$window_idx)
##
##      AIC      BIC   logLik deviance df.resid
##    758.9    807.8   -364.4    728.9      177
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.93131 -0.80195 -0.06154  0.81380  2.04131
##
## Random effects:
##  Groups                   Name        Variance Std.Dev.
##  working_dataG$window_idx (Intercept) 0.000    0.000
##  Residual                             2.608    1.615
## Number of obs: 192, groups:  working_dataG$window_idx, 24
##
## Fixed effects:
##                                    Estimate Std. Error t value
## (Intercept)                         1.76031    1.18296   1.488
## log(working_dataG$totalCommitters)  0.40372    0.15834   2.550
## log(working_dataG$totalcommits)    -0.14431    0.11819  -1.221
## working_dataG$projectAge           -0.02443    0.02768  -0.883
## working_dataG$turnover             -1.38135    0.55466  -2.490
## working_dataG$tenureMedian         -0.09215    0.06655  -1.385
## working_dataG$centrality            0.02872    0.27424   0.105
## working_dataG$tenureDiversity       0.01451    0.04720   0.308
## working_dataG$stCongruence          0.38007    0.34847   1.091
## working_dataG$truckFactor           0.04206    0.08775   0.479
## working_dataG$expertise            -0.37246    0.38274  -0.973
## working_dataG$female                0.03992    0.02095   1.906
## working_dataG$blauGender            1.94591    1.61696   1.203
```

```
##
## Correlation matrix not shown by default, as p = 13 > 12.
## Use print(summary(organisationT), correlation=TRUE)  or
##     vcov(summary(organisationT))        if you need it
```

```
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```
# Applying anova
Anova(organisationT)
```

```
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: working_dataG$organisation
##                                   Chisq Df Pr(>Chisq)
## log(working_dataG$totalCommitters) 6.5014  1    0.01078 *
## log(working_dataG$totalcommits)    1.4908  1    0.22210
## working_dataG$projectAge           0.7789  1    0.37747
## working_dataG$turnover             6.2023  1    0.01276 *
## working_dataG$tenureMedian         1.9174  1    0.16614
## working_dataG$centrality           0.0110  1    0.91660
## working_dataG$tenureDiversity      0.0946  1    0.75846
## working_dataG$stCongruence         1.1896  1    0.27542
## working_dataG$truckFactor          0.2298  1    0.63167
## working_dataG$expertise            0.9470  1    0.33049
## working_dataG$female               3.6331  1    0.05664 .
## working_dataG$blauGender           1.4482  1    0.22881
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Save in a txt file
sink("globe/output_organisation_globe_confounding_variables.txt")
print(summary(organisationT))
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: working_dataG$organisation ~ log(working_dataG$totalCommitters) +
##     log(working_dataG$totalcommits) + working_dataG$projectAge +
##     working_dataG$turnover + working_dataG$tenureMedian + working_dataG$centrality +
##     working_dataG$tenureDiversity + working_dataG$stCongruence +
##     working_dataG$truckFactor + working_dataG$expertise + working_dataG$female +
##     working_dataG$blauGender + (1 | working_dataG$window_idx)
##
##      AIC      BIC   logLik deviance df.resid
##    758.9    807.8   -364.4    728.9      177
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.93131 -0.80195 -0.06154  0.81380  2.04131
##
## Random effects:
##  Groups                    Name        Variance Std.Dev.
##  working_dataG$window_idx (Intercept) 0.000    0.000
##  Residual                             2.608    1.615
## Number of obs: 192, groups:  working_dataG$window_idx, 24
##
## Fixed effects:
```

```
##                                    Estimate Std. Error t value
## (Intercept)                         1.76031    1.18296   1.488
## log(working_dataG$totalCommitters)  0.40372    0.15834   2.550
## log(working_dataG$totalcommits)    -0.14431    0.11819  -1.221
## working_dataG$projectAge           -0.02443    0.02768  -0.883
## working_dataG$turnover             -1.38135    0.55466  -2.490
## working_dataG$tenureMedian         -0.09215    0.06655  -1.385
## working_dataG$centrality            0.02872    0.27424   0.105
## working_dataG$tenureDiversity       0.01451    0.04720   0.308
## working_dataG$stCongruence          0.38007    0.34847   1.091
## working_dataG$truckFactor           0.04206    0.08775   0.479
## working_dataG$expertise            -0.37246    0.38274  -0.973
## working_dataG$female                0.03992    0.02095   1.906
## working_dataG$blauGender            1.94591    1.61696   1.203
##
## Correlation matrix not shown by default, as p = 13 > 12.
## Use print(summary(organisationT), correlation=TRUE)  or
##     vcov(summary(organisationT))        if you need it
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```
Anova(organisationT)
```

```
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: working_dataG$organisation
##                                    Chisq Df Pr(>Chisq)
## log(working_dataG$totalCommitters) 6.5014  1    0.01078 *
## log(working_dataG$totalcommits)    1.4908  1    0.22210
## working_dataG$projectAge           0.7789  1    0.37747
## working_dataG$turnover             6.2023  1    0.01276 *
## working_dataG$tenureMedian         1.9174  1    0.16614
## working_dataG$centrality           0.0110  1    0.91660
## working_dataG$tenureDiversity      0.0946  1    0.75846
## working_dataG$stCongruence         1.1896  1    0.27542
## working_dataG$truckFactor          0.2298  1    0.63167
## working_dataG$expertise            0.9470  1    0.33049
## working_dataG$female               3.6331  1    0.05664 .
## working_dataG$blauGender           1.4482  1    0.22881
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
sink()
```

Linear Mixed Model using lmer function on only random effect hofsetde

```
#----------------------------
#ONLY RANDOM EFFECT

# Applying a Linear Mixed Model using the lmer function
organisation <- lmer(working_data$organisation~(1 | working_data$window_idx ), REML=FALSE)
```

```
## boundary (singular) fit: see help('isSingular')
```

```
# Remove outlier
#romr.fnc(black, working_data, trim = 2.5)

# print result
print(summary(organisation))
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: working_data$organisation ~ (1 | working_data$window_idx)
##
##      AIC      BIC   logLik deviance df.resid
##    755.5    765.3   -374.8    749.5      189
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.3999 -0.8131 -0.2262  0.9475  1.5344
##
## Random effects:
##  Groups                   Name        Variance Std.Dev.
##  working_data$window_idx (Intercept) 0.000    0.000
##  Residual                             2.904    1.704
## Number of obs: 192, groups:  working_data$window_idx, 24
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)    2.385      0.123    19.4
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```
# Save in a txt file
sink("hofstede/output_organisation_hofstede_random.txt")
print(summary(organisation))
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: working_data$organisation ~ (1 | working_data$window_idx)
##
##      AIC      BIC   logLik deviance df.resid
##    755.5    765.3   -374.8    749.5      189
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.3999 -0.8131 -0.2262  0.9475  1.5344
##
## Random effects:
##  Groups                   Name        Variance Std.Dev.
##  working_data$window_idx (Intercept) 0.000    0.000
##  Residual                             2.904    1.704
## Number of obs: 192, groups:  working_data$window_idx, 24
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)    2.385      0.123    19.4
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```
sink()
```

Linear Mixed Model using lmer function on only random effect trompenaars

```
#----------------------------
#ONLY RANDOM EFFECT

# Applying a Linear Mixed Model using the lmer function
organisationT <- lmer(working_dataT$organisation~(1 | working_dataT$window_idx ), REML=FALSE)
```

```
## boundary (singular) fit: see help('isSingular')
```

```
# Remove outlier
#romr.fnc(blackT, working_dataT, trim = 2.5)

# print result
print(summary(organisationT))
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: working_dataT$organisation ~ (1 | working_dataT$window_idx)
##
##      AIC      BIC   logLik deviance df.resid
##    755.5    765.3   -374.8    749.5      189
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.3999 -0.8131 -0.2262  0.9475  1.5344
##
## Random effects:
##  Groups                    Name        Variance Std.Dev.
##  working_dataT$window_idx (Intercept) 0.000    0.000
##  Residual                             2.904    1.704
## Number of obs: 192, groups:  working_dataT$window_idx, 24
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)    2.385      0.123    19.4
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```
# Save in a txt file
sink("trompe/output_organisation_trompenaars_random.txt")
print(summary(organisationT))
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: working_dataT$organisation ~ (1 | working_dataT$window_idx)
##
##      AIC      BIC   logLik deviance df.resid
##    755.5    765.3   -374.8    749.5      189
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.3999 -0.8131 -0.2262  0.9475  1.5344
```

```
##
## Random effects:
##  Groups                   Name        Variance Std.Dev.
##  working_dataT$window_idx (Intercept) 0.000    0.000
##  Residual                             2.904    1.704
## Number of obs: 192, groups:  working_dataT$window_idx, 24
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)    2.385      0.123    19.4
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```
sink()
```

Linear Mixed Model using lmer function on only random effect globe

```
#----------------------------
#ONLY RANDOM EFFECT

# Applying a Linear Mixed Model using the lmer function
organisationG <- lmer(working_dataG$organisation~(1 | working_dataG$window_idx ), REML=FALSE)
```

```
## boundary (singular) fit: see help('isSingular')
```

```
# Remove outlier
#romr.fnc(organisationG, working_dataG, trim = 2.5)

# print result
print(summary(organisationG))
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: working_dataG$organisation ~ (1 | working_dataG$window_idx)
##
##      AIC      BIC   logLik deviance df.resid
##    755.5    765.3   -374.8    749.5      189
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.3999 -0.8131 -0.2262  0.9475  1.5344
##
## Random effects:
##  Groups                   Name        Variance Std.Dev.
##  working_dataG$window_idx (Intercept) 0.000    0.000
##  Residual                             2.904    1.704
## Number of obs: 192, groups:  working_dataG$window_idx, 24
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)    2.385      0.123    19.4
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```r
# Save in a txt file
sink("globe/output_organisation_globe_random.txt")
print(summary(organisationG))
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: working_dataG$organisation ~ (1 | working_dataG$window_idx)
##
##      AIC      BIC   logLik deviance df.resid
##    755.5    765.3   -374.8    749.5      189
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.3999 -0.8131 -0.2262  0.9475  1.5344
##
## Random effects:
##  Groups                   Name        Variance Std.Dev.
##  working_dataG$window_idx (Intercept) 0.000    0.000
##  Residual                             2.904    1.704
## Number of obs: 192, groups:  working_dataG$window_idx, 24
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)    2.385      0.123    19.4
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```r
sink()
```