Supplemental Online Material

A Comparison of Multilevel Imputation Schemes for Random Coefficient Models: Fully Conditional Specification and Joint Model Imputation with Random Covariance Matrices

Craig K. Enders[1], Timothy Hayes[2], and Han Du[1]

[1]University of California – Los Angeles

[2]Florida International University

Supplement A: jomo Syntax for Joint Model Imputation

```
# load packages
library(jomo)
# read raw data
rawdata <- read.table("~/desktop/example/math.csv", sep = ",")
names(rawdata) = c("school", "stanmath", "pctminor", "math", "mathse",
   "ability", "teachexp", "condition")
rawdata[rawdata == 999] <- NA
rawdata$icept <- 1
# imputation
set.seed(90291)
l1miss <- c("math", "mathse")
l2miss <- c("teachexp")
l2comp <- c("icept", "condition")
impdata <- jomo(rawdata[l1miss], Y2 = rawdata[l2miss],
   X2 = rawdata[l2comp], clus = rawdata$school, nburn = 2000,
   nbetween = 2000, nimp = 20, meth = "random")
# write data
impdata <- impdata[impdata$Imputation != 0,]
keepvars <- c("Imputation", "clus", "math", "mathse", "teachexp",
   "condition")
write.table(impdata[keepvars],"~/desktop/example/jomoimputations.csv",
   sep = ",", col.names = F, row.names = F)
```

Supplement B: Blimp Syntax for Fully Conditional Specification

```
DATA: ~/desktop/example/math.csv;
VARIABLES: school stanmath pctminor math mathse ability
   teachexp condition;
ORDINAL: ;
NOMINAL: ;
MISSING: 999;
MODEL: school ~ math:mathse teachexp condition;
NIMPS: 20;
THIN: 2000;
BURN: 2000;
SEED: 90291;
OUTFILE: ~/desktop/example/blimpimputations.csv;
OPTIONS: stacked;
```

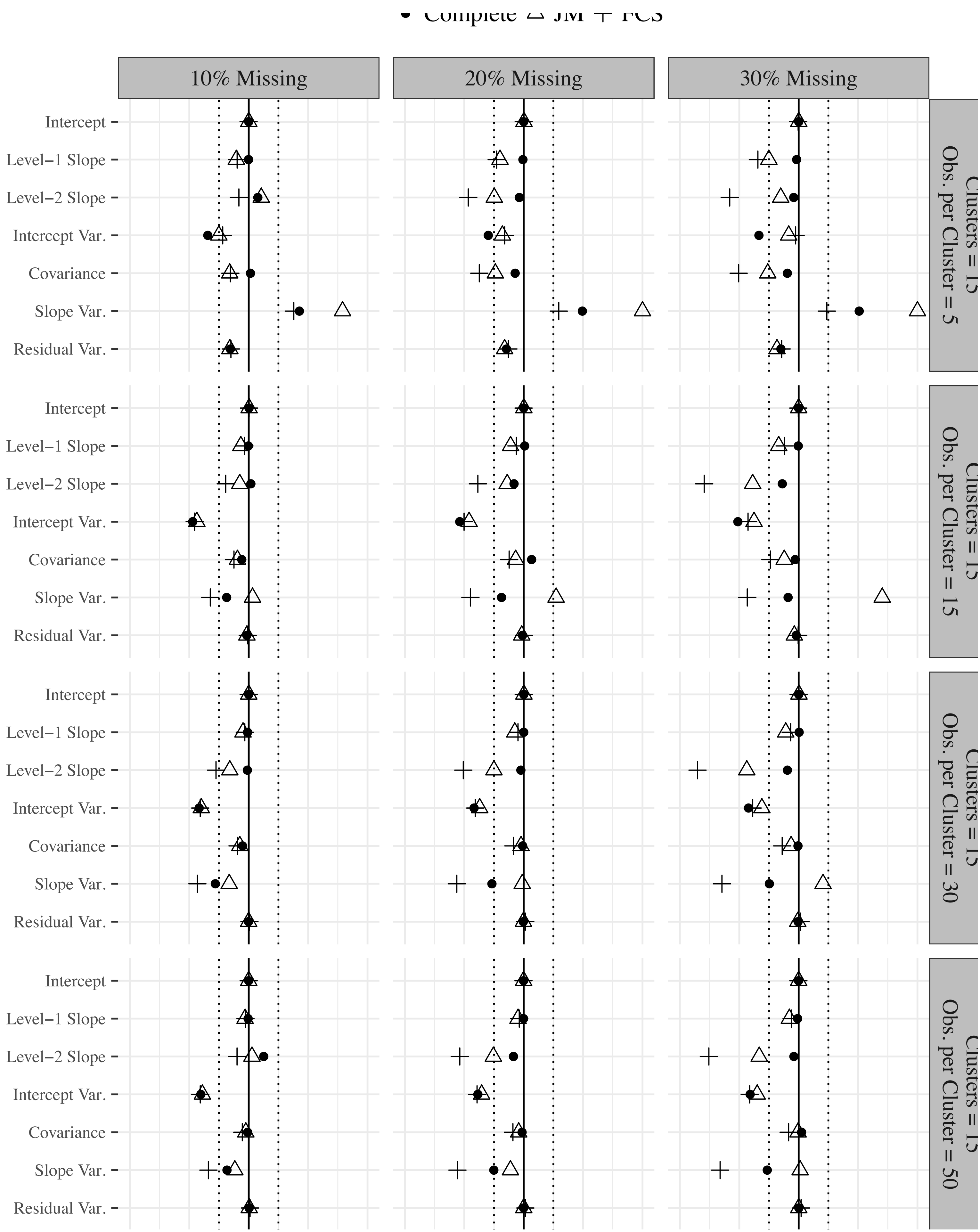
Supplement C: R Syntax for Analysis and Pooling Steps

```
# Required packages
library(mitml)
library(lme4)

# Read blimp data
filepath <- "~/desktop/example/blimpimputations.csv"
impdata <- read.csv(filepath, header = F)
names(impdata) <- c("imputation", "school", "stanmath", "pctminor",
"math", "mathse", "ability", "teachexp", "condition")

# Analyze data and pool estimates
model1 <- "math ~ mathse + teachexp + condition + (mathse|school)"
implist <-  as.mitml.list(split(impdata, impdata$imputation))
randomslope <- with(implist, lmer(model1, REML = F))
estimates <- testEstimates(randomslope, var.comp = T, df.com = NULL)
estimates

# Compare models with likelihood ratio test
model0 <- "math ~ mathse + teachexp + condition + (1|school)"
ranintercept <- with(implist, lmer(model0, REML = F))
testModels(randomslope, ranintercept, method = "D3")

# Read jomo data
filepath <- "~/desktop/example/jomoimputations.csv"
impdata <- read.csv(filepath, header = F)
names(impdata) <- c("imputation", "school", "math", "mathse",
"teachexp", "condition")

# Analyze data and pool estimates
model1 <- "math ~ mathse + teachexp + condition + (mathse|school)"
implist <-  as.mitml.list(split(impdata, impdata$imputation))
randomslope <- with(implist, lmer(model1, REML = F))
```

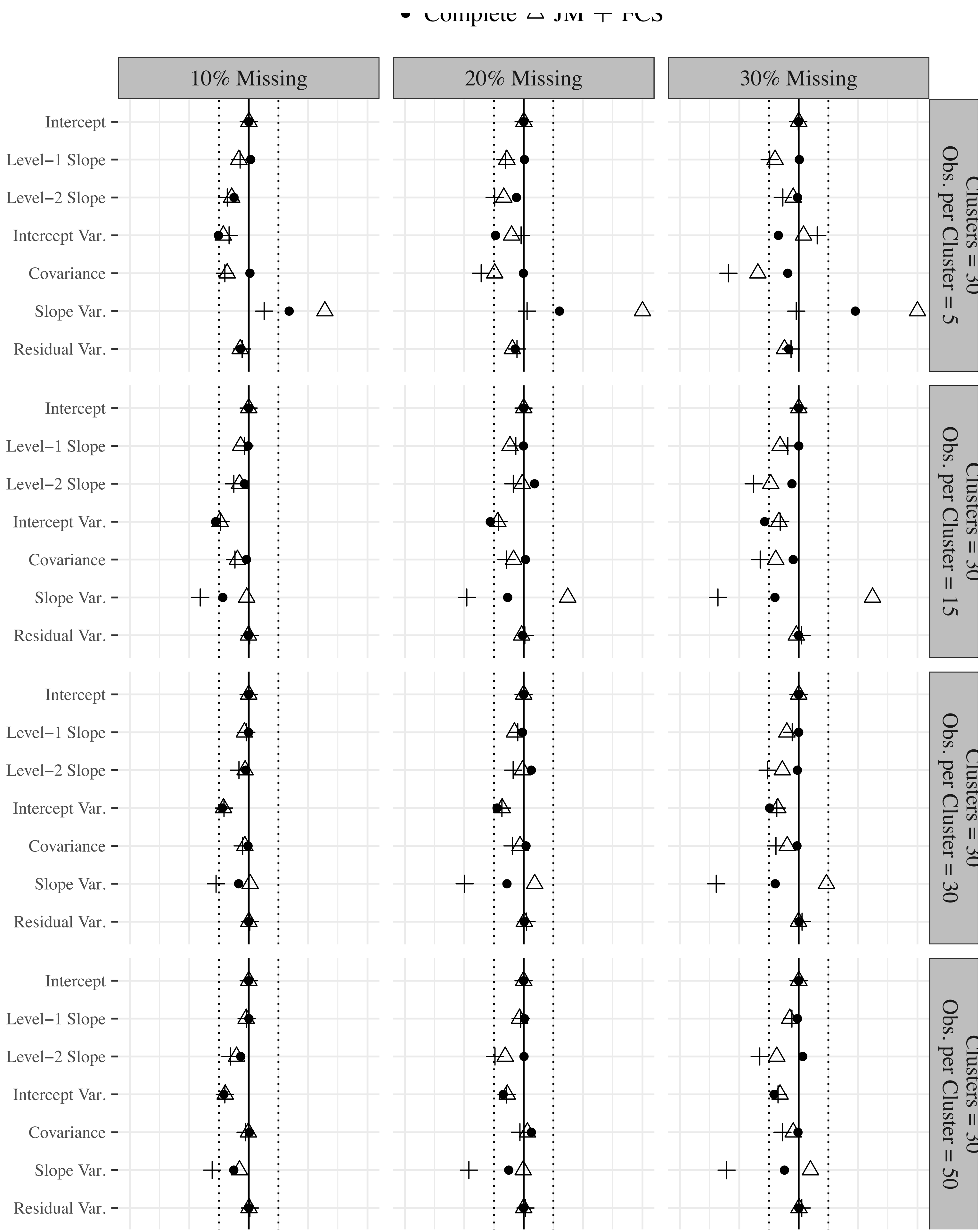
```
estimates <- testEstimates(randomslope, var.comp = T, df.com = NULL)
estimates

# Compare models with likelihood ratio test
model0 <- "math ~ mathse + teachexp + condition + (1|school)"
ranintercept <- with(implist, lmer(model0, REML = F))
testModels(randomslope, ranintercept, method = "D3")
```

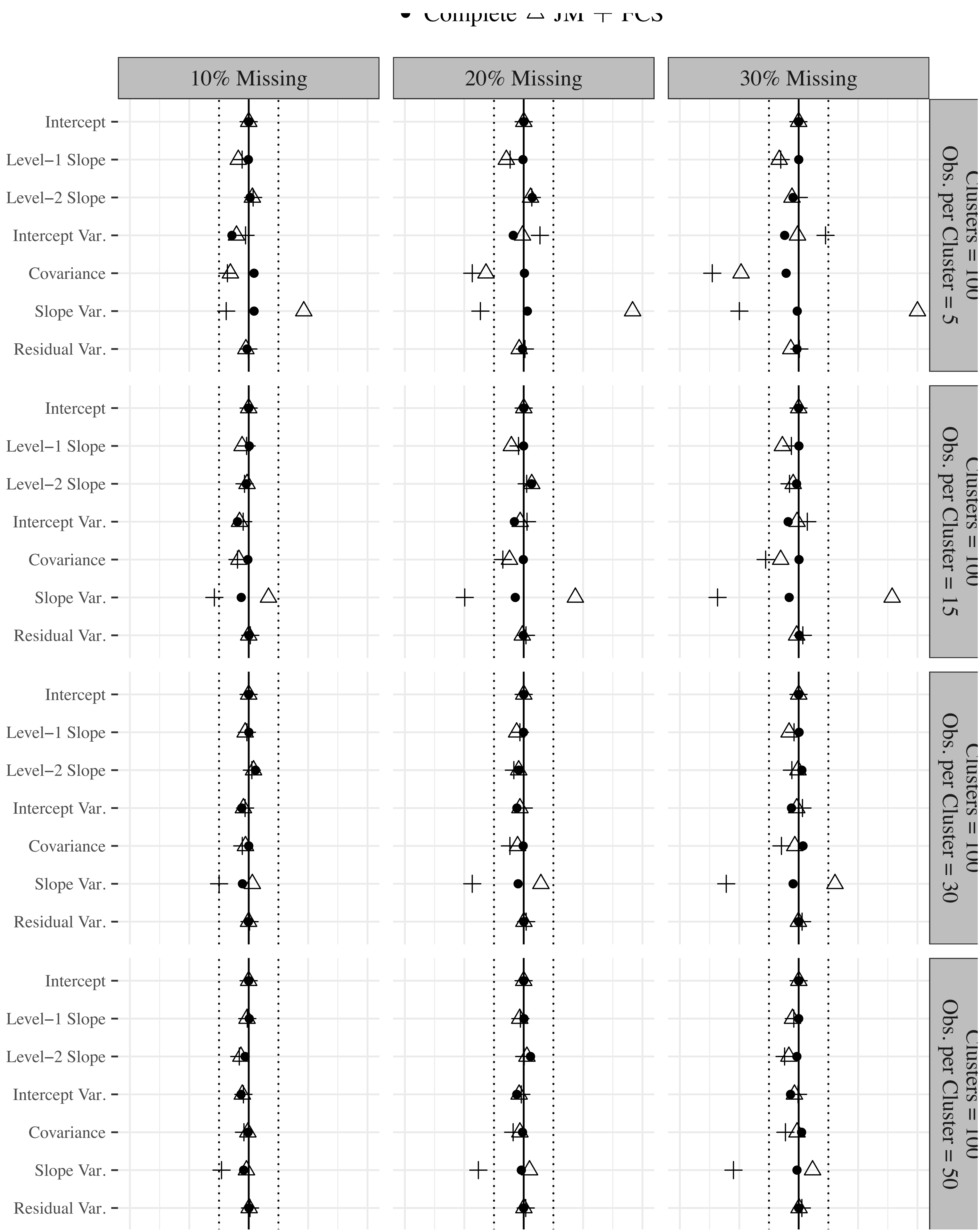
Supplement D: Trellis Plots of Relative Bias

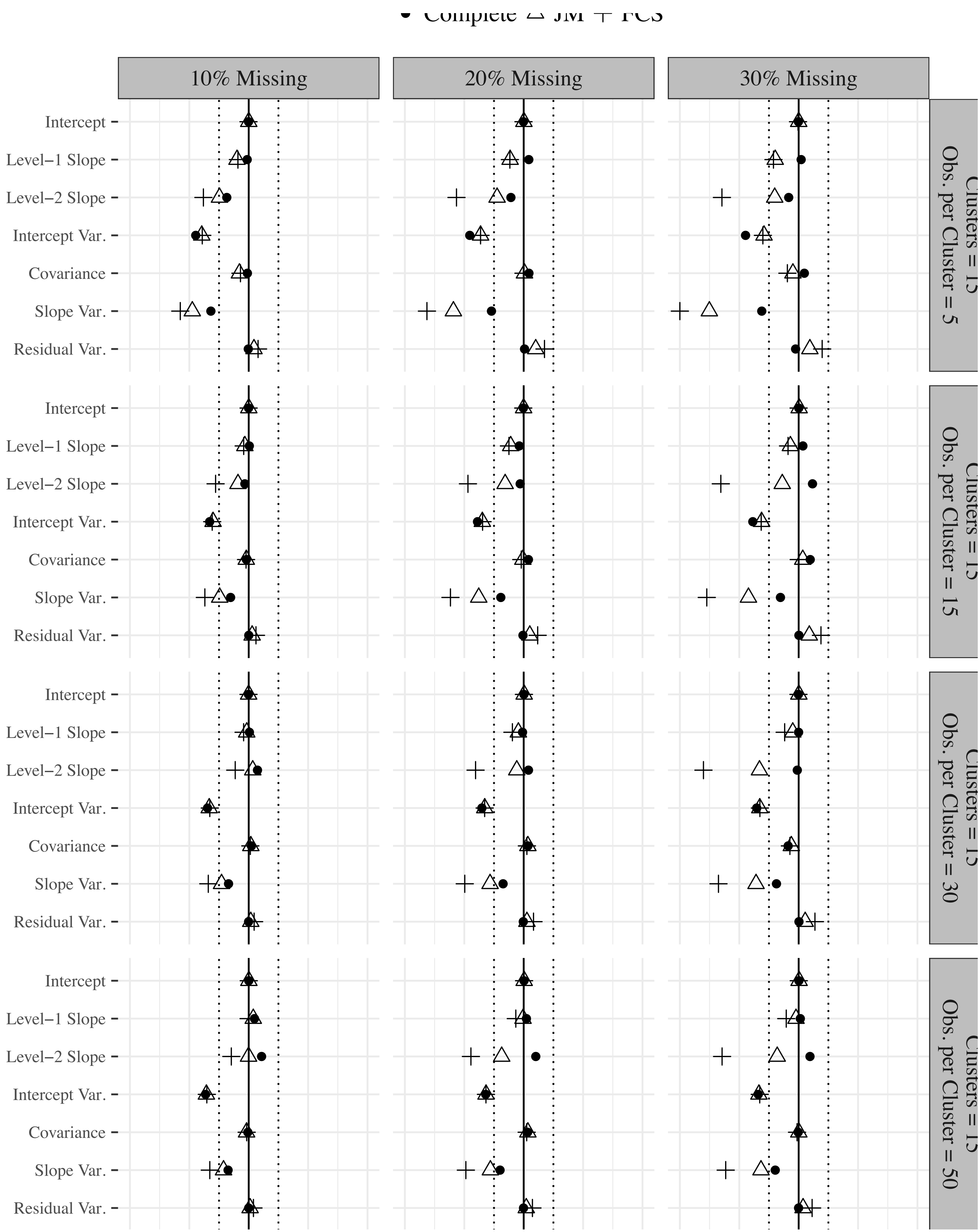
Uncorrelated Random Effects, Auxiliary as the Cause of Missingness



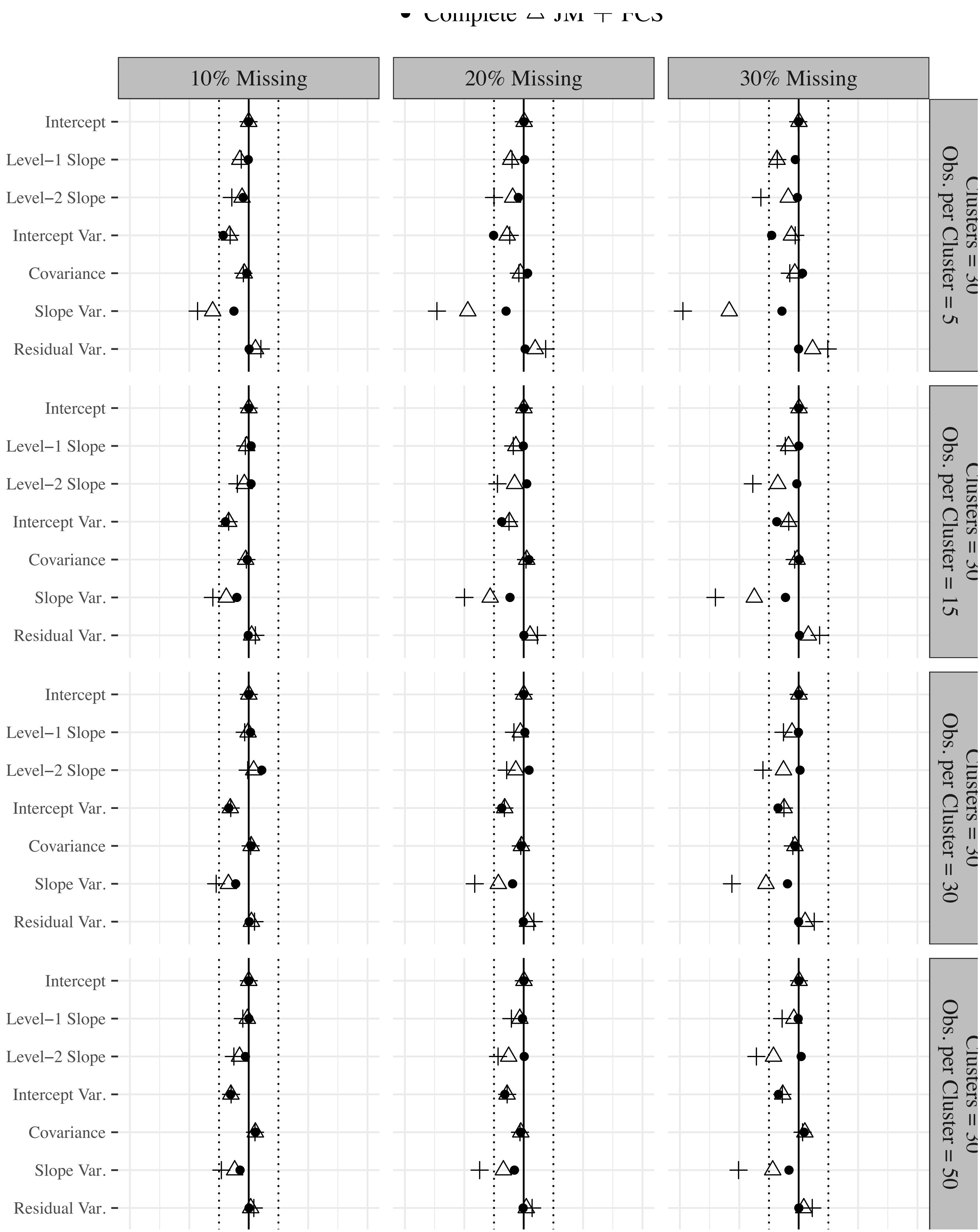
Complete JM FCS
10% Missing
20% Missing
30% Missing
Intercept
Level−1 Slope
Level−2 Slope
Intercept Var.
Covariance
Slope Var.
Residual Var.
Clusters = 15
Obs. per Cluster = 5
Obs. per Cluster = 15
Obs. per Cluster = 30
Obs. per Cluster = 50

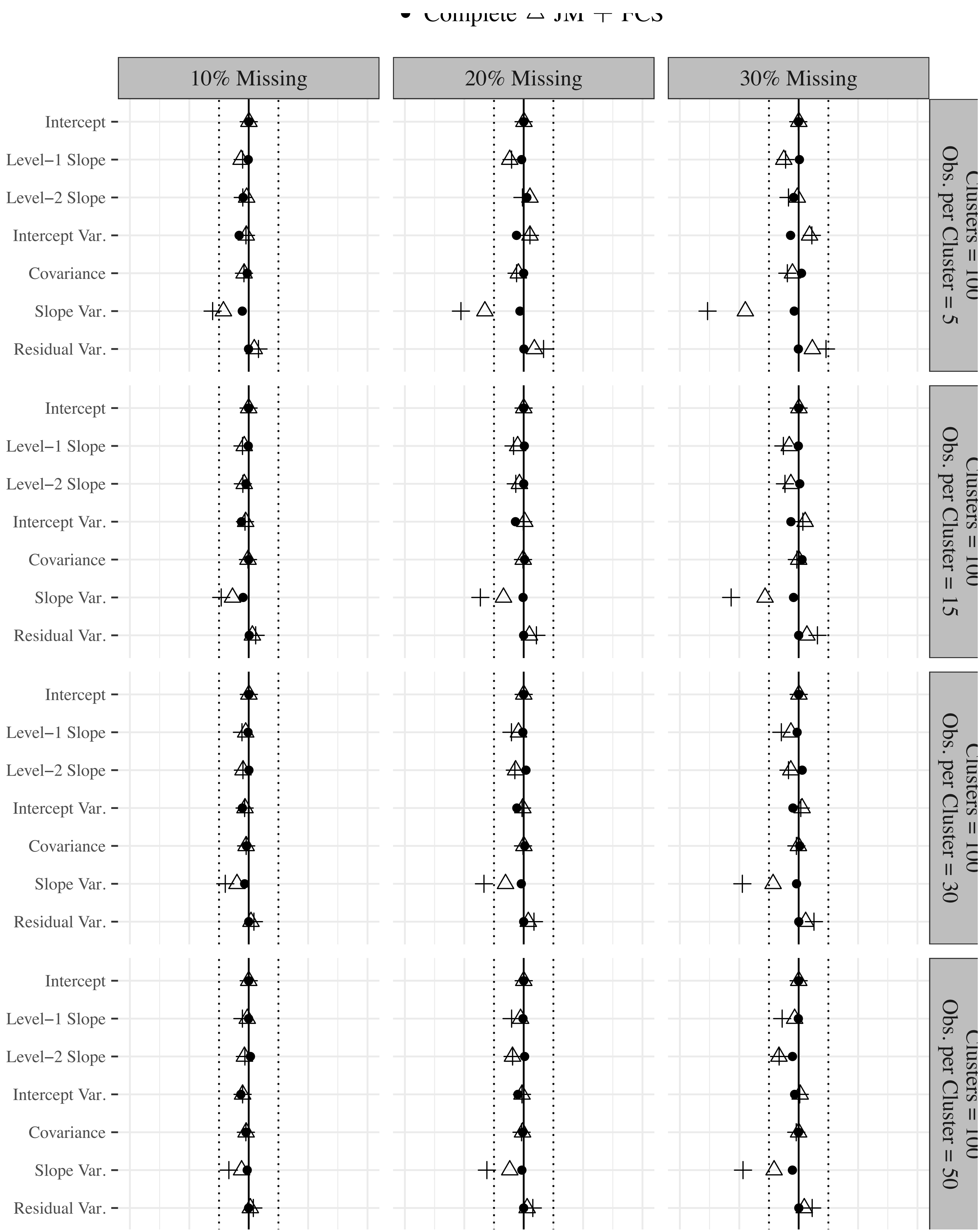

Complete JM FCS
10% Missing
20% Missing
30% Missing
Intercept
Level−1 Slope
Level−2 Slope
Intercept Var.
Covariance
Slope Var.
Residual Var.
Clusters = 30
Obs. per Cluster = 5
Obs. per Cluster = 15
Obs. per Cluster = 30
Obs. per Cluster = 50


● Complete △ JM + FCS
10% Missing
20% Missing
30% Missing
Intercept
Level−1 Slope
Level−2 Slope
Intercept Var.
Covariance
Slope Var.
Residual Var.
Clusters = 100
Obs. per Cluster = 5
Obs. per Cluster = 15
Obs. per Cluster = 30
Obs. per Cluster = 50



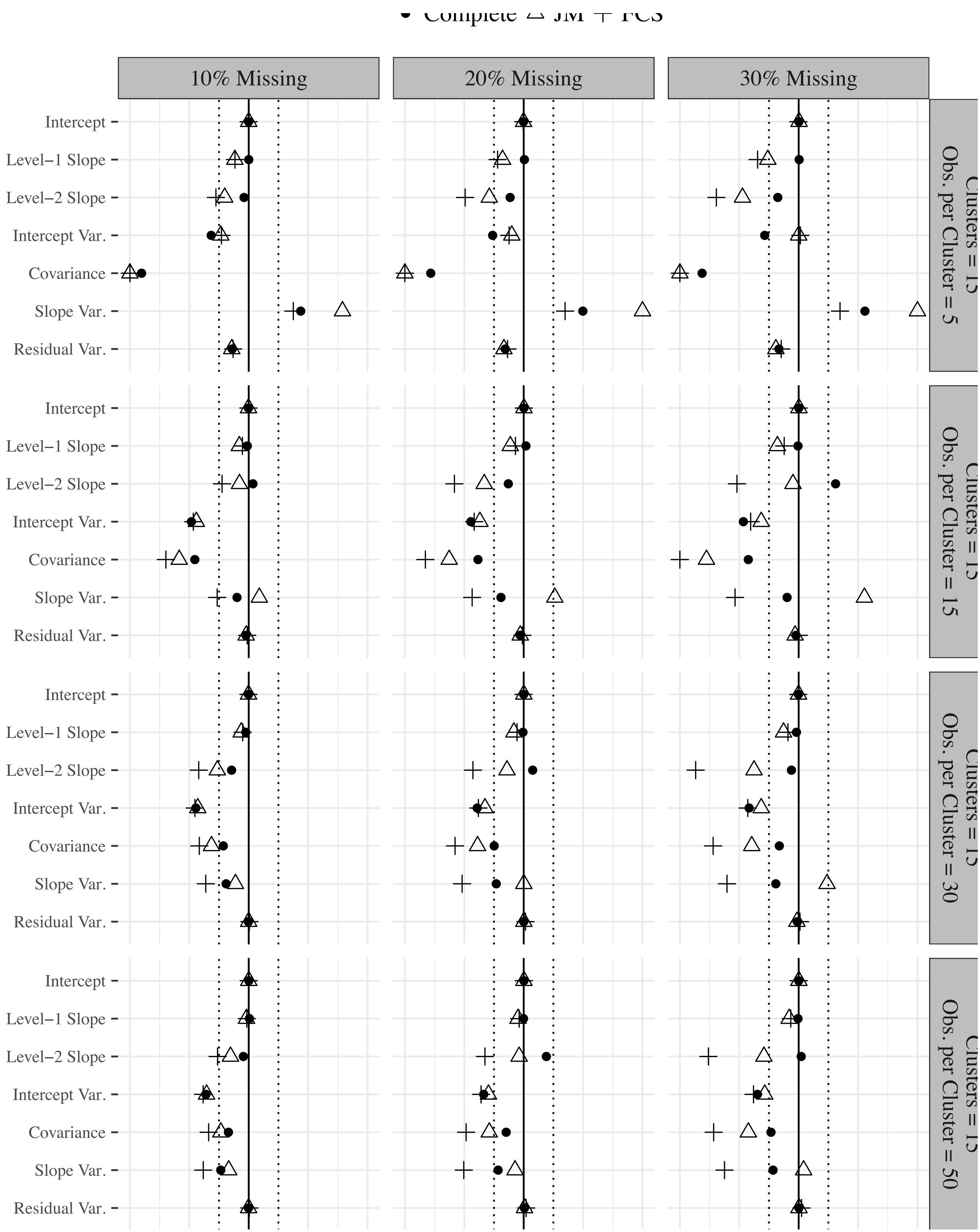
Complete JM FCS
10% Missing
20% Missing
30% Missing
Intercept
Level−1 Slope
Level−2 Slope
Intercept Var.
Covariance
Slope Var.
Residual Var.
Clusters = 15
Obs. per Cluster = 5
Obs. per Cluster = 15
Obs. per Cluster = 30
Obs. per Cluster = 50


Complete JM FCS
10% Missing
20% Missing
30% Missing
Intercept
Level–1 Slope
Level–2 Slope
Intercept Var.
Covariance
Slope Var.
Residual Var.
Clusters = 30
Obs. per Cluster = 5
Obs. per Cluster = 15
Obs. per Cluster = 30
Obs. per Cluster = 50

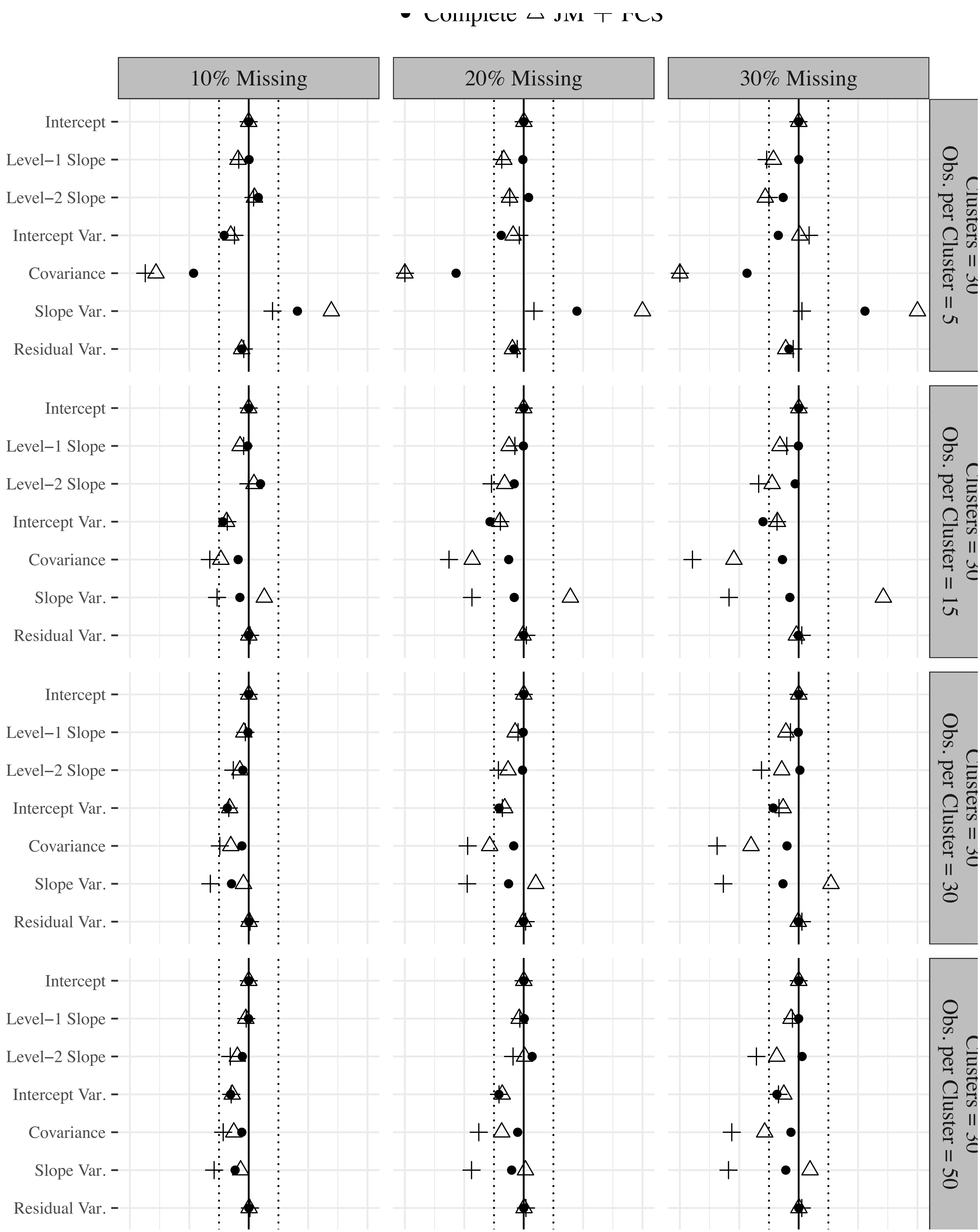

● Complete △ JM + FCS
10% Missing
20% Missing
30% Missing
Intercept
Level−1 Slope
Level−2 Slope
Intercept Var.
Covariance
Slope Var.
Residual Var.
Clusters = 100
Obs. per Cluster = 5
Clusters = 100
Obs. per Cluster = 15
Clusters = 100
Obs. per Cluster = 30
Clusters = 100
Obs. per Cluster = 50

Supplement E: Trellis Plots of Relative Bias

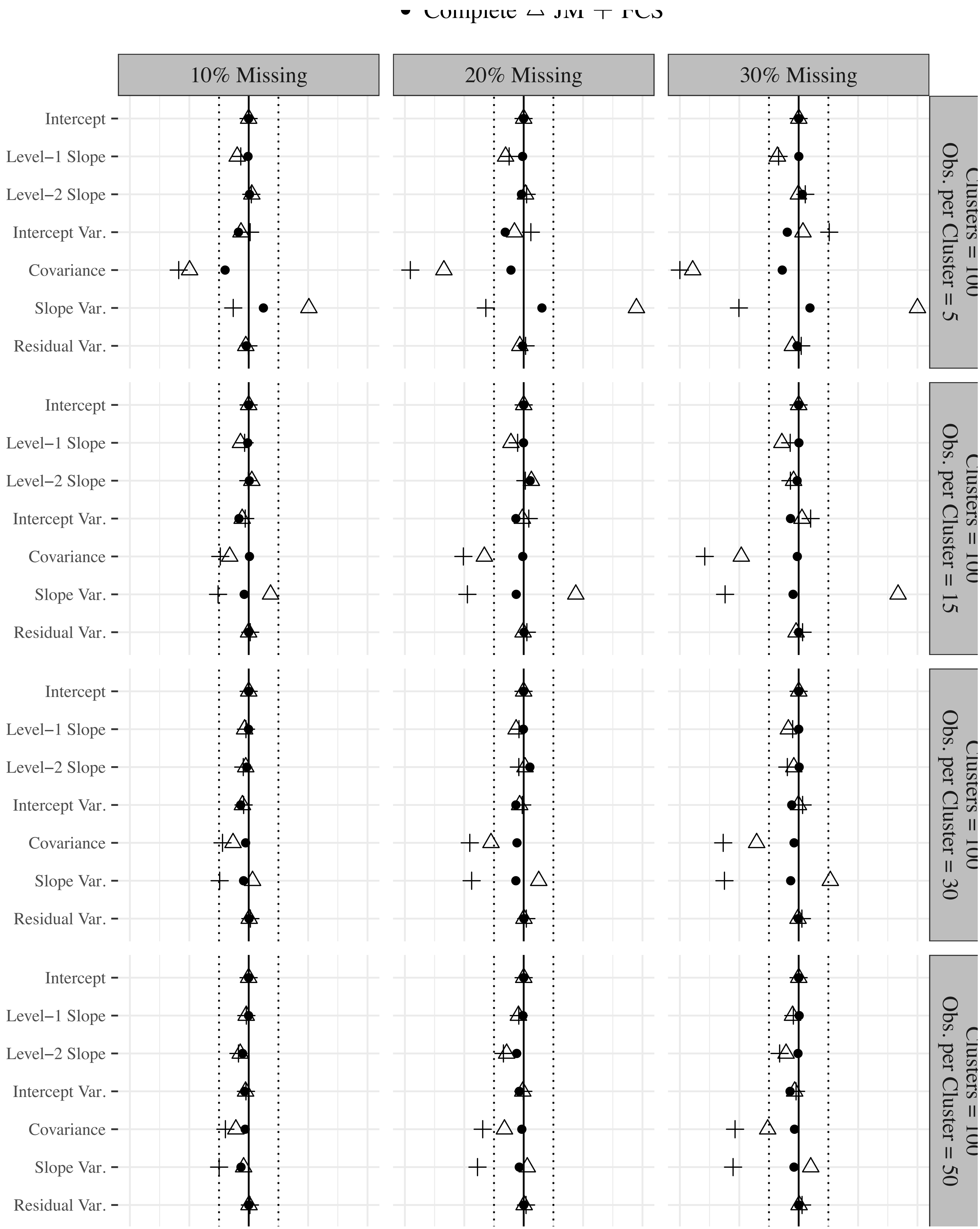
Correlated Random Effects, Auxiliary as the Cause of Missingness



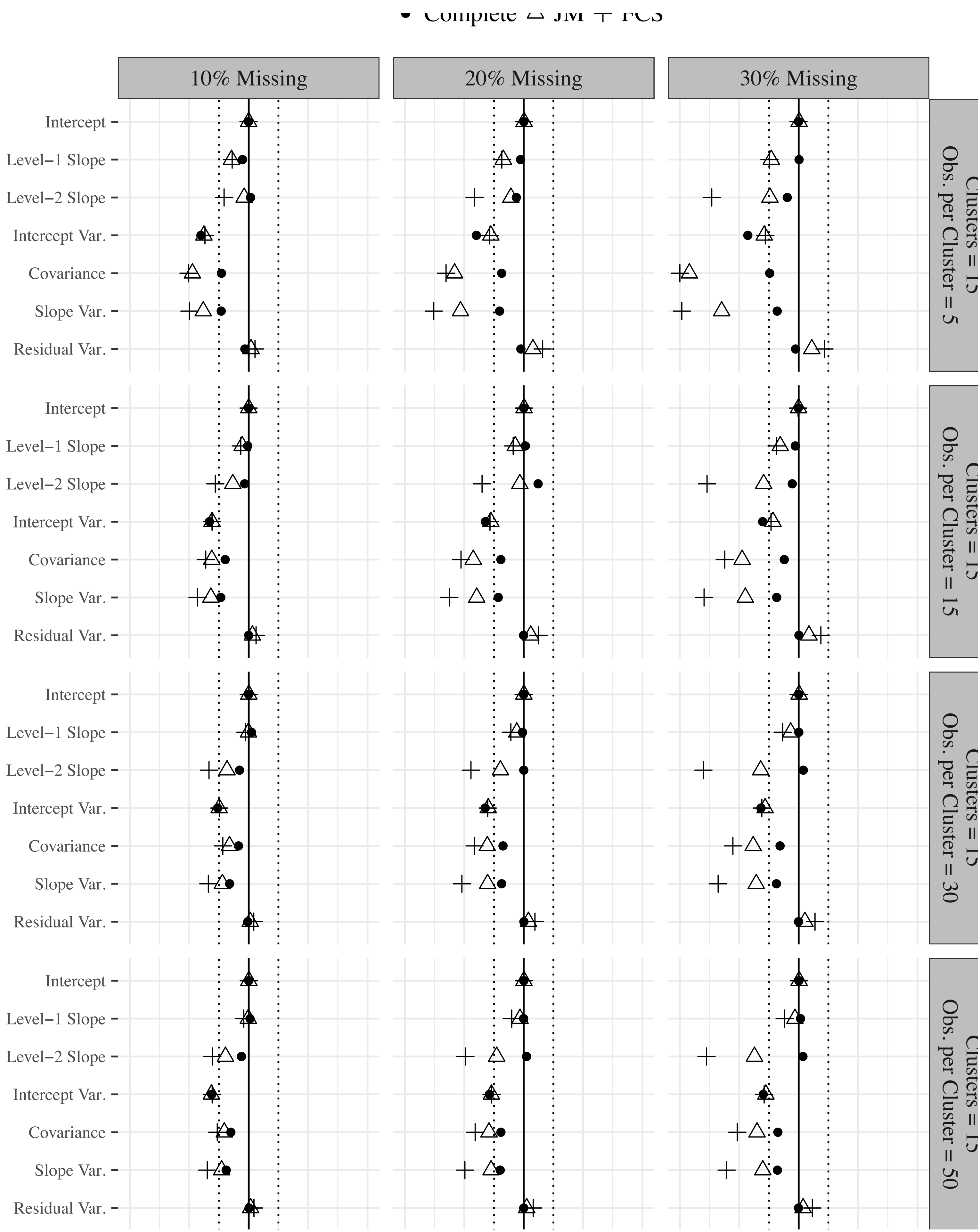
Complete JM FCS
10% Missing
20% Missing
30% Missing
Intercept
Level−1 Slope
Level−2 Slope
Intercept Var.
Covariance
Slope Var.
Residual Var.
Clusters = 15
Obs. per Cluster = 5
Obs. per Cluster = 15
Obs. per Cluster = 30
Obs. per Cluster = 50



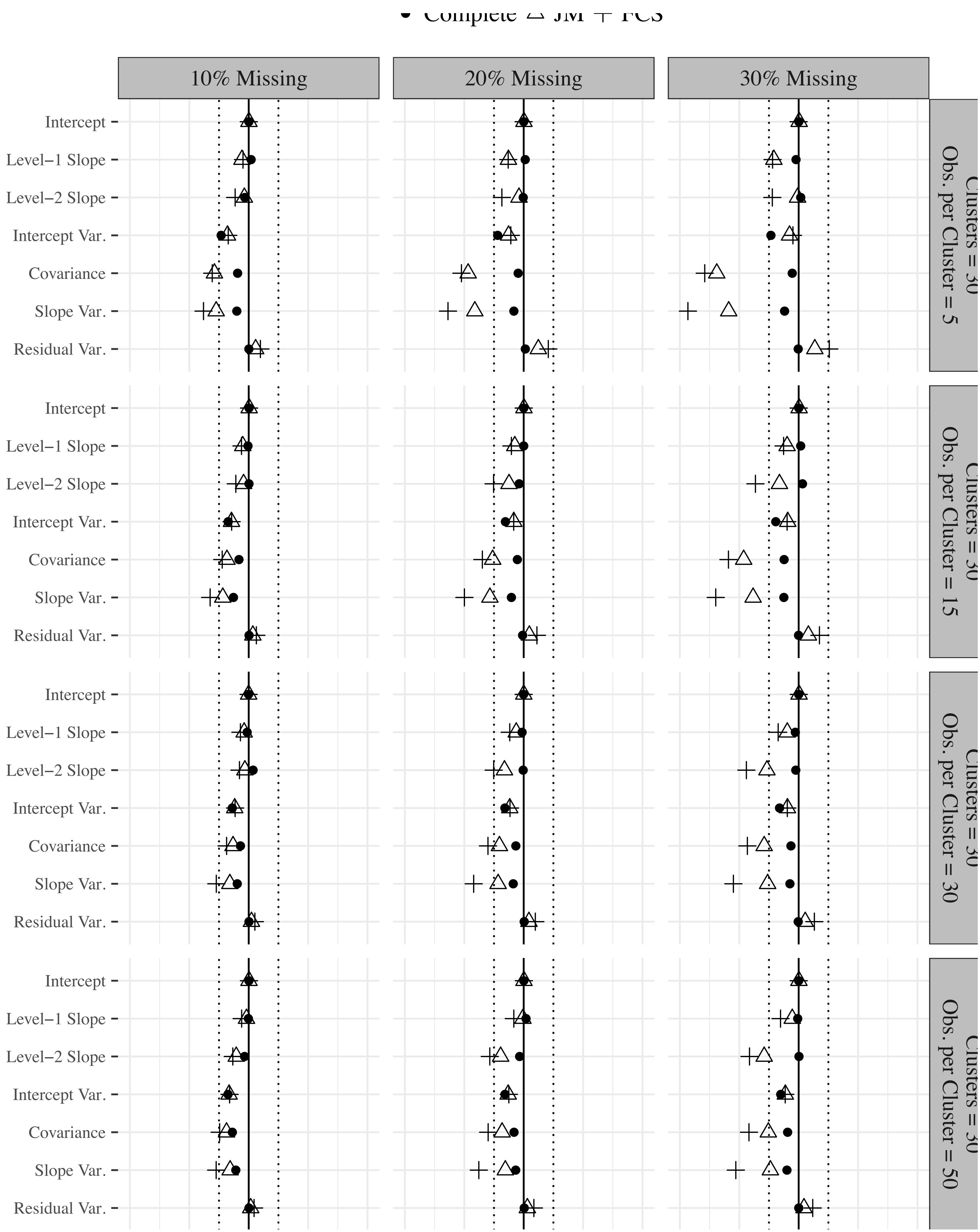
Complete JM FCS
10% Missing
20% Missing
30% Missing
Intercept
Level−1 Slope
Level−2 Slope
Intercept Var.
Covariance
Slope Var.
Residual Var.
Clusters = 30
Obs. per Cluster = 5
Obs. per Cluster = 15
Obs. per Cluster = 30
Obs. per Cluster = 50



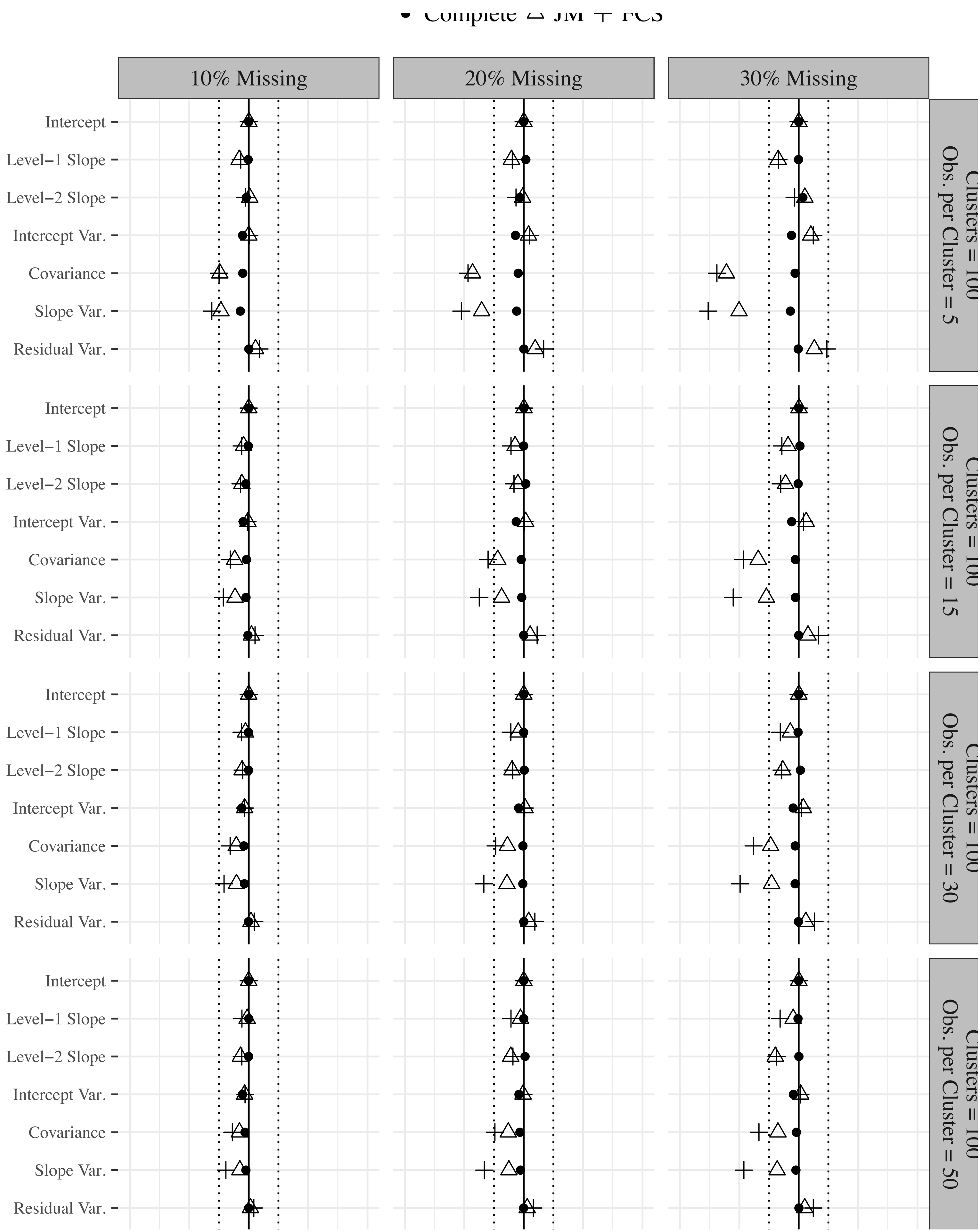
Complete JM FCS
10% Missing
20% Missing
30% Missing
Intercept
Level−1 Slope
Level−2 Slope
Intercept Var.
Covariance
Slope Var.
Residual Var.
Clusters = 100
Obs. per Cluster = 5
Obs. per Cluster = 15
Obs. per Cluster = 30
Obs. per Cluster = 50



Complete JM FCS
10% Missing
20% Missing
30% Missing
Intercept
Level−1 Slope
Level−2 Slope
Intercept Var.
Covariance
Slope Var.
Residual Var.
Clusters = 15
Obs. per Cluster = 5
Obs. per Cluster = 15
Obs. per Cluster = 30
Obs. per Cluster = 50



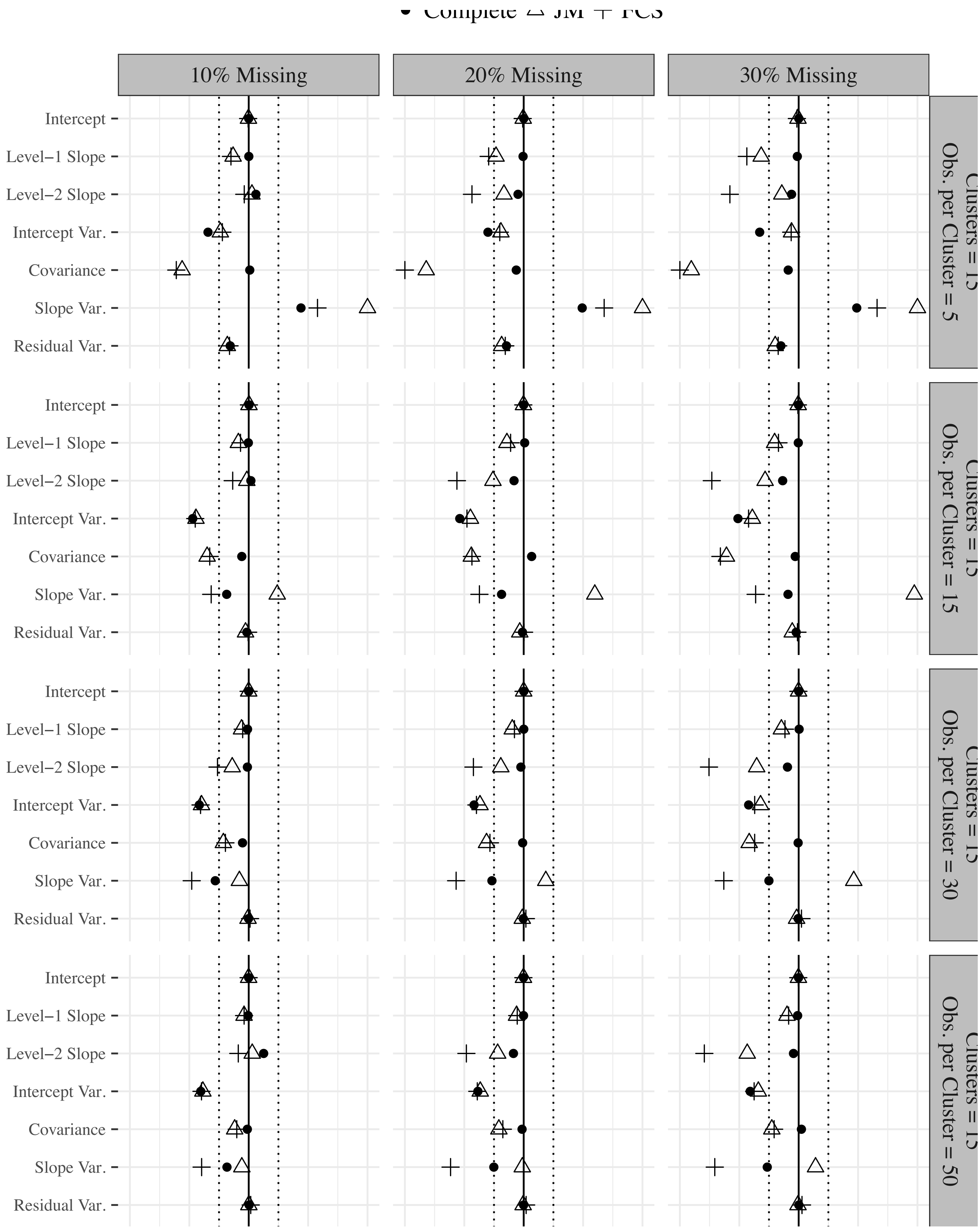
Complete JM FCS
10% Missing
20% Missing
30% Missing
Intercept
Level−1 Slope
Level−2 Slope
Intercept Var.
Covariance
Slope Var.
Residual Var.
Clusters = 30
Obs. per Cluster = 5
Obs. per Cluster = 15
Obs. per Cluster = 30
Obs. per Cluster = 50

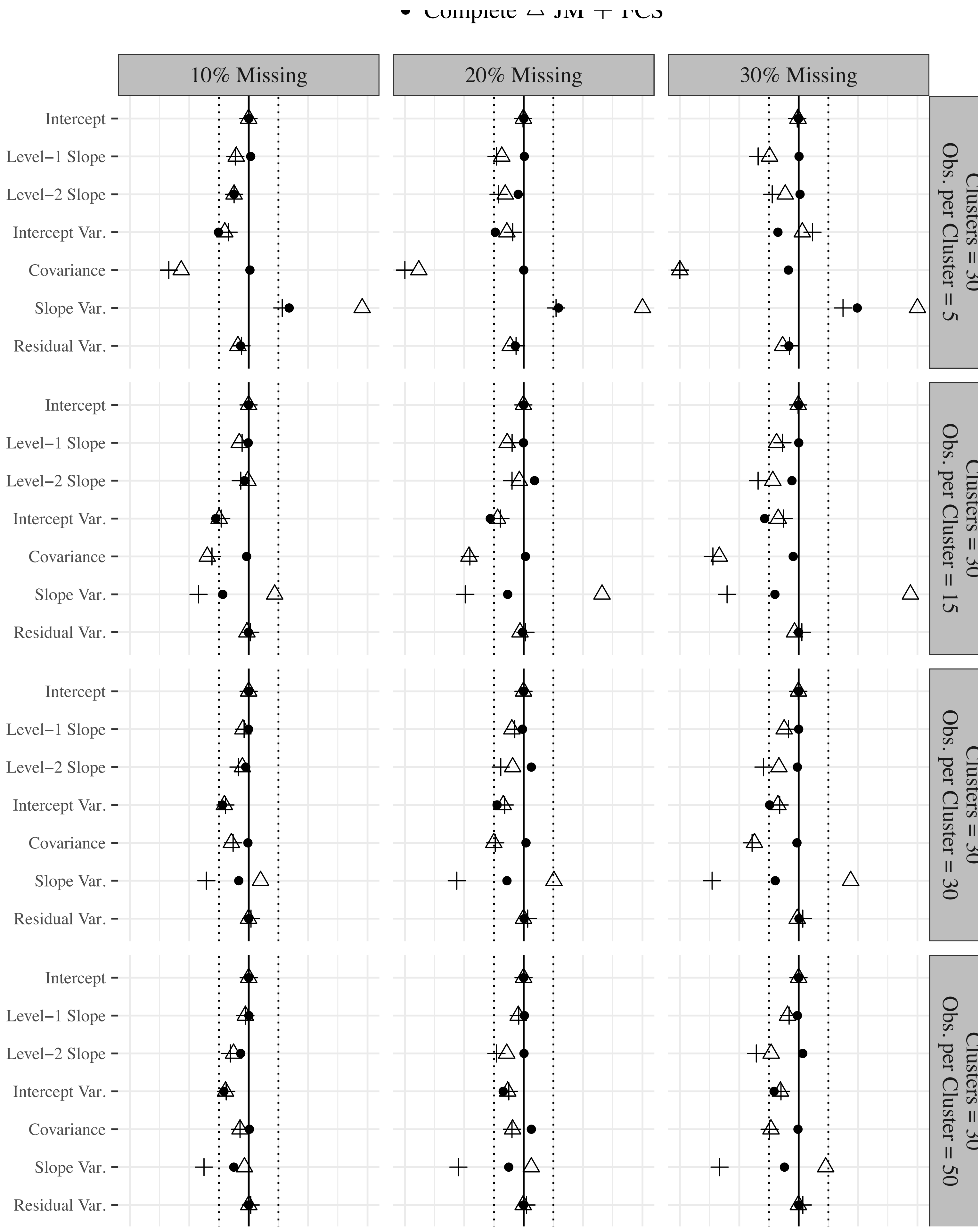


Complete JM FCS
10% Missing
20% Missing
30% Missing
Intercept
Level−1 Slope
Level−2 Slope
Intercept Var.
Covariance
Slope Var.
Residual Var.
Clusters = 100
Obs. per Cluster = 5
Obs. per Cluster = 15
Obs. per Cluster = 30
Obs. per Cluster = 50

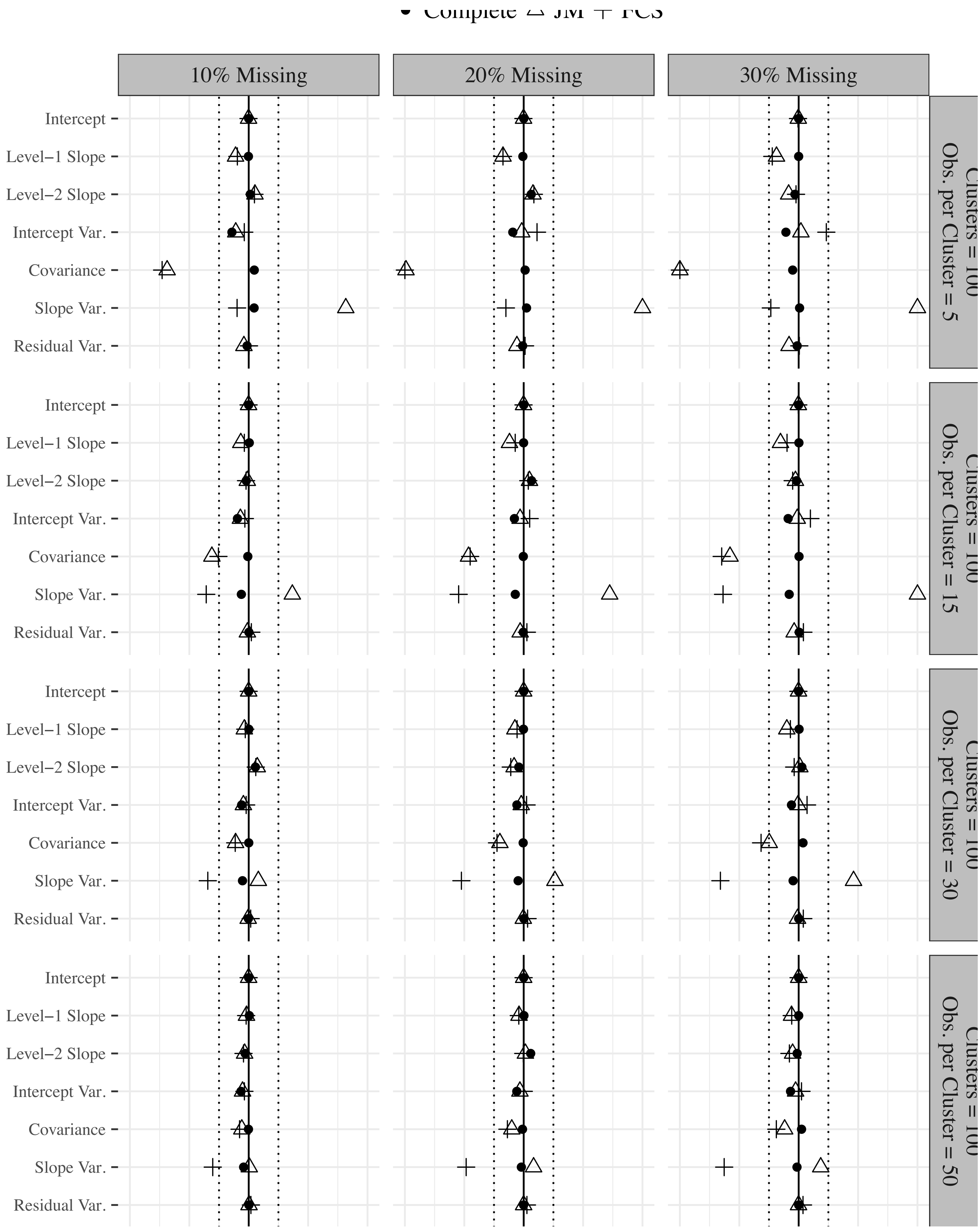
Supplement F: Trellis Plots of Relative Bias

Uncorrelated Random Effects, Outcome as the Cause of Missingness

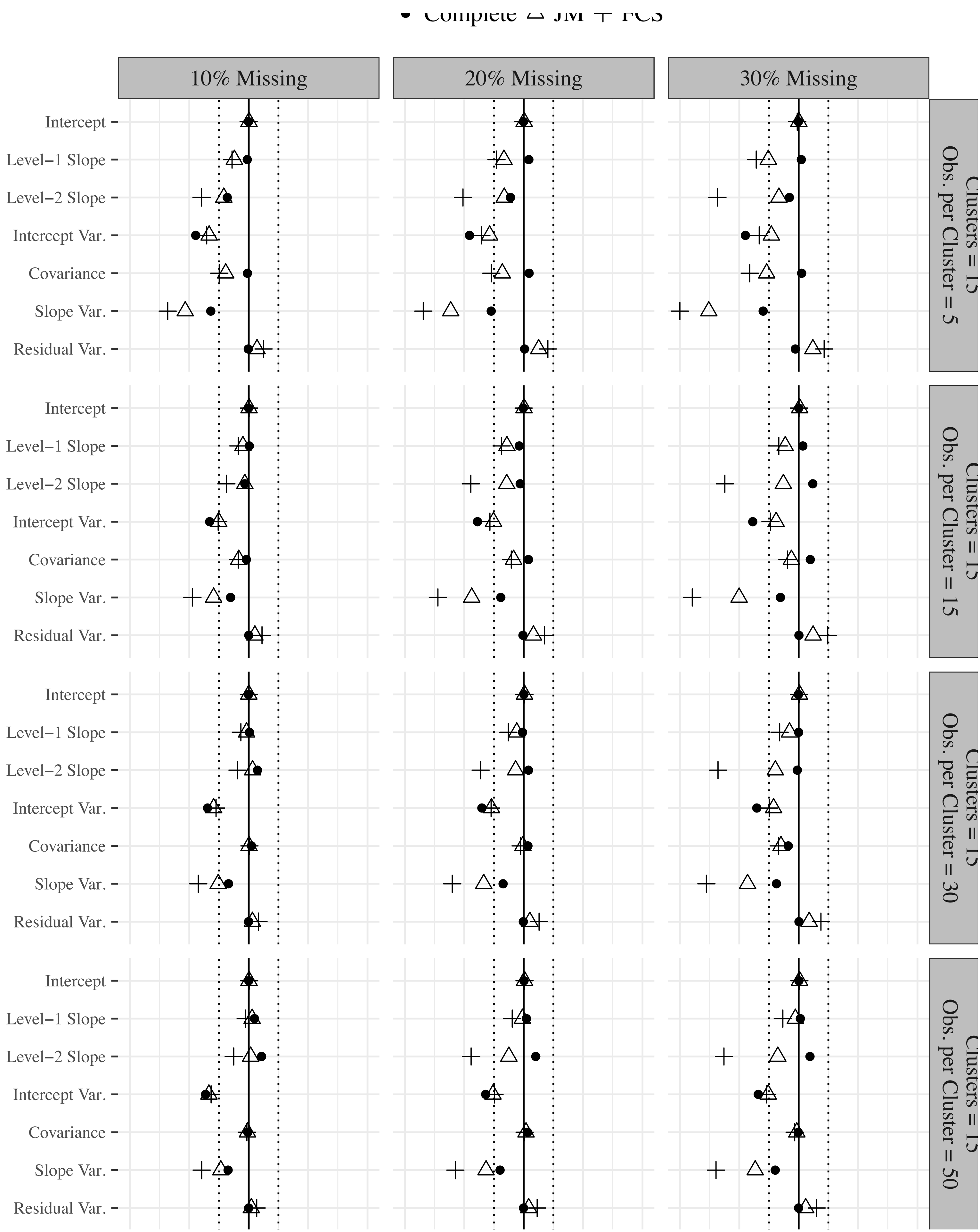

Complete
JM
FCS
10% Missing
20% Missing
30% Missing
Intercept
Level−1 Slope
Level−2 Slope
Intercept Var.
Covariance
Slope Var.
Residual Var.
Clusters = 15
Obs. per Cluster = 5
Obs. per Cluster = 15
Obs. per Cluster = 30
Obs. per Cluster = 50

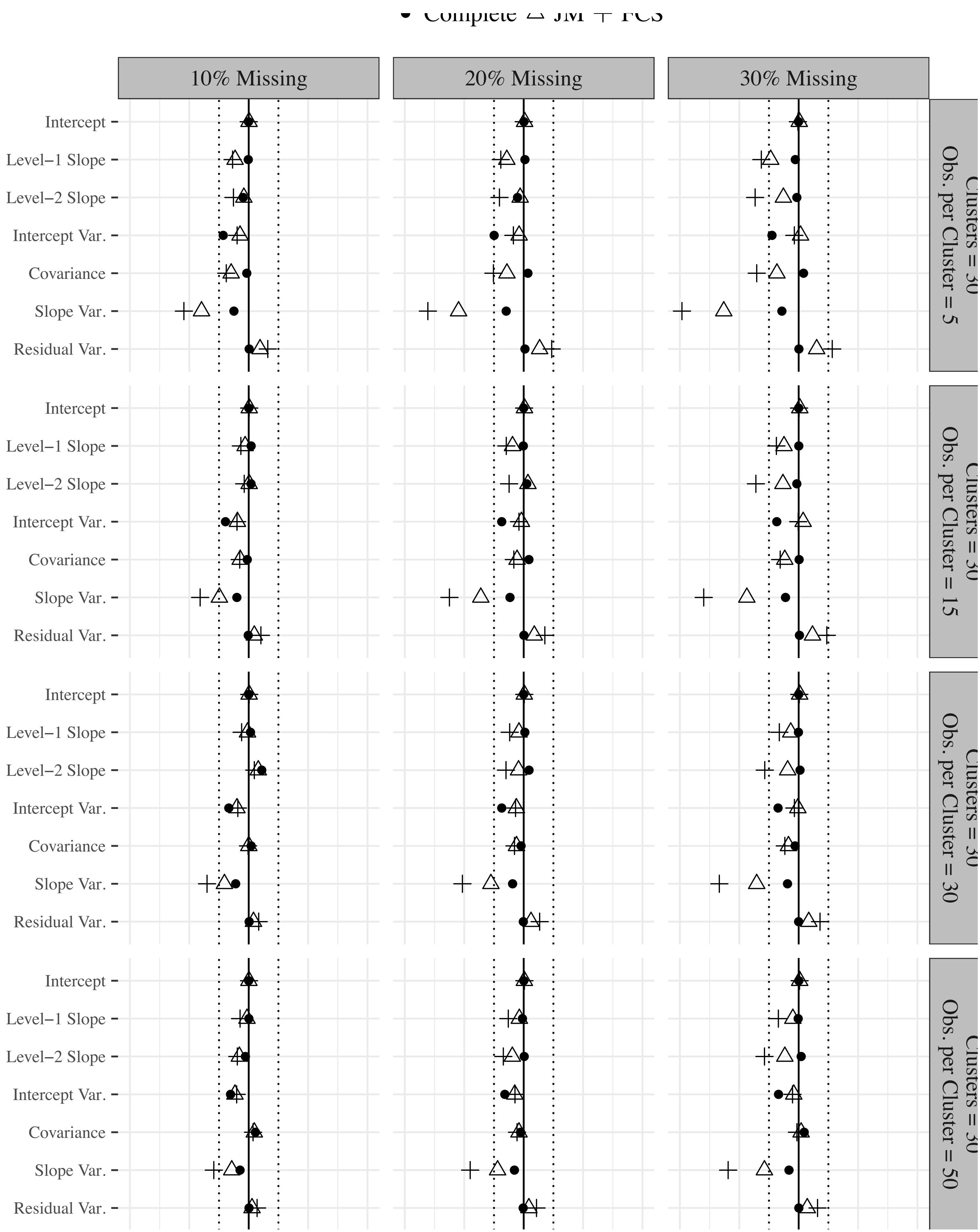


Complete
JM
FCS
10% Missing
20% Missing
30% Missing
Intercept
Level−1 Slope
Level−2 Slope
Intercept Var.
Covariance
Slope Var.
Residual Var.
Clusters = 30
Obs. per Cluster = 5
Obs. per Cluster = 15
Obs. per Cluster = 30
Obs. per Cluster = 50

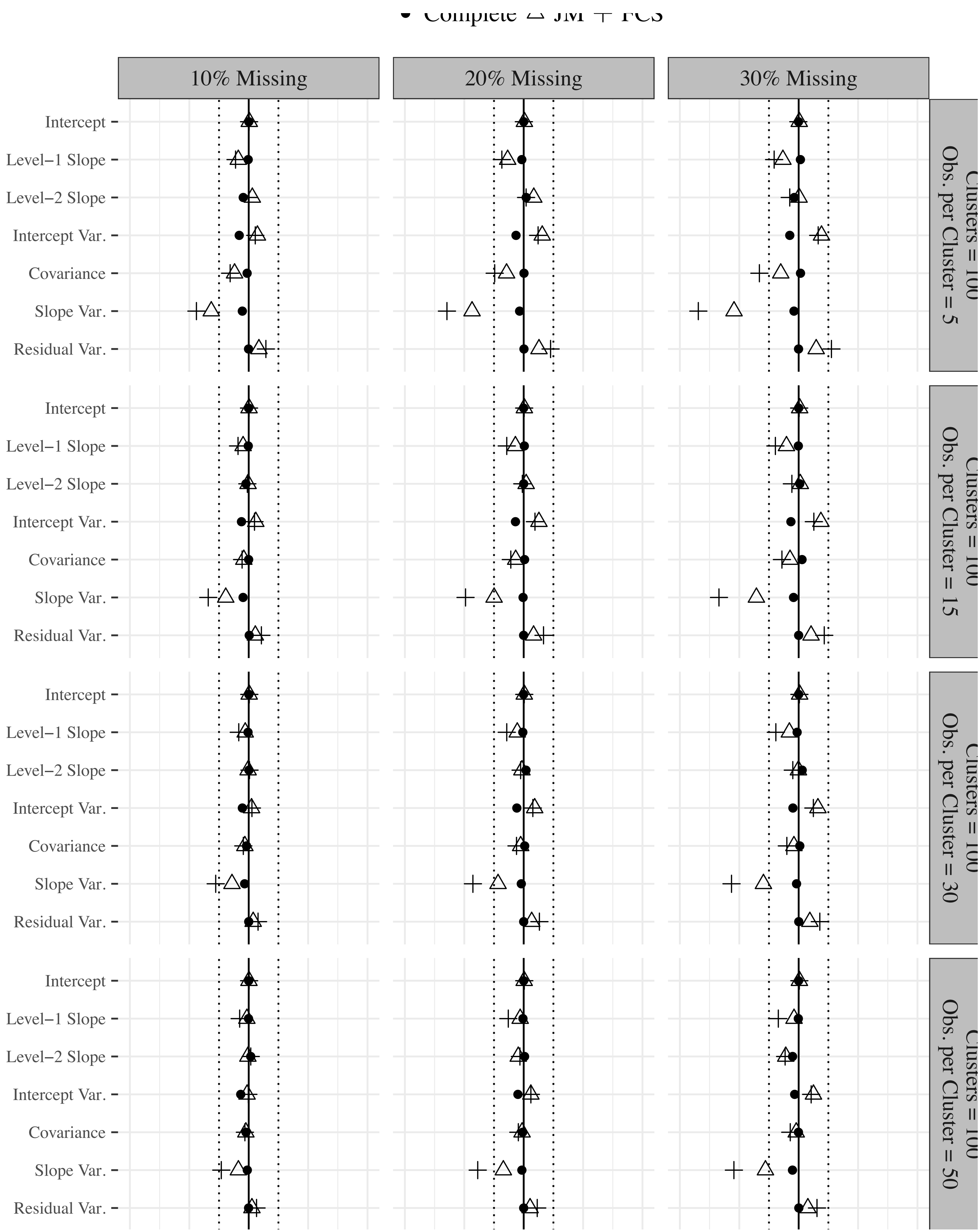


● Complete △ JM + FCS
10% Missing
20% Missing
30% Missing
Intercept
Level−1 Slope
Level−2 Slope
Intercept Var.
Covariance
Slope Var.
Residual Var.
Clusters = 100
Obs. per Cluster = 5
Obs. per Cluster = 15
Obs. per Cluster = 30
Obs. per Cluster = 50


Complete JM FCS
10% Missing
20% Missing
30% Missing
Intercept
Level−1 Slope
Level−2 Slope
Intercept Var.
Covariance
Slope Var.
Residual Var.
Clusters = 15
Obs. per Cluster = 5
Obs. per Cluster = 15
Obs. per Cluster = 30
Obs. per Cluster = 50

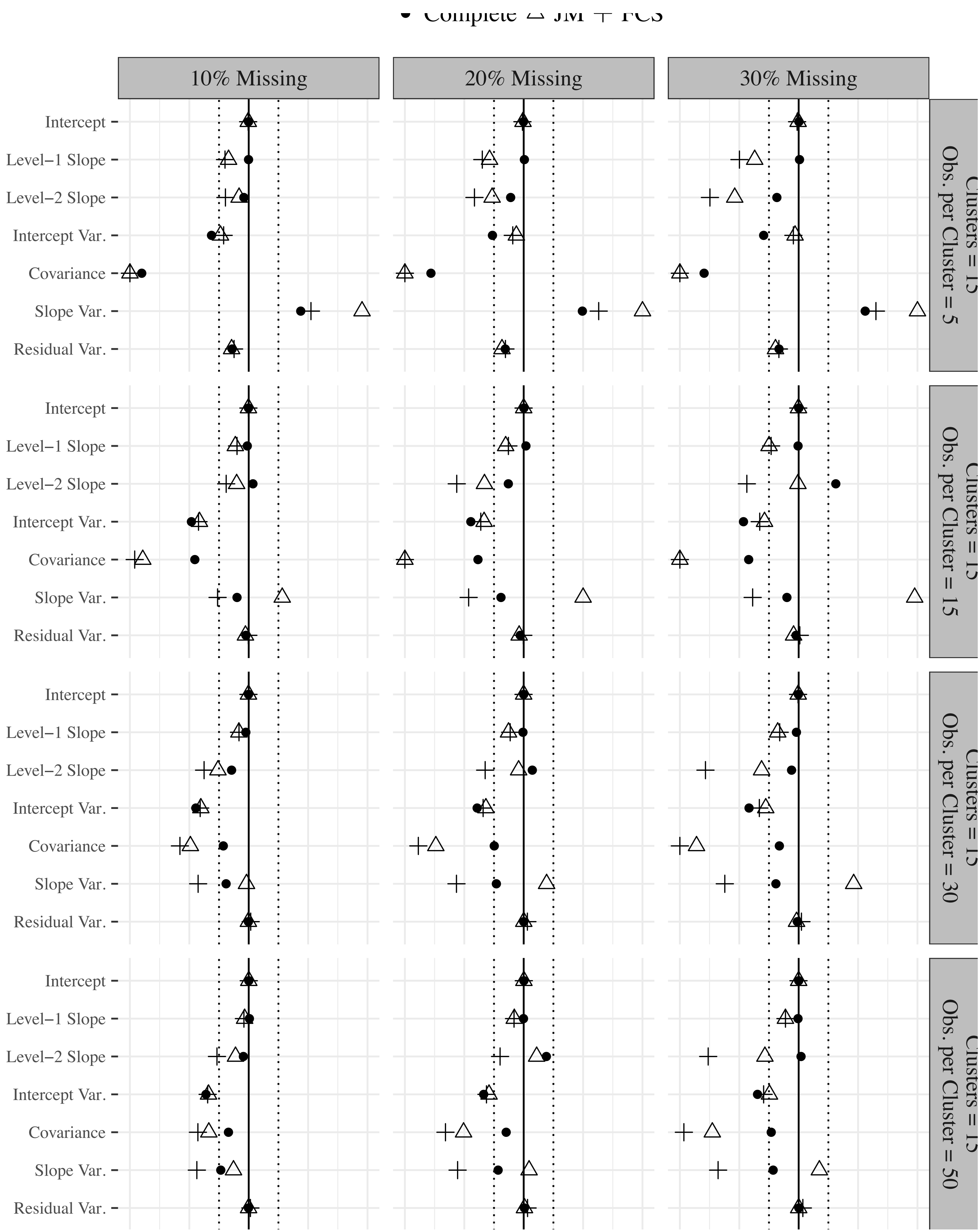


Complete JM FCS
10% Missing
20% Missing
30% Missing
Intercept
Level−1 Slope
Level−2 Slope
Intercept Var.
Covariance
Slope Var.
Residual Var.
Clusters = 30
Obs. per Cluster = 5
Obs. per Cluster = 15
Obs. per Cluster = 30
Obs. per Cluster = 50

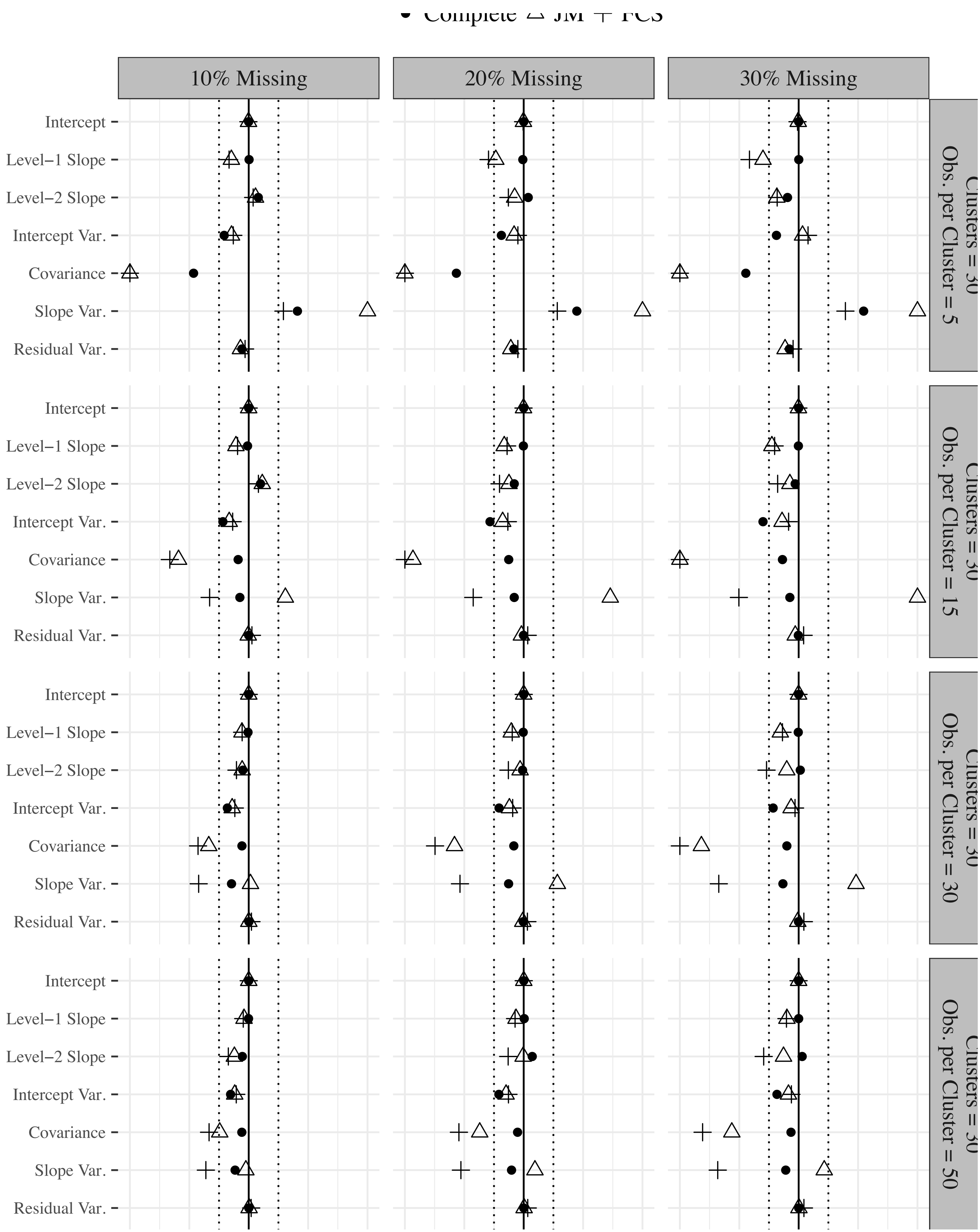

Complete JM FCS
10% Missing
20% Missing
30% Missing
Intercept
Level−1 Slope
Level−2 Slope
Intercept Var.
Covariance
Slope Var.
Residual Var.
Clusters = 100
Obs. per Cluster = 5
Obs. per Cluster = 15
Obs. per Cluster = 30
Obs. per Cluster = 50

Supplement G: Trellis Plots of Relative Bias

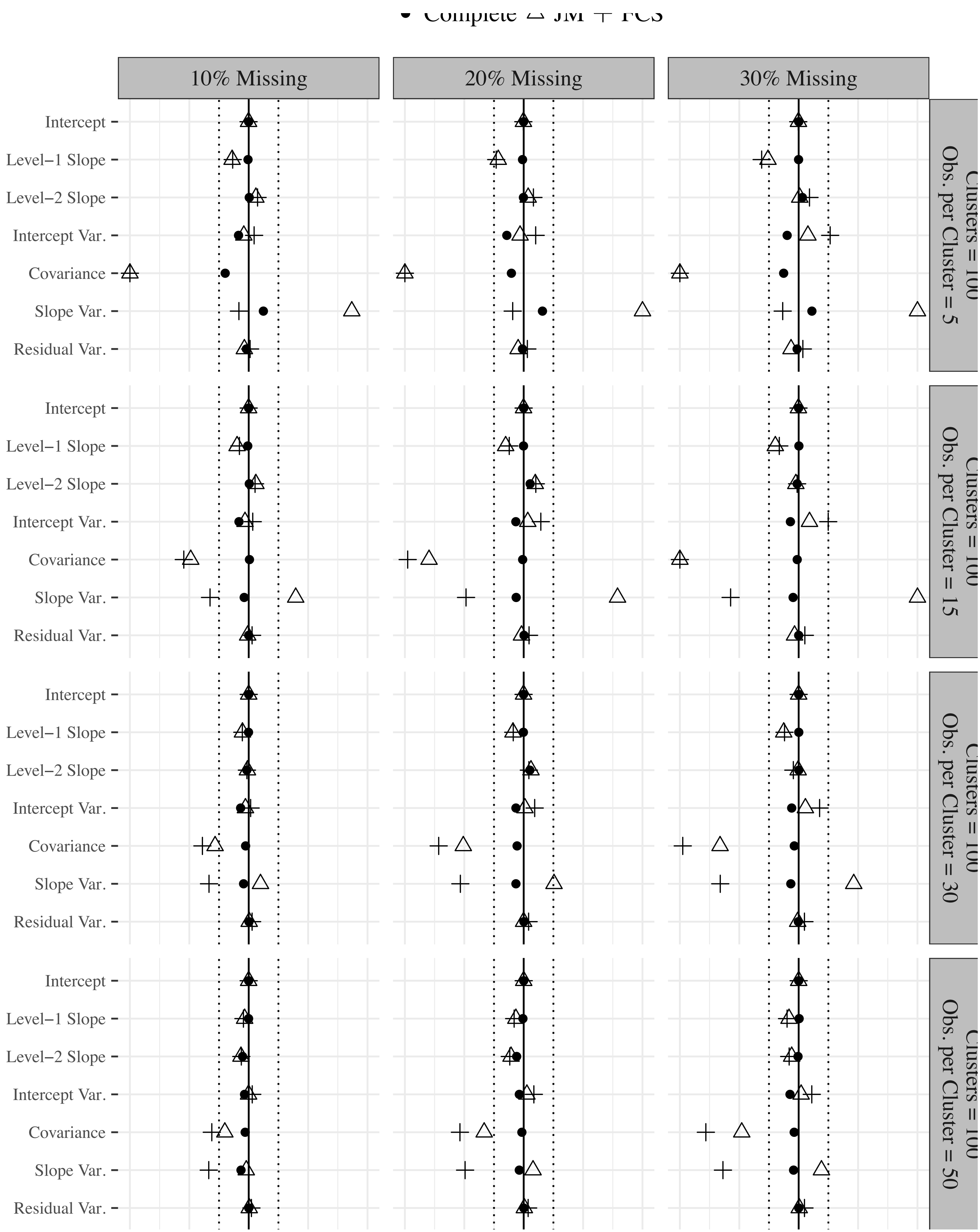
Correlated Random Effects, Outcome as the Cause of Missingness



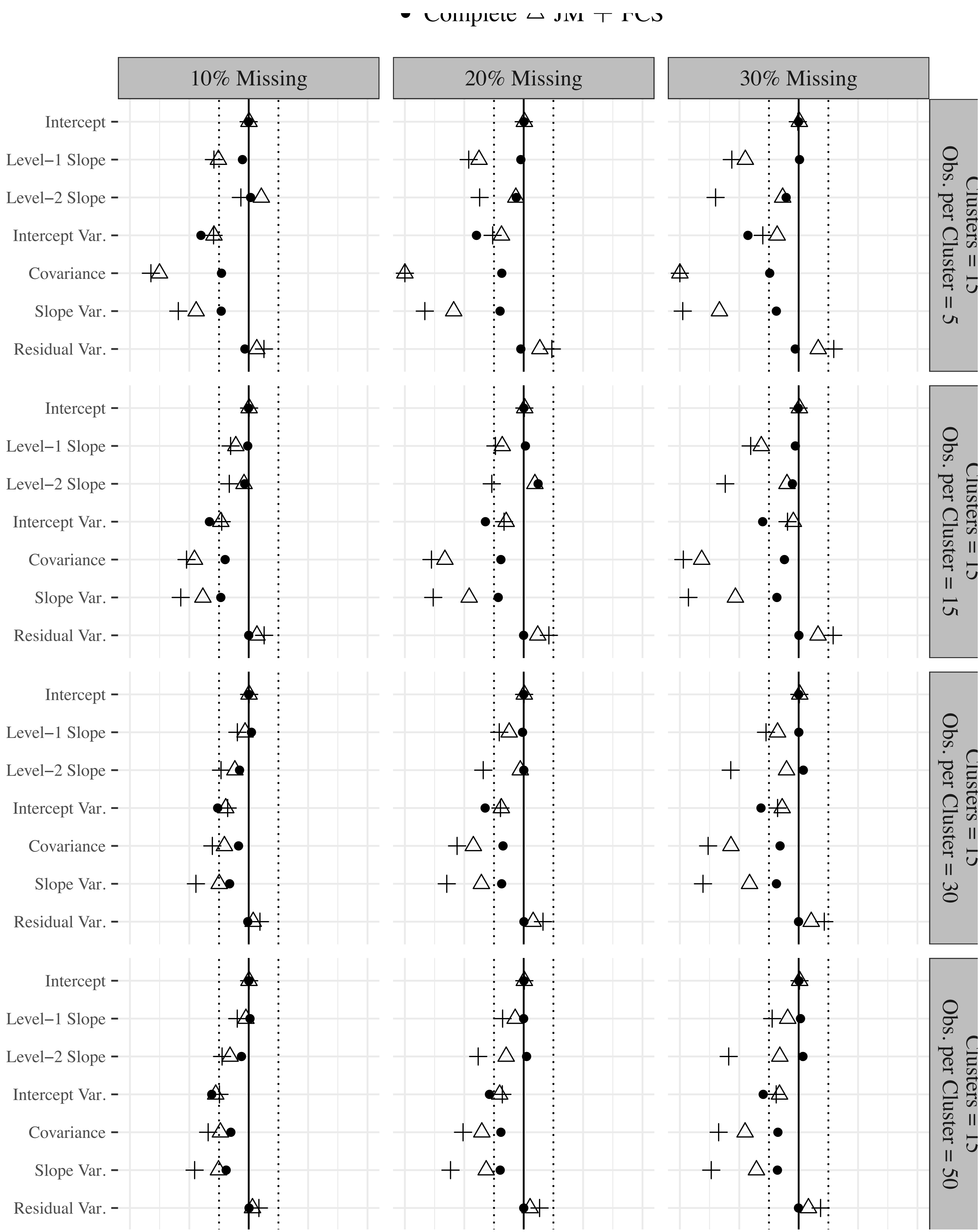
Complete JM FCS
10% Missing
20% Missing
30% Missing
Intercept
Level−1 Slope
Level−2 Slope
Intercept Var.
Covariance
Slope Var.
Residual Var.
Clusters = 15
Obs. per Cluster = 5
Obs. per Cluster = 15
Obs. per Cluster = 30
Obs. per Cluster = 50

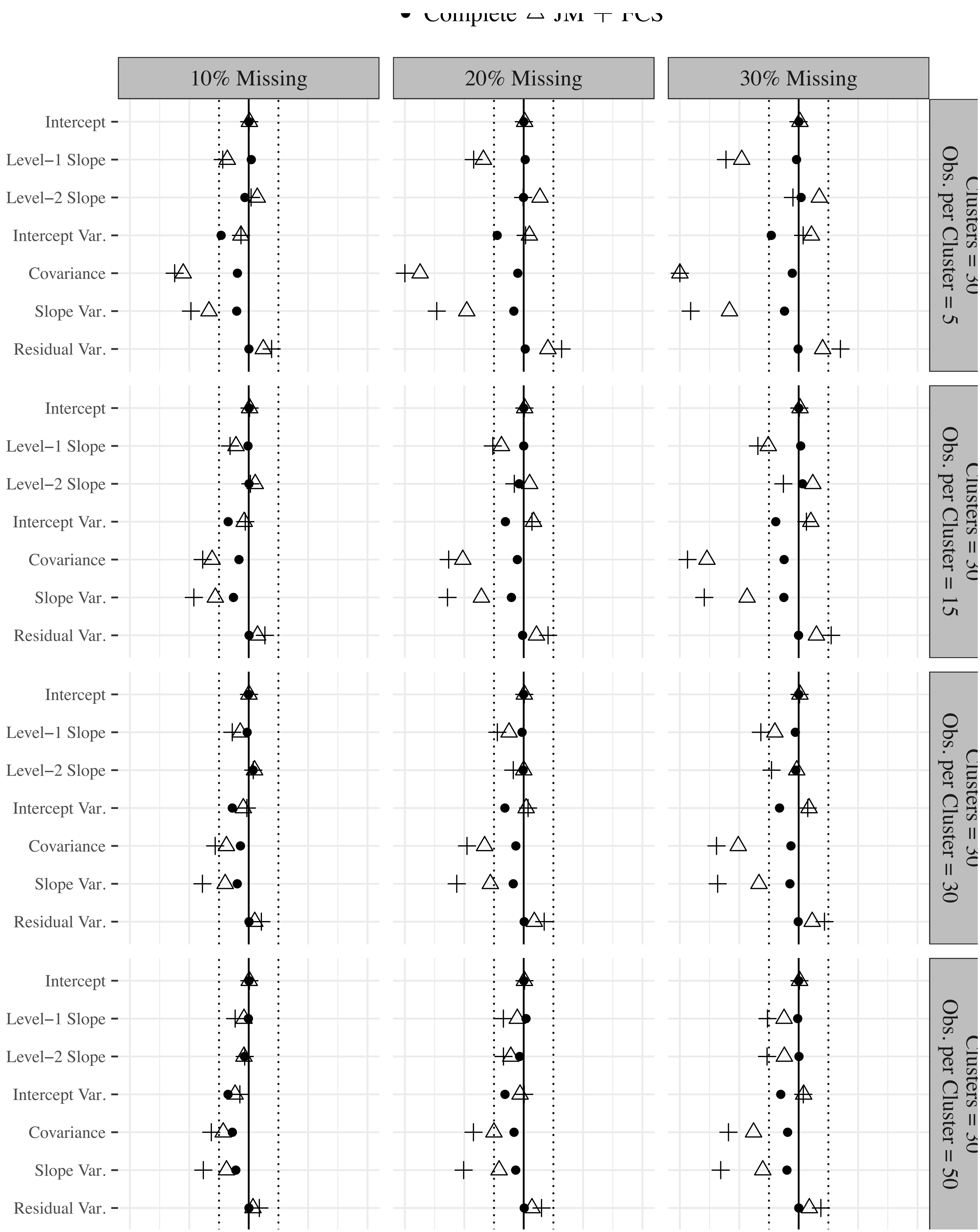


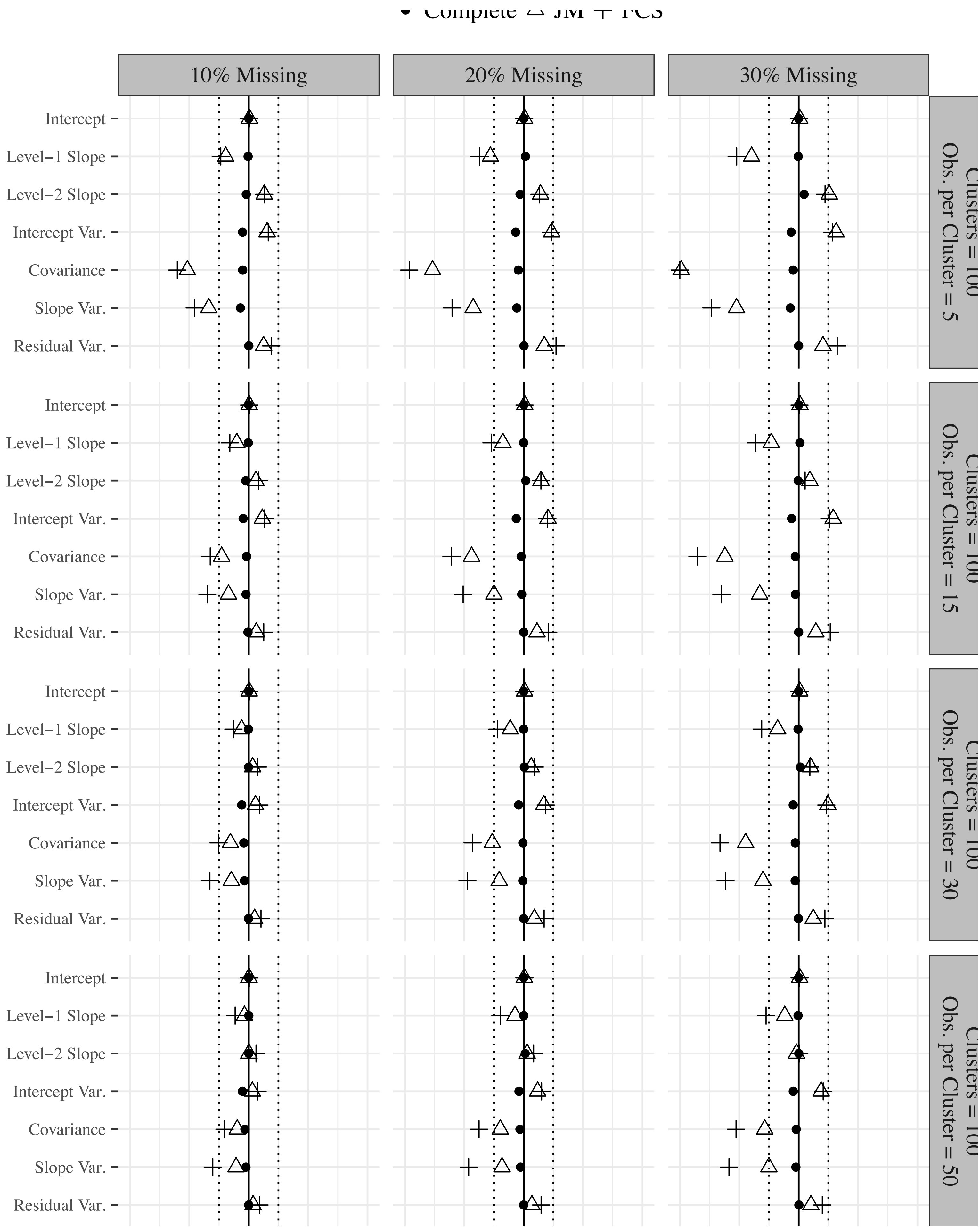
Complete JM FCS
10% Missing
20% Missing
30% Missing
Intercept
Level−1 Slope
Level−2 Slope
Intercept Var.
Covariance
Slope Var.
Residual Var.
Clusters = 30
Obs. per Cluster = 5
Obs. per Cluster = 15
Obs. per Cluster = 30
Obs. per Cluster = 50



Complete JM FCS
10% Missing
20% Missing
30% Missing
Intercept
Level−1 Slope
Level−2 Slope
Intercept Var.
Covariance
Slope Var.
Residual Var.
Clusters = 100
Obs. per Cluster = 5
Obs. per Cluster = 15
Obs. per Cluster = 30
Obs. per Cluster = 50

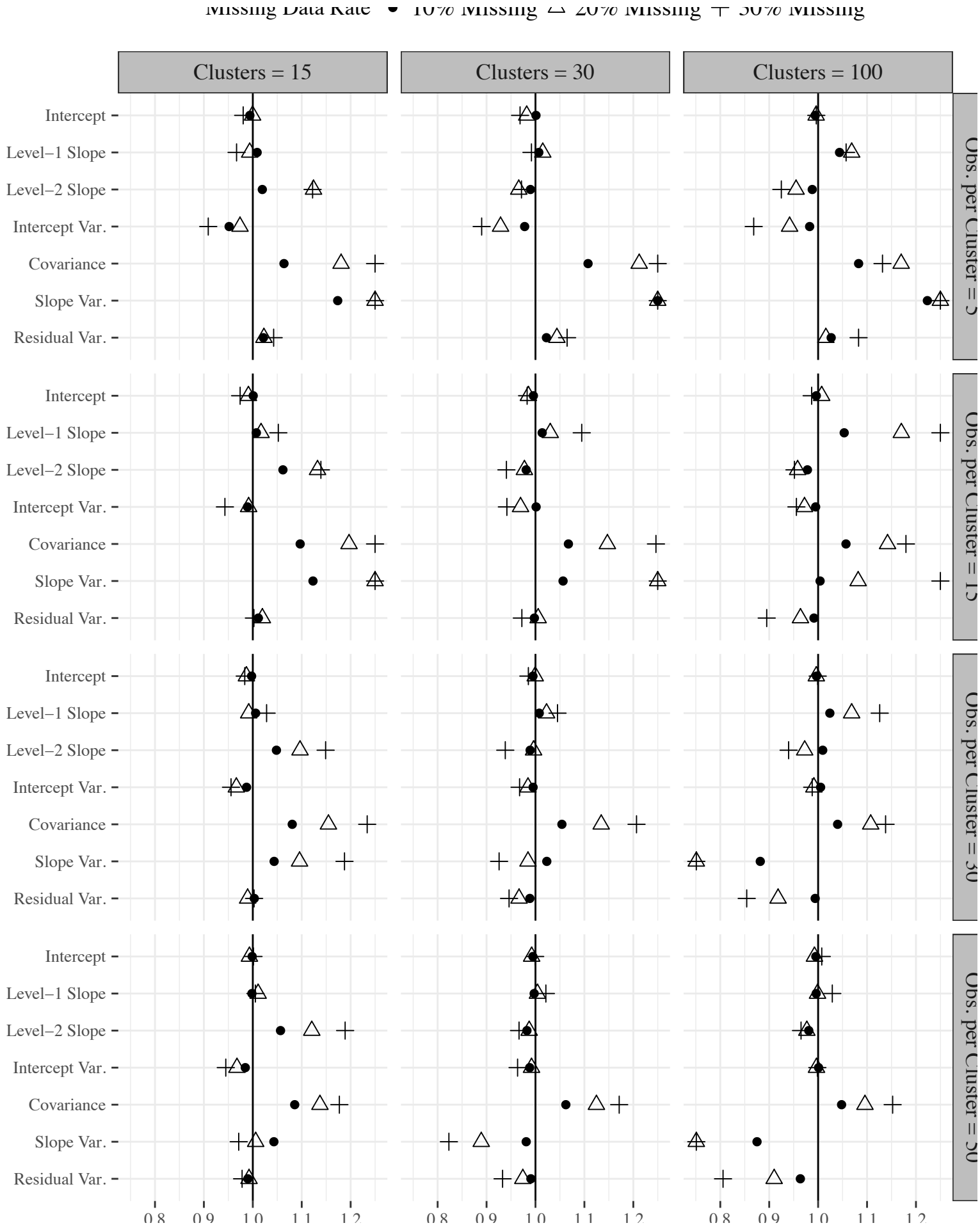

Complete JM FCS
10% Missing
20% Missing
30% Missing
Intercept
Level−1 Slope
Level−2 Slope
Intercept Var.
Covariance
Slope Var.
Residual Var.
Clusters = 15
Obs. per Cluster = 5
Obs. per Cluster = 15
Obs. per Cluster = 30
Obs. per Cluster = 50

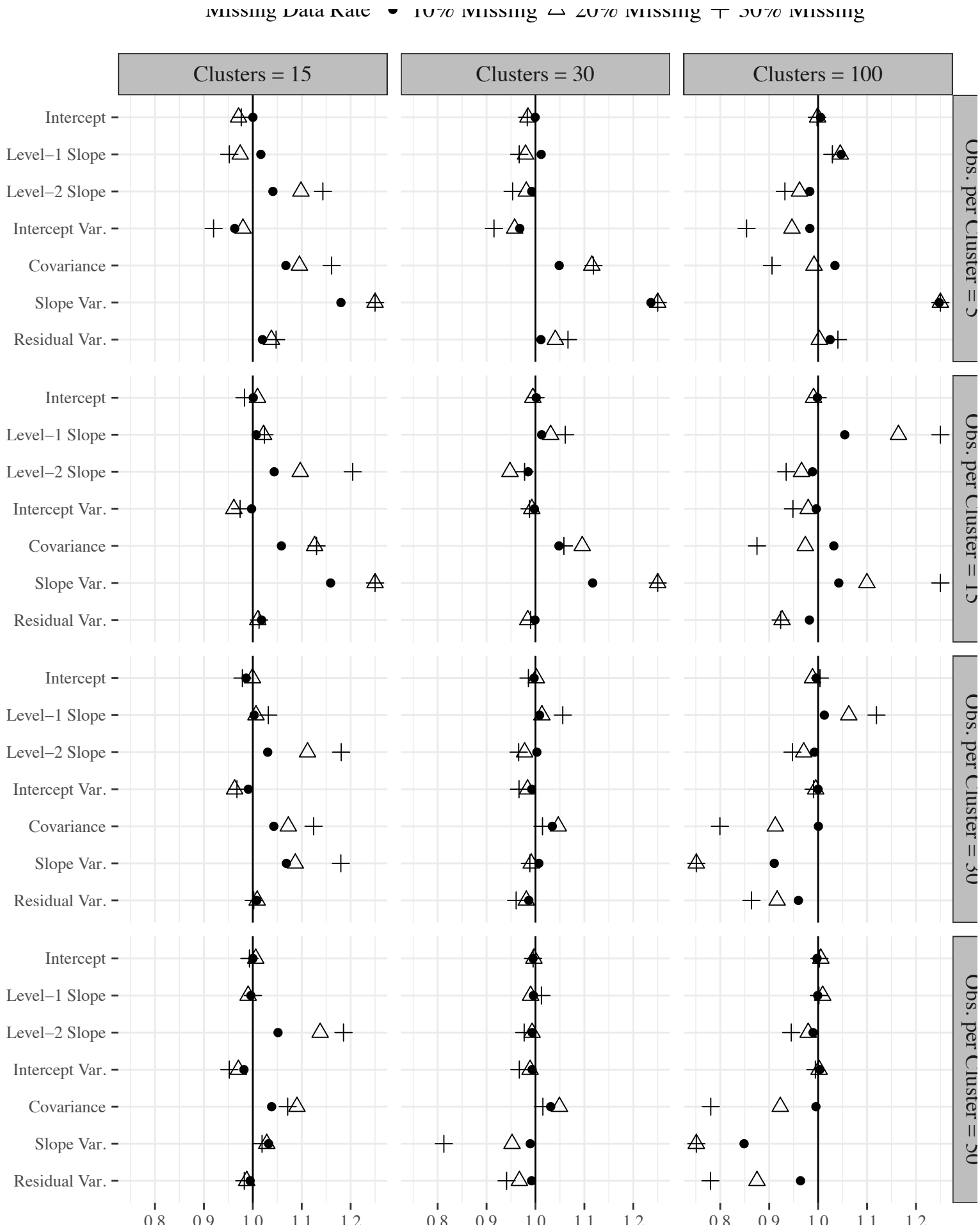

Complete
JM
FCS
10% Missing
20% Missing
30% Missing
Intercept
Level−1 Slope
Level−2 Slope
Intercept Var.
Covariance
Slope Var.
Residual Var.
Clusters = 30
Obs. per Cluster = 5
Obs. per Cluster = 15
Obs. per Cluster = 30
Obs. per Cluster = 50

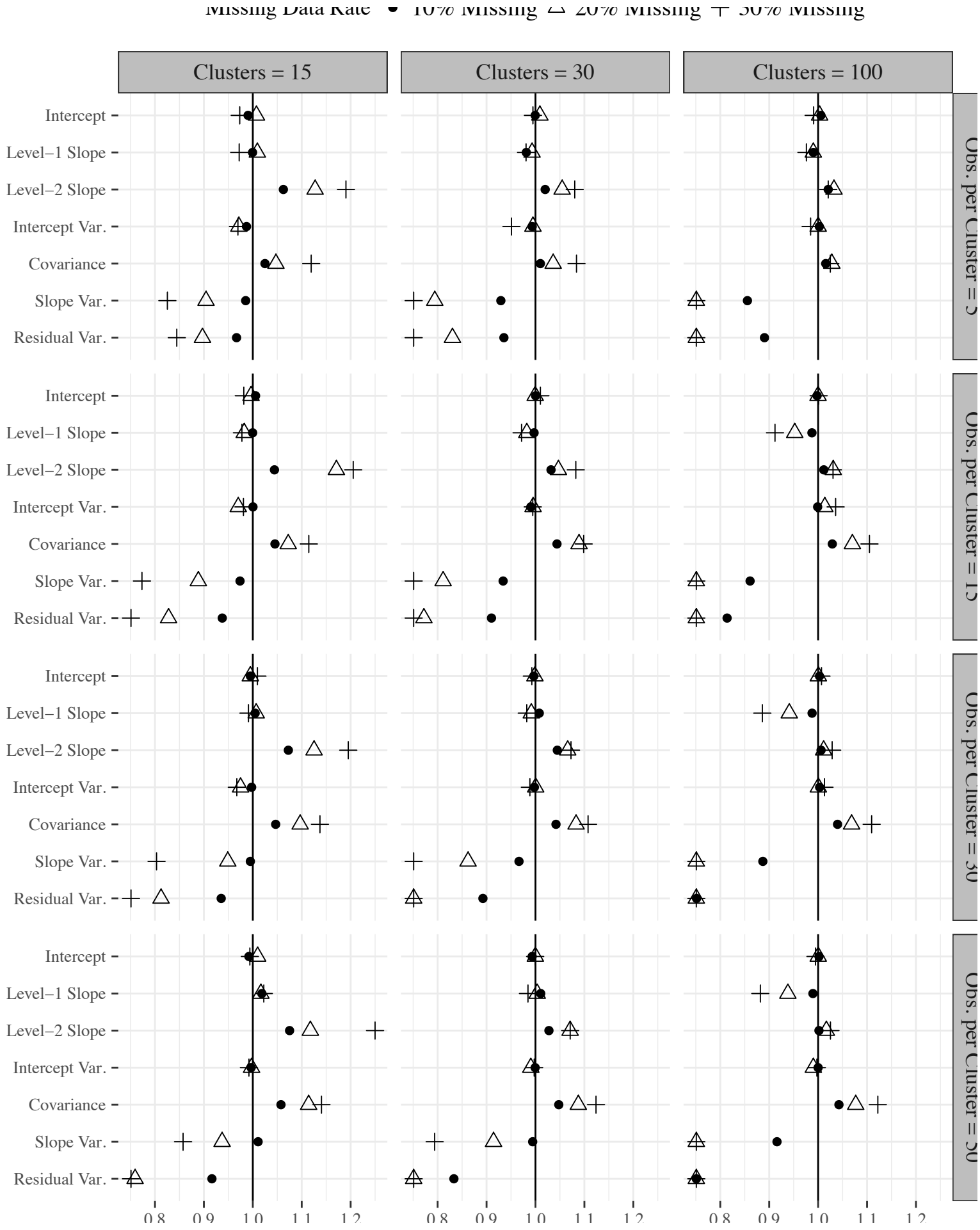


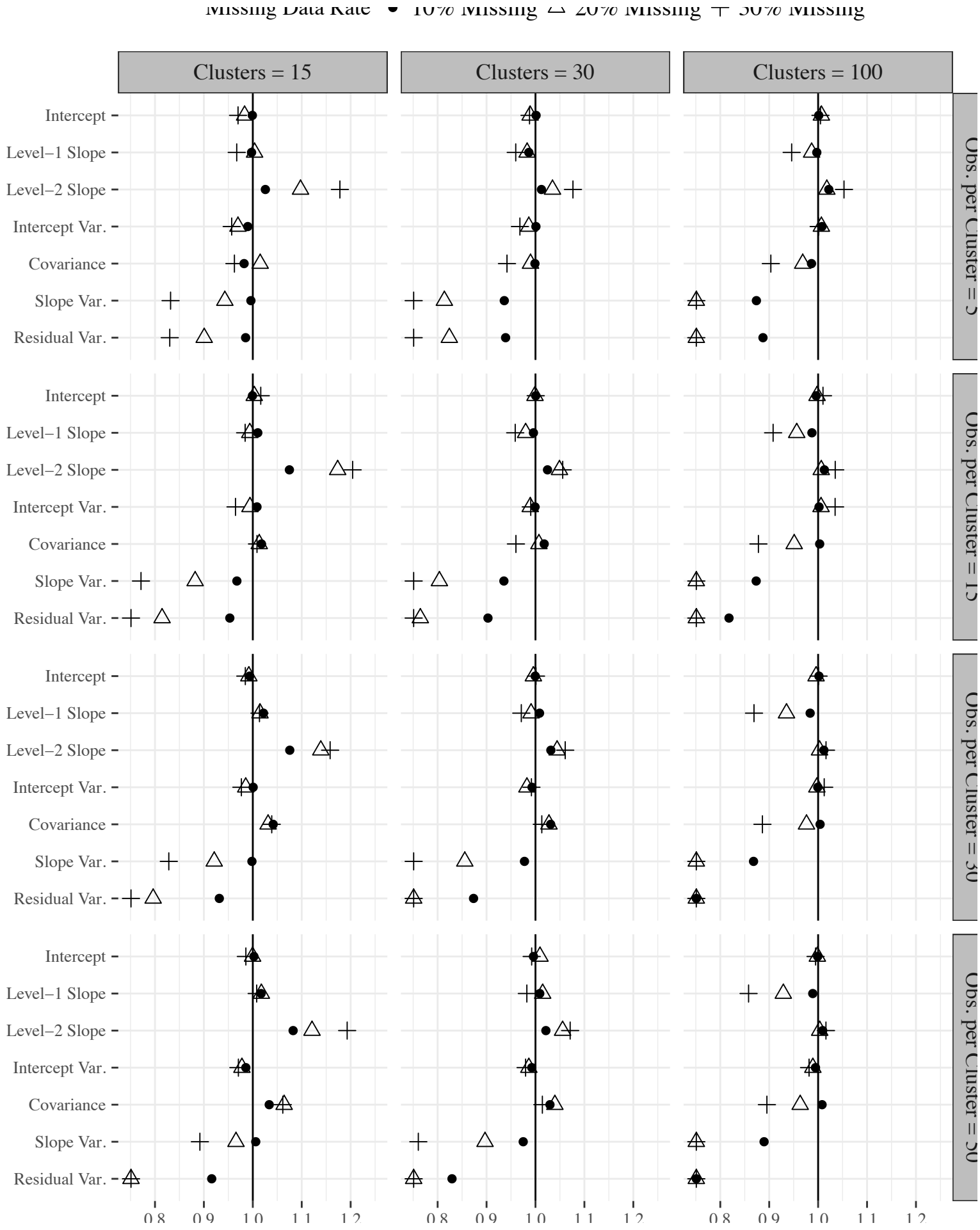
Complete JM FCS
10% Missing
20% Missing
30% Missing
Intercept
Level−1 Slope
Level−2 Slope
Intercept Var.
Covariance
Slope Var.
Residual Var.
Clusters = 100
Obs. per Cluster = 5
Obs. per Cluster = 15
Obs. per Cluster = 30
Obs. per Cluster = 50

Supplement H: Trellis Plots of Mean Squared Error Ratios


Missing Data Rate
10% Missing
20% Missing
30% Missing
Clusters = 15
Clusters = 30
Clusters = 100
Obs. per Cluster = 5
Obs. per Cluster = 15
Obs. per Cluster = 30
Obs. per Cluster = 50
Intercept
Level−1 Slope
Level−2 Slope
Intercept Var.
Covariance
Slope Var.
Residual Var.
0.8
0.9
1.0
1.1
1.2


Missing Data Rate
10% Missing
20% Missing
30% Missing
Clusters = 15
Clusters = 30
Clusters = 100
Obs. per Cluster = 5
Obs. per Cluster = 15
Obs. per Cluster = 30
Obs. per Cluster = 50
Intercept
Level−1 Slope
Level−2 Slope
Intercept Var.
Covariance
Slope Var.
Residual Var.
0.8
0.9
1.0
1.1
1.2


Missing Data Rate
10% Missing
20% Missing
30% Missing
Clusters = 15
Clusters = 30
Clusters = 100
Obs. per Cluster = 5
Obs. per Cluster = 15
Obs. per Cluster = 30
Obs. per Cluster = 50
Intercept
Level−1 Slope
Level−2 Slope
Intercept Var.
Covariance
Slope Var.
Residual Var.
0.8
0.9
1.0
1.1
1.2


Missing Data Rate
10% Missing
20% Missing
30% Missing
Clusters = 15
Clusters = 30
Clusters = 100
Obs. per Cluster = 5
Obs. per Cluster = 15
Obs. per Cluster = 30
Obs. per Cluster = 50
Intercept
Level−1 Slope
Level−2 Slope
Intercept Var.
Covariance
Slope Var.
Residual Var.
0.8
0.9
1.0
1.1
1.2

Supplement I: Blimp Syntax for Substantive Model-Compatible Imputation

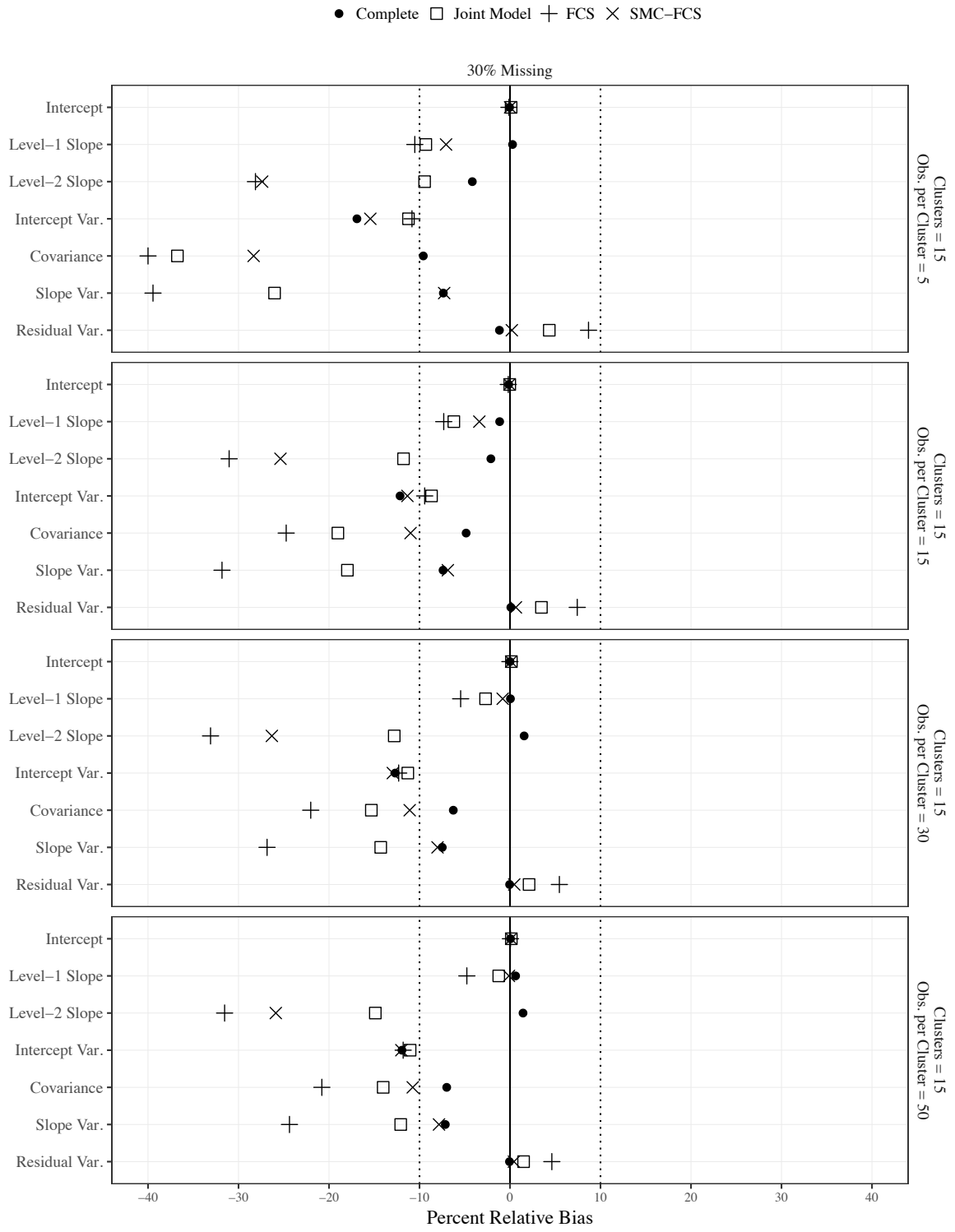
The syntax below applies substantive model-compatible imputation from Enders et al. (2018) to an artificial data set from the simulations reported in the body of the paper. Additional information about Blimp code is available in the user's guide (Keller & Enders, 2018).

```
DATA: ~/desktop/data.csv;
VARNAMES: id a1 a2 y x w;
! specify random slope with y:x;
MODEL: id ~ a1 a2 y:x w;
! specify outcome variable from the substantive analysis;
OUTCOME: y;
BURN: 2500;
THIN: 2500;
NIMPS: 20;
MISSING: 999;
SEED: 90291;
OUTFILE: ~/desktop/imp*.csv;
OPTIONS: separate;
```

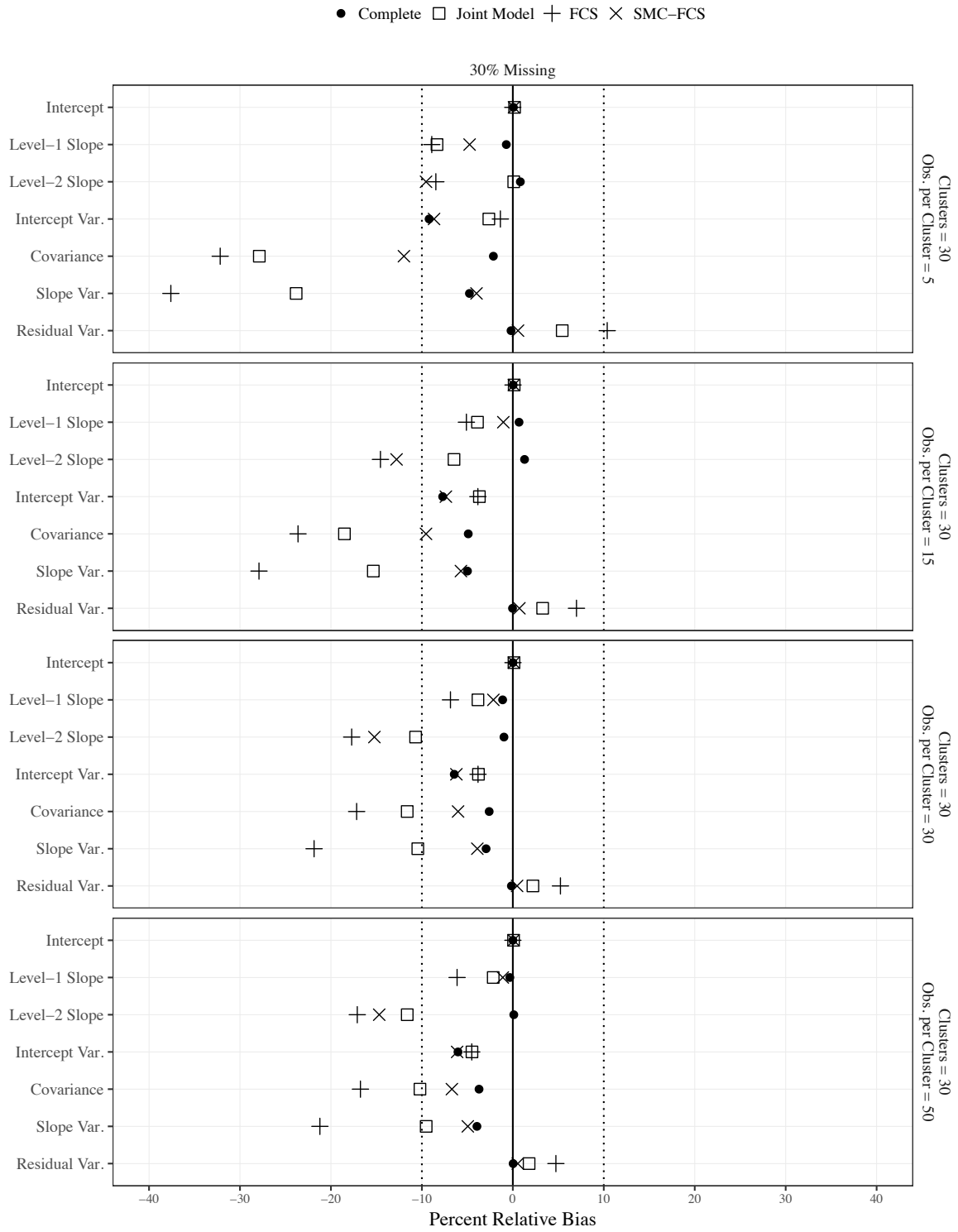
Supplement J: Simulation Results Evaluating Substantive Model-Compatible Imputation

To provide a preliminary evaluation of substantive model-compatible fully conditional specification (SMC-FCS), we applied the procedure to the subset of simulation conditions with an ICC = .50 and correlated random effects. The trellis plots display relative bias values for this new procedure as well as the conventional imputation approaches evaluated in the manuscript. Additional details on SMC-FCS can be found in Enders et al. (2018).

Supplemental Figure 1. Average relative bias values for design cells with ICC = .50 and 15 clusters. Relative bias is defined as the difference between an average estimate and the true value expressed as a proportion of the true value. The dashed lines represent bias values of ± 0.10.



Supplemental Figure 2. Average relative bias values for design cells with ICC = .50 and 30 clusters. Relative bias is defined as the difference between an average estimate and the true value expressed as a proportion of the true value. The dashed lines represent bias values of ± 0.10.



Supplemental Figure 3. Average relative bias values for design cells with ICC = .50 and 100 clusters. Relative bias is defined as the difference between an average estimate and the true value expressed as a proportion of the true value. The dashed lines represent bias values of ± 0.10.