Tech 2, Descriptives

### Data

The datasets in thie file were worked up in the project file [01\_CRI\_Data\_Prep.Rm](https://github.com/nbreznau/CRI/blob/master/01_CRI_Data_Prep.Rmd). In the [CRI.Rproj](https://github.com/nbreznau/CRI/blob/master/CRI.Rproj). Project and data available via GitHub.

1. *cri.csv* = model-level data, numerical values
2. *cri\_team.csv* = team-level data

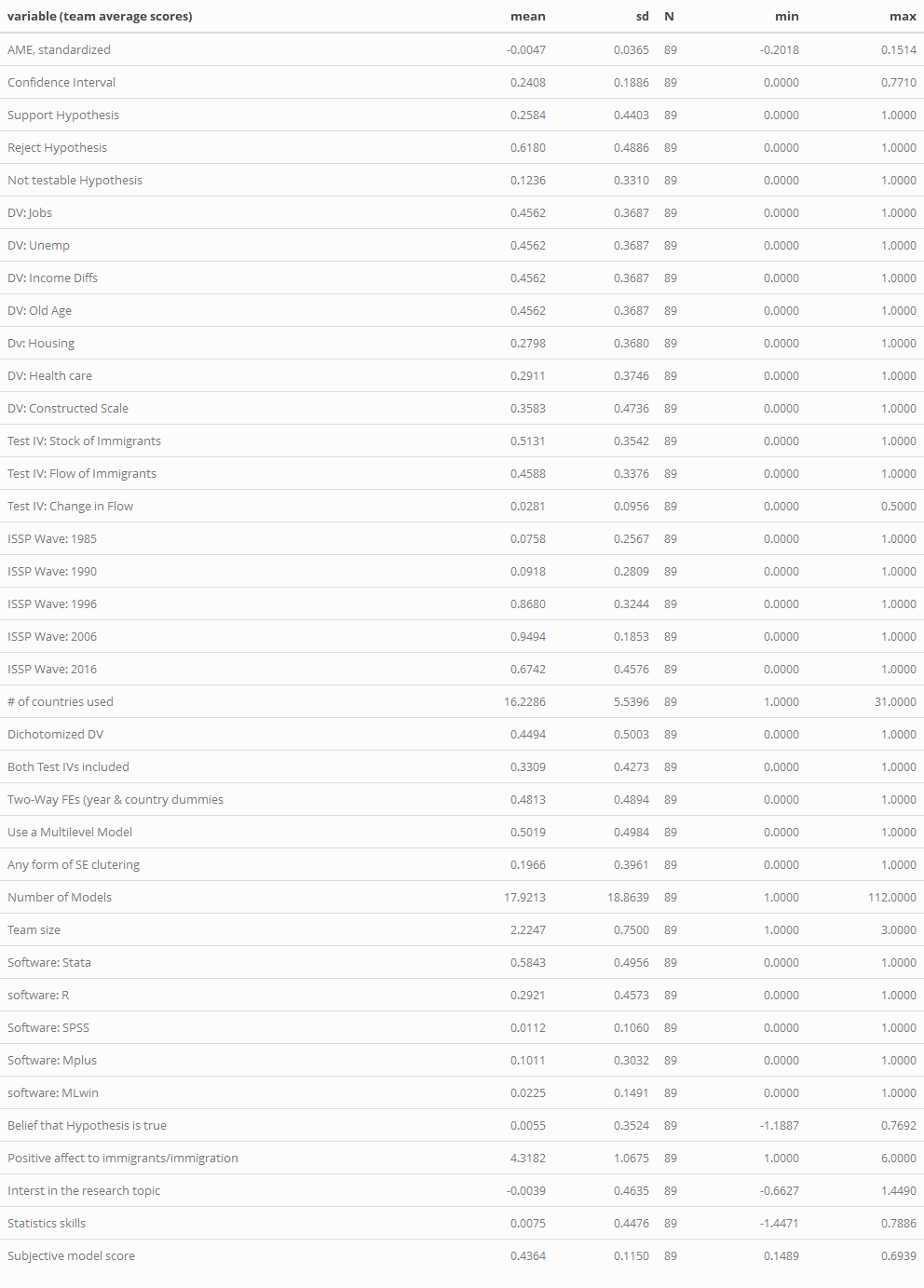
### Descriptives

There were 196 participants in 73 teams that produced 1261 models.

There were 16 teams that came to more than one conclusion. For example, finding support of the hypothesis for immigration stock models, but finding rejection of the hypothesis for immigration flow models. Therefore, we have 89 observations (‘results’) in the team-level data that derive from 72 teams plus we include the conclusion from the study by Brooks and Manza ([2006](https://doi.org/10.1177/000312240607100306)) that provides an example of a state of the art study for testing this hypothesis.

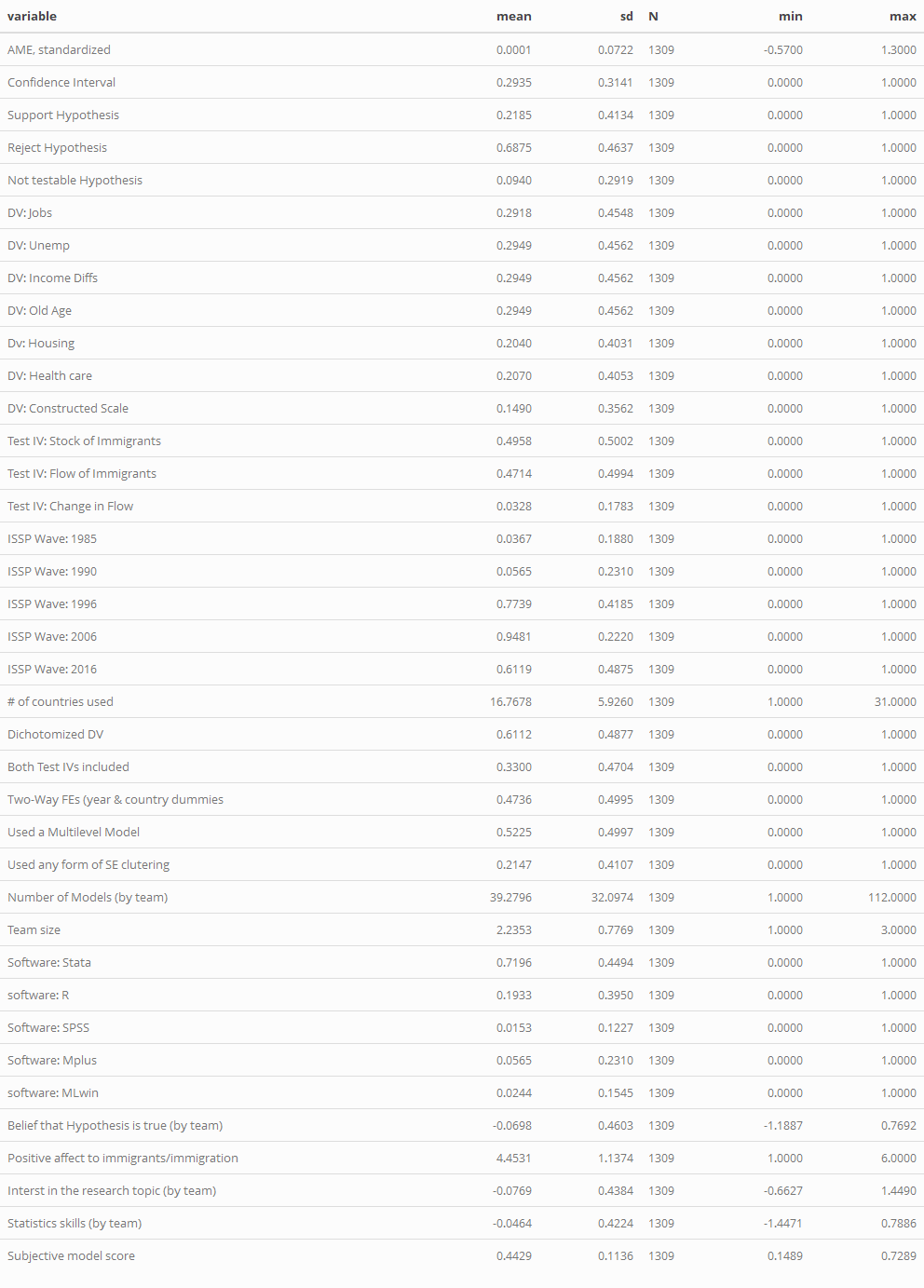
#### Table 1. Team-Results-Level Descriptive Statistics

knitr::include\_graphics("results/Table1.png")



#### Table 2. Model-Level Descriptive Statistics

knitr::include\_graphics("results/Table2.png")



### Main Findings - Tables 1-3

# Inlcude team 105 in the upper panel (all teams). They conducted measurement models and determined it could not (should not) be tested. Team 1 included as well, they had non-convergence.

This chunk preps some of the stats for Tables 1-3, these tables will be completed in 05\_CRI\_Main\_Analyses.Rmd

# remove original study and non-conclusive results  
cri0 <- subset(cri, u\_teamid != 0)  
cri\_team0 <- subset(cri\_team, u\_teamid != 0 & u\_teamid != 1)  
  
tm\_avg <- round(mean(cri\_team0$inv\_weight, na.rm=T),2)  
tm\_sd <- round(sd(cri\_team0$inv\_weight, na.rm=T),2)  
tm\_min <- min(cri\_team0$inv\_weight, na.rm=T)  
tm\_max <- max(cri\_team0$inv\_weight, na.rm=T)

This chunk creates the matrix and saves it for export to

# create table frame  
tbl1 <- matrix(nrow = 14, ncol = 3)  
tbl1[1:14,1] <- c("","","Test Results","Positive","Insignificant","Negative","(test failed)", "TOTAL","Sujective Conclusion: Hypothesis is?","Supported", "Not testable/inconclusive", "Rejected", "TOTAL","")  
tbl1[2,2] <- paste0("(out of ",length(unique(cri\_team0$u\_teamid))," teams)")  
tbl1[2,3] <- paste0("(out of ",length(cri0$AME)," models)")  
tbl1[1,2:3] <- c("Average Rate","Rate[a]")  
tbl1[14,1] <- paste0("Of the ", length(unique(cri\_team0$u\_teamid)), " teams in the CRI,", sum(duplicated(cri\_team0$u\_teamid)), " of them treated stock and flow measures as independent tests of the hypothesis. Therefore, there are ", length(cri\_team0$u\_teamid), " team-level observed tests, each with an independent subjective conclusion. Of these ", length(cri\_team0$u\_teamid), " team-level results there was an average of ", tm\_avg, " test models per team [sd =", tm\_sd, ", min =", tm\_min, ", max =", tm\_max, "].")  
  
# fill in descriptive results  
  
# need to calculate rate carefully because teams 1 and 105 had failed tests  
  
tbl1[4,2] <- mean(cri\_team0$pos\_test\_pct\_p05, na.rm = T)  
tbl1[4,3] <- mean(cri0$AME\_sup\_p05, na.rm = T)  
tbl1[5,2] <- mean(cri\_team0$ns\_test\_pct\_p05, na.rm = T)  
tbl1[5,3] <- mean(cri0$AME\_ns\_p05, na.rm = T)  
tbl1[6,2] <- mean(cri\_team0$neg\_test\_pct\_p05, na.rm = T)  
tbl1[6,3] <- mean(cri0$AME\_neg\_p05, na.rm = T)  
tbl1[7,2] <- 1- (sum(as.numeric(tbl1[4:6,2]), na.rm = T))  
tbl1[7,3] <- 1- (sum(as.numeric(tbl1[4:6,3]), na.rm = T))  
  
tbl1[8,2:3] <- 1  
  
tbl1[10,2] <- (mean(cri\_team0$Hsup, na.rm = T)\*87)/89  
tbl1[11,2] <- ((mean(cri\_team0$Hnotest, na.rm = T)\*87)+2)/89  
tbl1[12,2] <- (mean(cri\_team0$Hrej, na.rm = T)\*87)/89  
tbl1[13,2] <- 1  
  
write.csv(tbl1, file = "results/tbl1.csv")

#### Correlation heatmap for Tbl1 obj and subj results

cor\_team <- select(cri\_team0, Hsup, Hno, Hrej, pos\_test\_pct\_p05, ns\_test\_pct\_p05, neg\_test\_pct\_p05)  
cormat <- round(cor(cor\_team),2)  
  
sq\_cormat <- cormat[-c(4:6),-c(1:3)]  
colnames(sq\_cormat) <- c("Positive","Not Sig.","Negative")  
rownames(sq\_cormat) <- c("Supported","Inconclusive","Rejected")  
sq\_cormat\_melted <- melt(sq\_cormat)  
  
agg\_png(filename = "results/Tbl1\_cor.png", res = 144, height = 600, width = 700)  
ggplot(data = sq\_cormat\_melted, aes(X1, X2, fill = value))+  
 geom\_tile(color = "white")+  
 scale\_fill\_gradient2(low = "maroon", high = "green", mid = "grey",   
 midpoint = 0, limit = c(-0.68,0.68), space = "Lab",   
 name="Pearson\nCorrelation") +  
 theme\_minimal() +  
 coord\_fixed() +  
 xlab("Regression Coefficients, 95% CI\n(percentage by team)") +  
 ylab("Subjective Team Conclusion\nabout Hypothesis Test") +  
 geom\_text(aes(X1, X2, label = value), color = "black", size = 4) +  
 theme(panel.grid.major = element\_blank(),  
 axis.title.x = element\_text(vjust = -0.8),  
 panel.border = element\_blank(),  
 panel.background = element\_blank(),  
 axis.ticks = element\_blank())  
dev.off()

## png   
## 2

knitr::include\_graphics("results/Tbl1\_cor.png")

