

Brier score

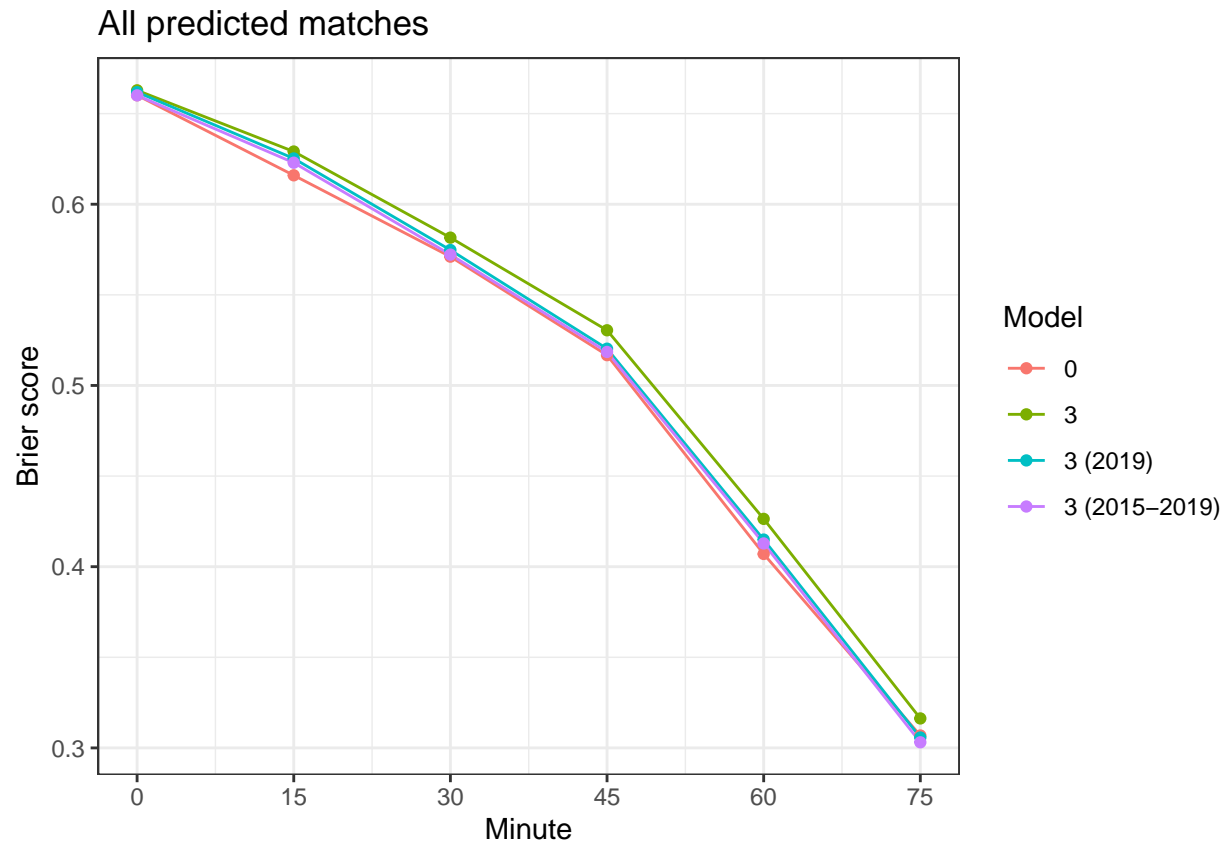
```
library(dplyr)
library(ggplot2)

load("data/HDA.RData")

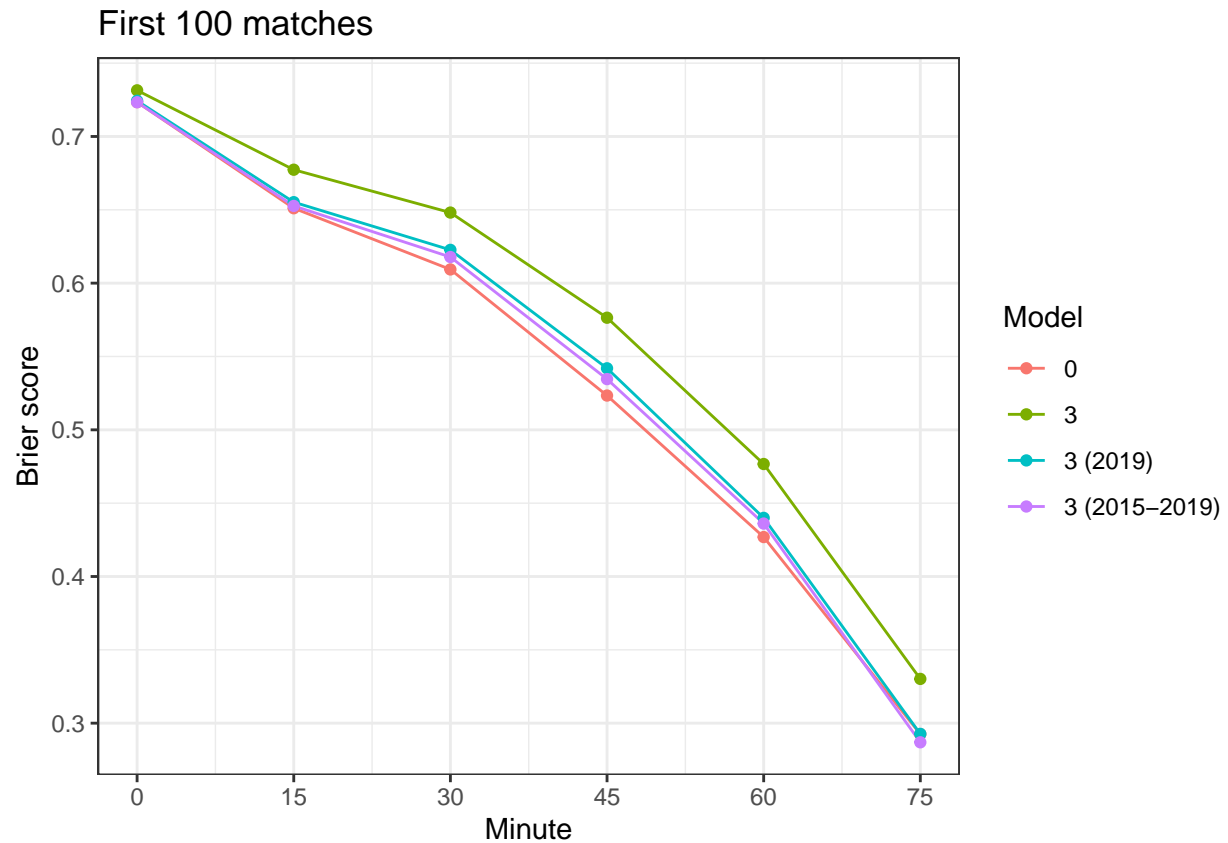
nrow(HDA)
```

```
## [1] 350
```

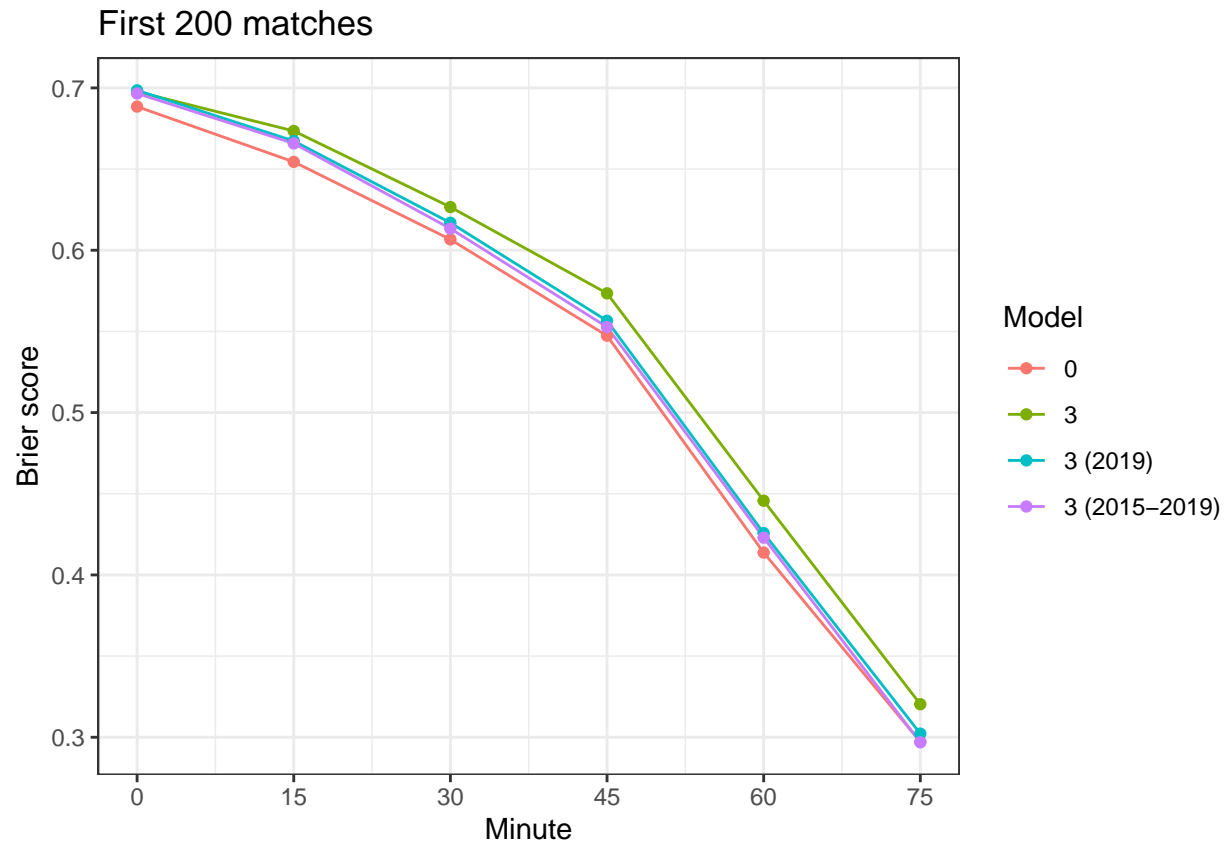
```
tibble(Brier = apply(HDA[,c(105:128)], 2, mean),
       Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 4)),
       Model = factor(c(rep("0", 6), rep("3", 6), rep("3 (2019)", 6),
                        rep("3 (2015-2019)", 6)),
                     levels = c("0", "3", "3 (2019)", "3 (2015-2019)"))) %>%
  ggplot(aes(x = Minute, y = Brier, col = Model)) +
  geom_line() +
  geom_point() +
  scale_x_continuous(breaks = c(0, 15, 30, 45, 60, 75)) +
  theme_bw() +
  ggtitle("All predicted matches") +
  ylab("Brier score")
```



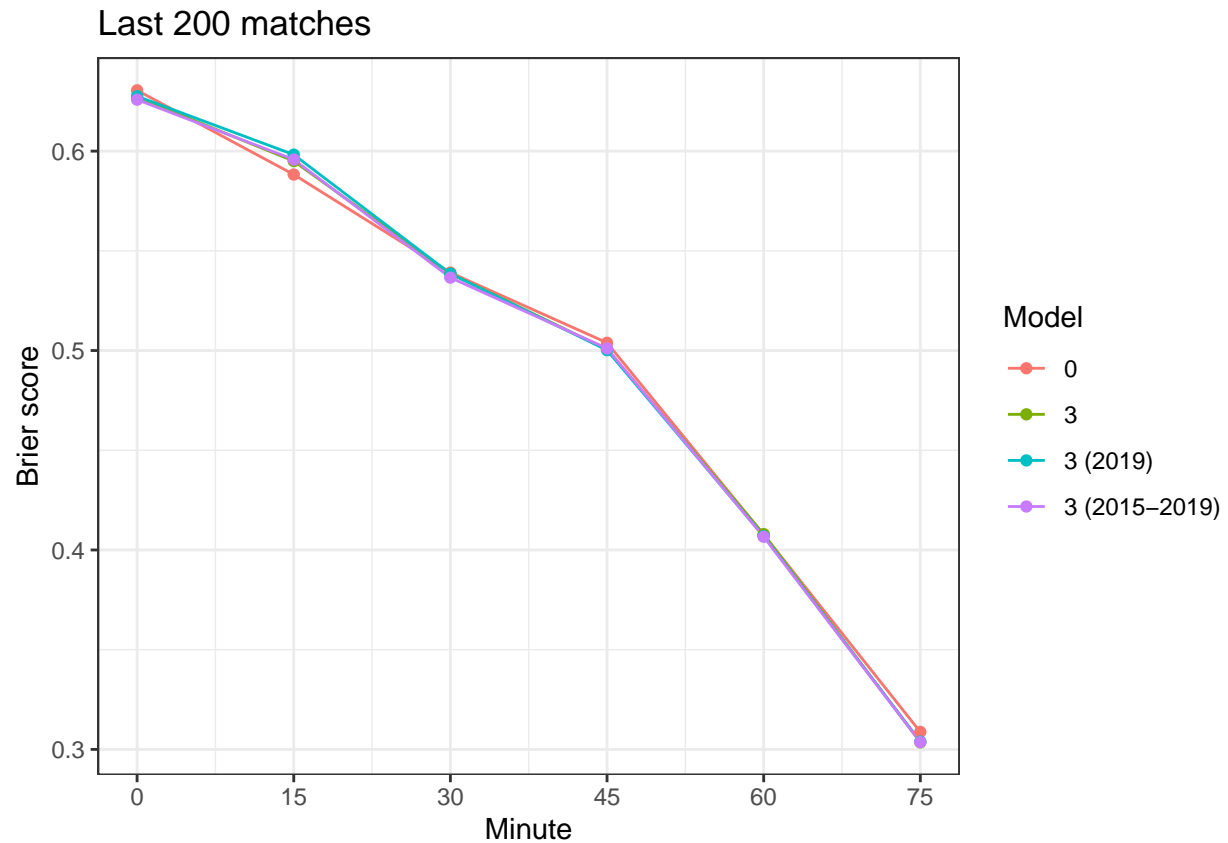
```
tibble(Brier = apply(HDA[c(1:100),c(105:128)], 2, mean),
  Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 4)),
  Model = factor(c(rep("0", 6), rep("3", 6), rep("3 (2019)", 6),
    rep("3 (2015-2019)", 6)),
    levels = c("0", "3", "3 (2019)", "3 (2015-2019)"))) %>%
  ggplot(aes(x = Minute, y = Brier, col = Model)) +
  geom_line() +
  geom_point() +
  scale_x_continuous(breaks = c(0, 15, 30, 45, 60, 75)) +
  theme_bw() +
  ggtitle("First 100 matches") +
  ylab("Brier score")
```



```
tibble(Brier = apply(HDA[c(1:200),c(105:128)], 2, mean),
  Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 4)),
  Model = factor(c(rep("0", 6), rep("3", 6), rep("3 (2019)", 6),
    rep("3 (2015-2019)", 6)),
    levels = c("0", "3", "3 (2019)", "3 (2015-2019)"))) %>%
  ggplot(aes(x = Minute, y = Brier, col = Model)) +
  geom_line() +
  geom_point() +
  scale_x_continuous(breaks = c(0, 15, 30, 45, 60, 75)) +
  theme_bw() +
  ggtitle("First 200 matches") +
  ylab("Brier score")
```



```
tibble(Brier = apply(HDA[c(151:350),c(105:128)], 2, mean),
  Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 4)),
  Model = factor(c(rep("0", 6), rep("3", 6), rep("3 (2019)", 6),
    rep("3 (2015-2019)", 6)),
    levels = c("0", "3", "3 (2019)", "3 (2015-2019)"))) %>%
  ggplot(aes(x = Minute, y = Brier, col = Model)) +
  geom_line() +
  geom_point() +
  scale_x_continuous(breaks = c(0, 15, 30, 45, 60, 75)) +
  theme_bw() +
  ggtitle("Last 200 matches") +
  ylab("Brier score")
```



```
tibble(Brier = apply(HDA[c(251:350),c(105:128)], 2, mean),
  Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 4)),
  Model = factor(c(rep("0", 6), rep("3", 6), rep("3 (2019)", 6),
    rep("3 (2015-2019)", 6)),
    levels = c("0", "3", "3 (2019)", "3 (2015-2019)"))) %>%
  ggplot(aes(x = Minute, y = Brier, col = Model)) +
  geom_line() +
  geom_point() +
  scale_x_continuous(breaks = c(0, 15, 30, 45, 60, 75)) +
  theme_bw() +
  ggtitle("Last 100 matches") +
  ylab("Brier score")
```

