Final

Peyton Taylor

5/6/2018

R Markdown

```
## Loading required package: quantmod
```

```
## Loading required package: xts
```

```
## Loading required package: zoo
```

```
##
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
## Loading required package: TTR
```

```
## Version 0.4-0 included new data defaults. See ?getSymbols.
```

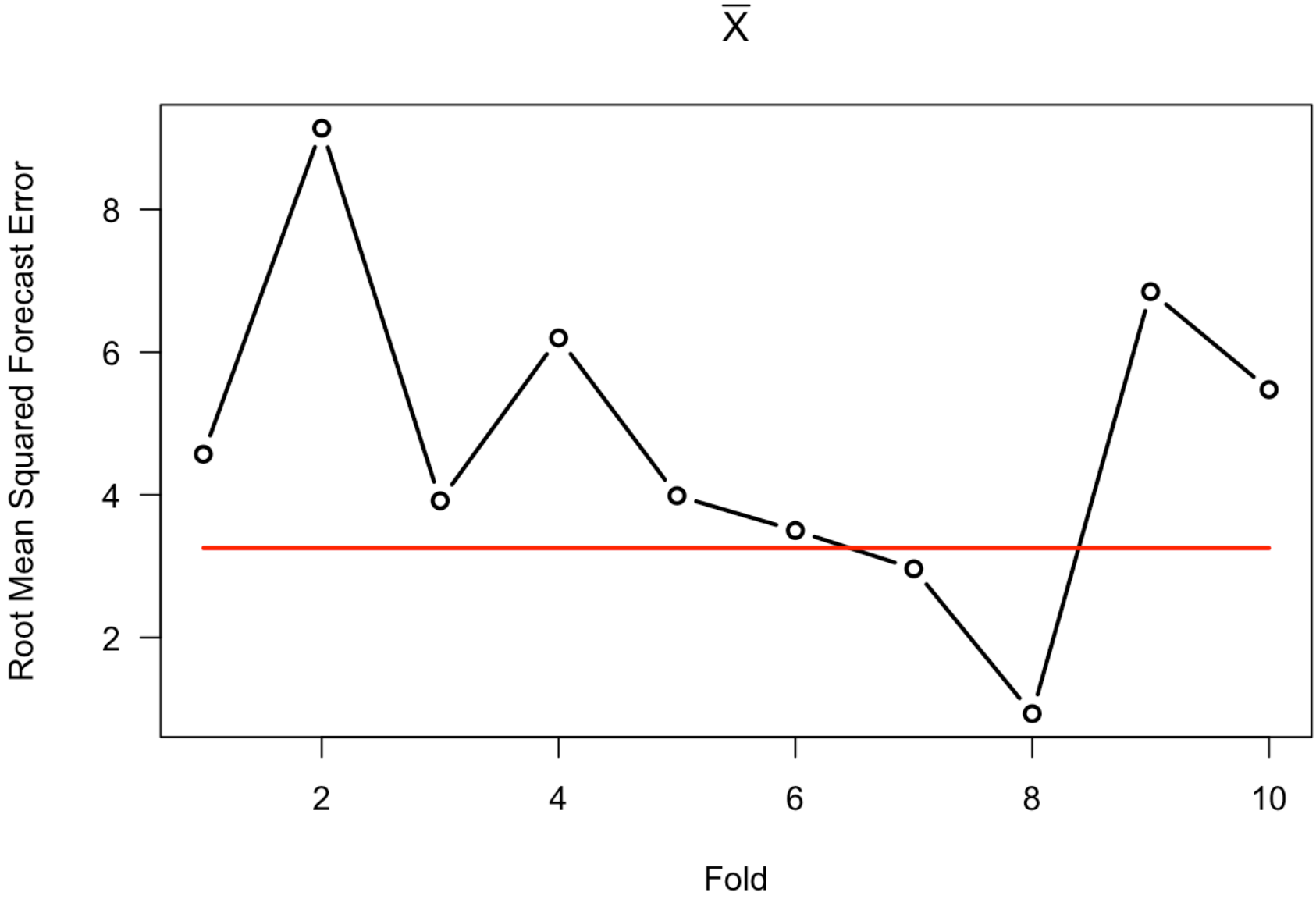
Model 1 (rmsfe)

```
avg_rev<-rep(NA,10)
rmsfe<-rep(NA,10)

for (i in 1:10) {
  avg_rev[i]<-mean(final_data$revenue[!(final_data$folds==i)])
  rmsfe[i]<-my_accuracy_msfe(actual=final_data$revenue[final_data$folds==i], predicti
on=avg_rev[i])
}

plot(x=1:10,
     y=rmsfe,
     type="b",
     xlab="Fold",
     ylab="Root Mean Squared Forecast Error",
     las=1,
     lwd=2,
     main=expression(bar(X)))

points(x=1:10, y=x_bar, type ="l", col= "red", lwd =2)
```

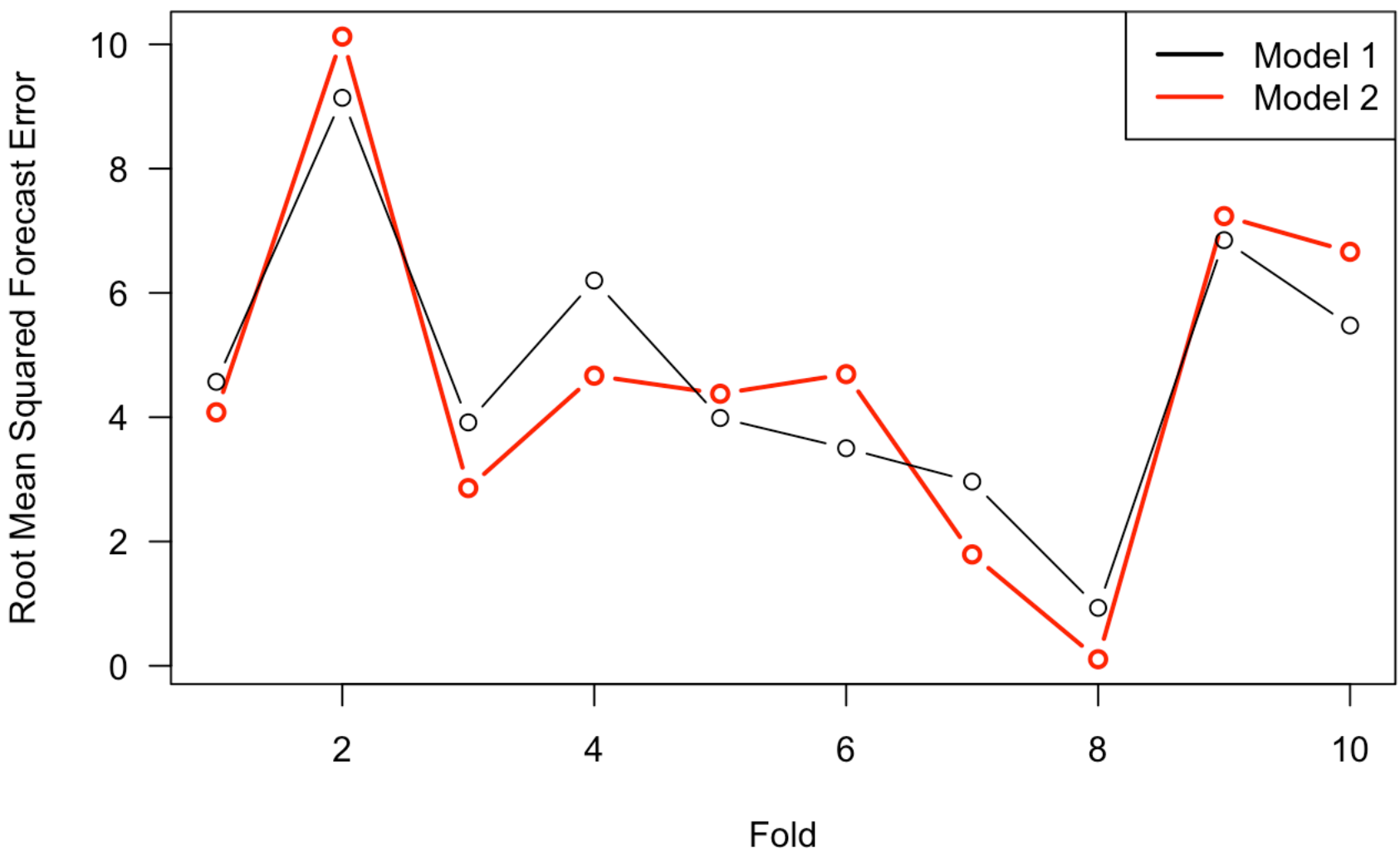


Model 2 (day, price)

```
rmsfe_price<-rep(NA,10)
for (i in 1:10) {
  model_3<-lm(revenue~day+price, data=final_data[!(final_data$folds==i),])
  rmsfe_price[i]<-my_accuracy_msfe(actual=final_data$revenue[final_data$folds==i], pr
ediction = predict(model_3, newdata=final_data[final_data$folds==i,]))
}

plot(x=1:10,
     y=rmsfe_price,
     ylim=range(c(rmsfe_price,rmsfe)),
     type="b",
     col = "red",
     xlab="Fold",
     ylab = "Root Mean Squared Forecast Error",
     las=1,
     lwd=2,
     main=expression(revenue[i]==beta[0]+beta[1]*day[i]+beta[2]*price[i]+epsilon[i]))

points(x=1:10,
       y=rmsfe,
       type="b",
       lwd=1,
       col="black")
legend("topright",legend=c("Model 1","Model 2"),col=c("black","red"),lty=c(1,1),lwd=c
(2,2))
```

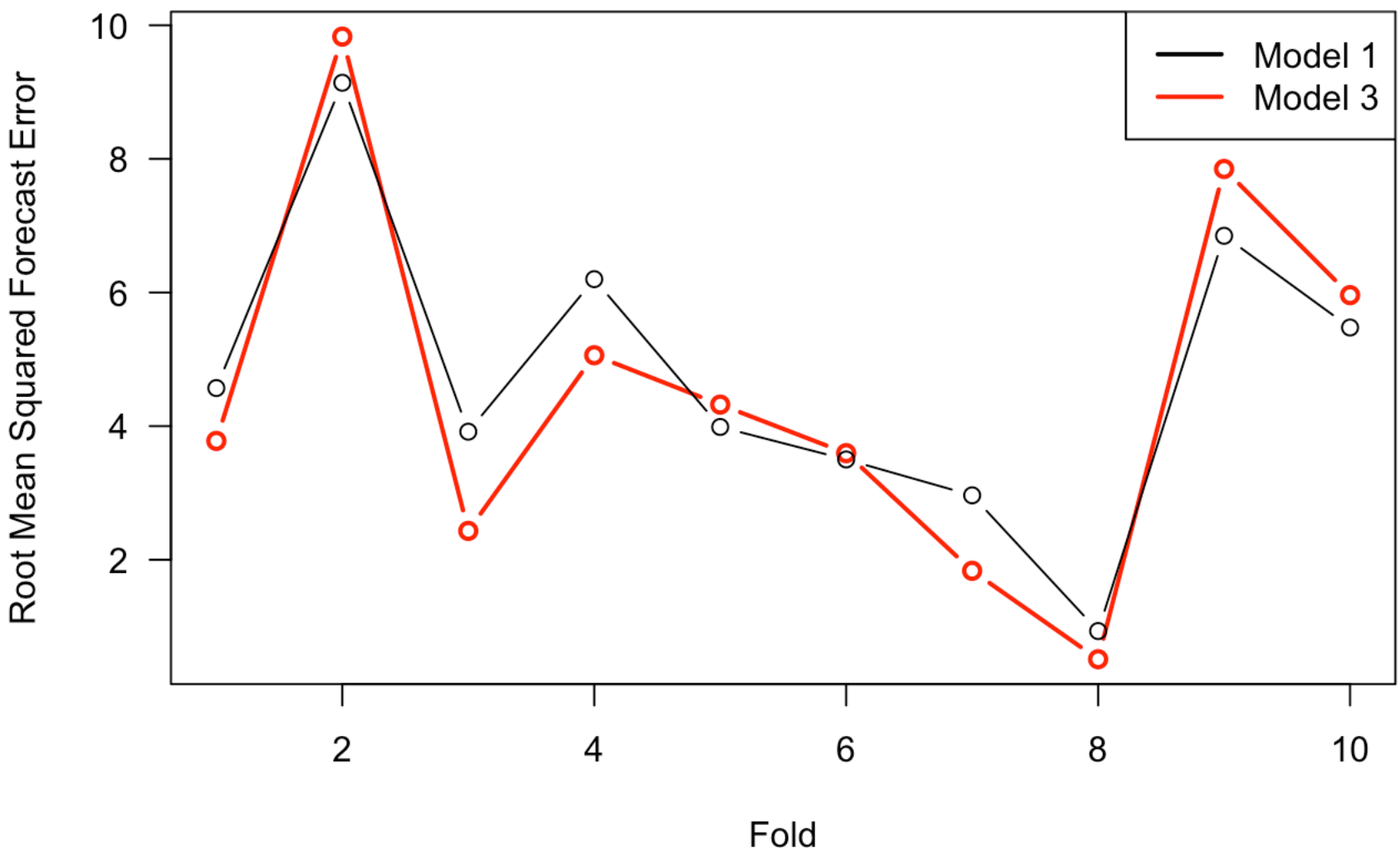


Model 3 (day, price, unit, manufactuer)

```
rmsfe_Manu<-rep(NA,10)
for (i in 1:10) {
  model_4<-lm(revenue~day+price+unit+manufacturer, data=final_data[!(final_data$folds
==i),])
  rmsfe_Manu[i]<-my_accuracy_msfe(actual=final_data$revenue[final_data$folds==i], pre
diction = predict(model_4, newdata=final_data[final_data$folds==i,]))
}

plot(x=1:10,
     y=rmsfe_Manu,
     ylim=range(c(rmsfe_Manu,rmsfe)),
     type="b",
     col = "red",
     xlab="Fold",
     ylab = "Root Mean Squared Forecast Error",
     las=1,
     lwd=2,
     main=expression(revenue[i]==beta[0]+beta[1]*day[i]+beta[2]*price[i]+Psi*Manu[i]+
epsilon[i]))

points(x=1:10,
       y=rmsfe,
       type="b",
       lwd=1,
       col="black")
legend("topright",legend=c("Model 1","Model 3"),col=c("black","red"),lty=c(1,1),lwd=c
(2,2))
```


revenue$_i = \beta_0 + \beta_1 day_i + \beta_2 price_i + \Psi Manu_i + \varepsilon_i$
Root Mean Squared Forecast Error
Fold
Model 1
Model 3
10
8
6
4
2
2
4
6
8
10

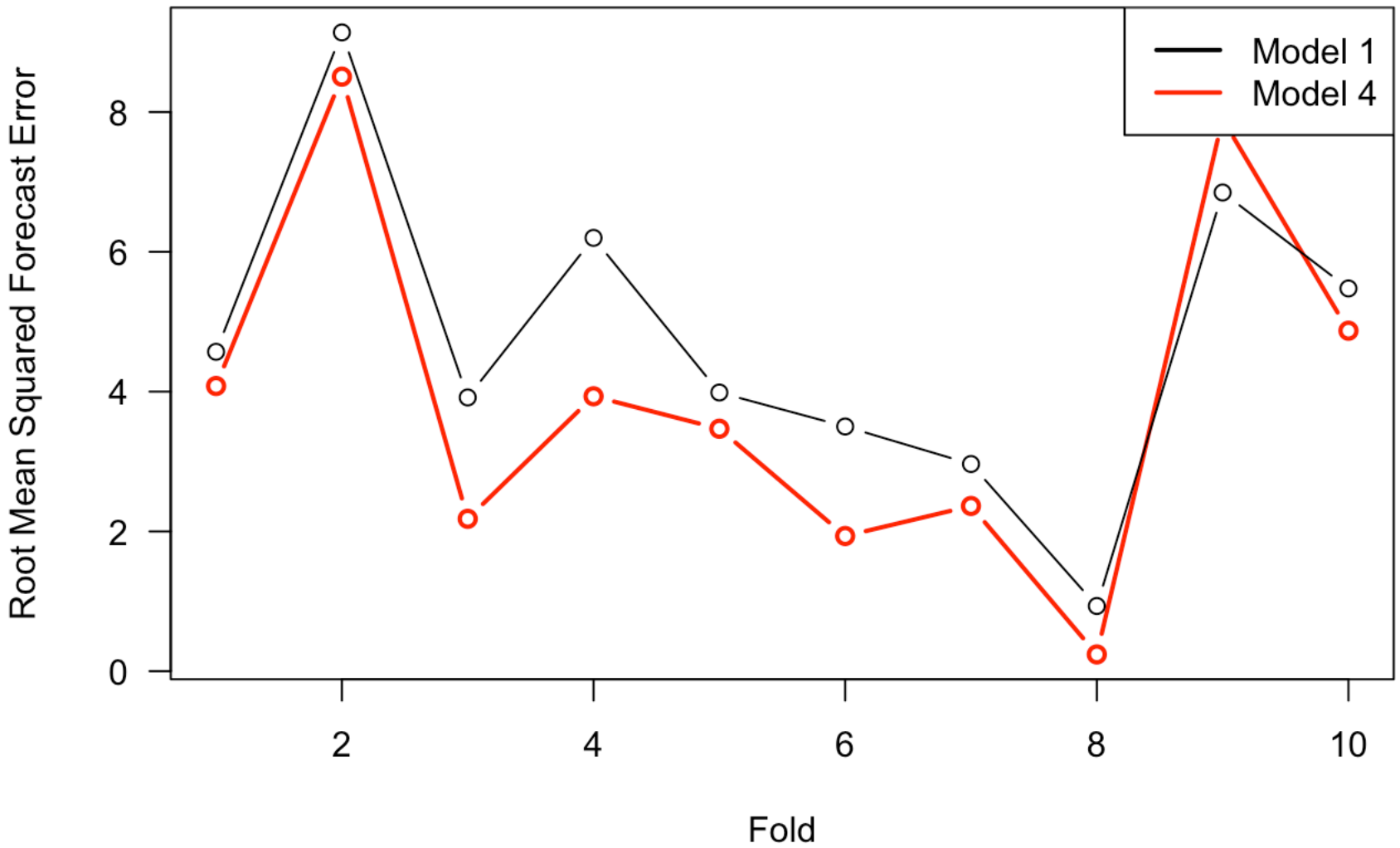
Model 4

```
rmsfe_pharmForm<-rep(NA,10)
for (i in 1:10) {
  model_4<-lm(revenue~day+price+unit+manufacturer+pharmForm, data=final_data[!(final_
data$folds==i),])
  rmsfe_pharmForm[i]<-my_accuracy_msfe(actual=final_data$revenue[final_data$folds==i]
, prediction = predict(model_4, newdata=final_data[final_data$folds==i,]))
}

plot(x=1:10,
     y=rmsfe_pharmForm,
     ylim=range(c(rmsfe_pharmForm,rmsfe)),
     type="b",
     col = "red",
     xlab="Fold",
     ylab = "Root Mean Squared Forecast Error",
     las=1,
     lwd=2,
     main=expression(revenue[i]==beta[0]+beta[1]*day[i]+beta[2]*price[i]+ beta[3]*uni
t+beta[4]*pharmForm+Psi*Manu[i]+epsilon[i]))

points(x=1:10,
       y=rmsfe,
       type="b",
       lwd=1,
       col="black")
legend("topright",legend=c("Model 1","Model 4"),col=c("black","red"),lty=c(1,1),lwd=c
(2,2))
```

$$revenue_i = \beta_0 + \beta_1 day_i + \beta_2 price_i + \beta_3 unit + \beta_4 pharmForm + \Psi Manu_i + \varepsilon_i$$



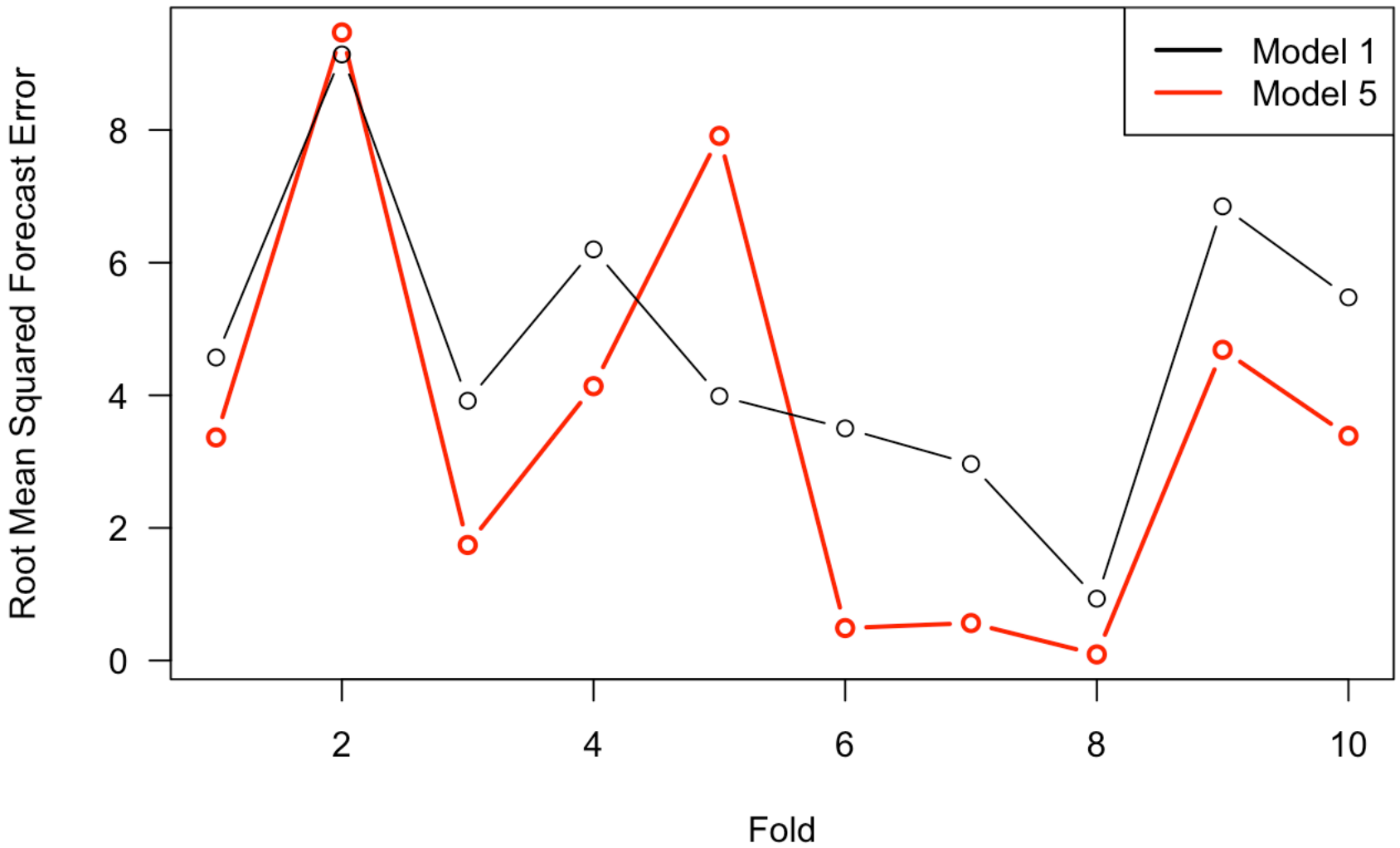
Model 5

```
rmsfe_salesIndex<-rep(NA,10)
for (i in 1:10) {
  model_4<-lm(revenue~day+manufacturer+basket+availability+price+adFlag+unit+rrp+sale
sIndex, data=final_data[!(final_data$folds==i),])
  rmsfe_salesIndex[i]<-my_accuracy_msfe(actual=final_data$revenue[final_data$folds==i
], prediction = predict(model_4, newdata=final_data[final_data$folds==i,]))
}

plot(x=1:10,
     y=rmsfe_salesIndex,
     ylim=range(c(rmsfe_salesIndex,rmsfe)),
     type="b",
     col = "red",
     xlab="Fold",
     ylab = "Root Mean Squared Forecast Error",
     las=1,
     lwd=2,
     main=expression(revenue[i]==beta[0]+beta[1]*day[i]+beta[2]+beta[3]*basket[i]+bet
a[4]*availability[i]+beta[5]*price[i]+beta[6]*adFlag[i]+beta[7]*unit[i]+beta[8]*rrp[i
]+beta[7]*salesIndex[i]+Psi*Manu[i]+epsilon[i]))

points(x=1:10,
       y=rmsfe,
       type="b",
       lwd=1,
       col="black")
legend("topright",legend=c("Model 1","Model 5"),col=c("black","red"),lty=c(1,1),lwd=c
(2,2))
```

$_1day_i + \beta_2 + \beta_3 basket_i + \beta_4 availability_i + \beta_5 price_i + \beta_6 adFlag_i + \beta_7 unit_i + \beta_8 rrp_i + \beta_7 sale$



Conclusions

```
rmsfe_df<-data.frame(rmsfe, rmsfe_price,rmsfe_Manu, rmsfe_pharmForm, rmsfe_salesIndex
)
names(rmsfe_df)<-c("Model 1", "Model 2", "Model 3", "Model 4", "Model 5")

print(rmsfe_df)
```

```
##       Model 1   Model 2   Model 3   Model 4   Model 5
## 1  4.5696489  4.079452 3.7787695 4.0806685 3.3636899
## 2  9.1393247 10.124195 9.8302006 8.5059964 9.4716338
## 3  3.9165905  2.859901 2.4318634 2.1803530 1.7416602
## 4  6.1996462  4.667970 5.0611914 3.9334032 4.1372497
## 5  3.9873222  4.377180 4.3221644 3.4687683 7.9112811
## 6  3.5009320  4.692997 3.5953739 1.9340677 0.4894435
## 7  2.9639786  1.791672 1.8348672 2.3638851 0.5647584
## 8  0.9328009  0.106041 0.5111613 0.2401237 0.0922336
## 9  6.8488365  7.235943 7.8490691 7.8987794 4.6854876
## 10 5.4758496  6.661664 5.9606138 4.8713189 3.3887961
```

```
outcomes<-c(length(which(rmsfe_price<rmsfe)==TRUE),
            length(which(rmsfe_Manu<rmsfe)==TRUE),
            length(which(rmsfe_pharmForm<rmsfe)==TRUE),
            length(which(rmsfe_salesIndex<rmsfe)==TRUE))

names(outcomes)<-c("Model 2 Wins", "Model 3 Wins", "Model 4 Wins", "Model 5 Wins")

print(outcomes)
```

```
## Model 2 Wins Model 3 Wins Model 4 Wins Model 5 Wins
##            5            5            9            8
```

After testing several models, Model #4 had the best performance with nine out of ten wins when compared against Model #1.