# Probability of Recession

## William Chiu

## 2022-11-13

## Summary

Forecast the probability of a recession in the next 3 months using the following predictors:

1. Spread between 10Y CMT and Effective Federal Funds Rate
2. YOY change in Unemployment Rate
3. YOY growth in CPI-U

## Extract Historical Data

Refer to this vignette for FRED data access.

```
library(tidyverse)
library(lubridate)
library(caTools)
library(scam)
library(fredr)
library(effects)
library(car)
library(MLmetrics)

series_id <- c("FEDFUNDS", "GS10", "USREC", "UNRATE", "CPIAUCSL")

full_data <- map_dfr(series_id, function(x) {
  fredr(
    series_id = x,
    observation_start = as.Date("1950-01-01"),
    observation_end = as.Date("2022-12-01")
  )
})
```

## Pivot Wider

```
full_data_wide_raw <- full_data %>%
  arrange(date) %>%
  select(date, series_id, value) %>%
  pivot_wider(id_cols=date, names_from = series_id, values_from = value)
```

## Calculate Features/Predictors

```
full_data_wide_features <- full_data_wide_raw %>%
  mutate(SPRD_10YCMT_FEDFUNDS = GS10 - FEDFUNDS,
         D_UNRATE = UNRATE - lag(UNRATE, 12),
         G_CPIU = (CPIAUCSL / lag(CPIAUCSL, 12) - 1) * 100
         )
```

## Recession in next 3 months

```
full_data_wide <- full_data_wide_features %>%
  arrange(date) %>%
  mutate(USREC_LEAD1 = lead(USREC, 1),
         USREC_LEAD2 = lead(USREC, 2),
         USREC_LEAD3 = lead(USREC, 3),
         USREC_3MOS = pmax(USREC_LEAD1, USREC_LEAD2, USREC_LEAD3)) %>%
  drop_na() %>%
  select(-USREC, -USREC_LEAD1, -USREC_LEAD2, -USREC_LEAD3)
```

## Split Train/Test

```
set.seed(111)

train_id <- sample.split(full_data_wide$USREC_3MOS, SplitRatio = 0.80)

train_data <- full_data_wide[train_id,]
test_data <- full_data_wide[!train_id,]

summary(train_data)
```

```
##       date                   UNRATE          CPIAUCSL           GS10
##  Min.   :1954-07-01    Min.   : 3.400   Min.   : 26.71   Min.   : 0.620
##  1st Qu.:1971-04-01    1st Qu.: 4.600   1st Qu.: 40.10   1st Qu.: 3.460
##  Median :1989-07-01    Median : 5.600   Median :124.50   Median : 5.100
##  Mean   :1988-08-23    Mean   : 5.843   Mean   :126.69   Mean   : 5.585
##  3rd Qu.:2005-12-01    3rd Qu.: 6.900   3rd Qu.:199.10   3rd Qu.: 7.390
##  Max.   :2022-07-01    Max.   :14.700   Max.   :295.33   Max.   :15.320
##     FEDFUNDS      SPRD_10YCMT_FEDFUNDS    D_UNRATE           G_CPIU
##  Min.   : 0.05   Min.   :-6.140       Min.   :-8.70000   Min.   :-1.959
##  1st Qu.: 1.75   1st Qu.: 0.280       1st Qu.:-0.70000   1st Qu.: 1.684
##  Median : 4.09   Median : 1.190       Median :-0.30000   Median : 2.849
##  Mean   : 4.53   Mean   : 1.055       Mean   :-0.03721   Mean   : 3.489
##  3rd Qu.: 6.14   3rd Qu.: 2.110       3rd Qu.: 0.30000   3rd Qu.: 4.365
##  Max.   :19.10   Max.   : 3.850       Max.   :11.10000   Max.   :14.589
##    USREC_3MOS
##  Min.   :0.0000
##  1st Qu.:0.0000
##  Median :0.0000
```

```
##  Mean    :0.1501
##  3rd Qu.:0.0000
##  Max.    :1.0000
```

## Logistic Regression
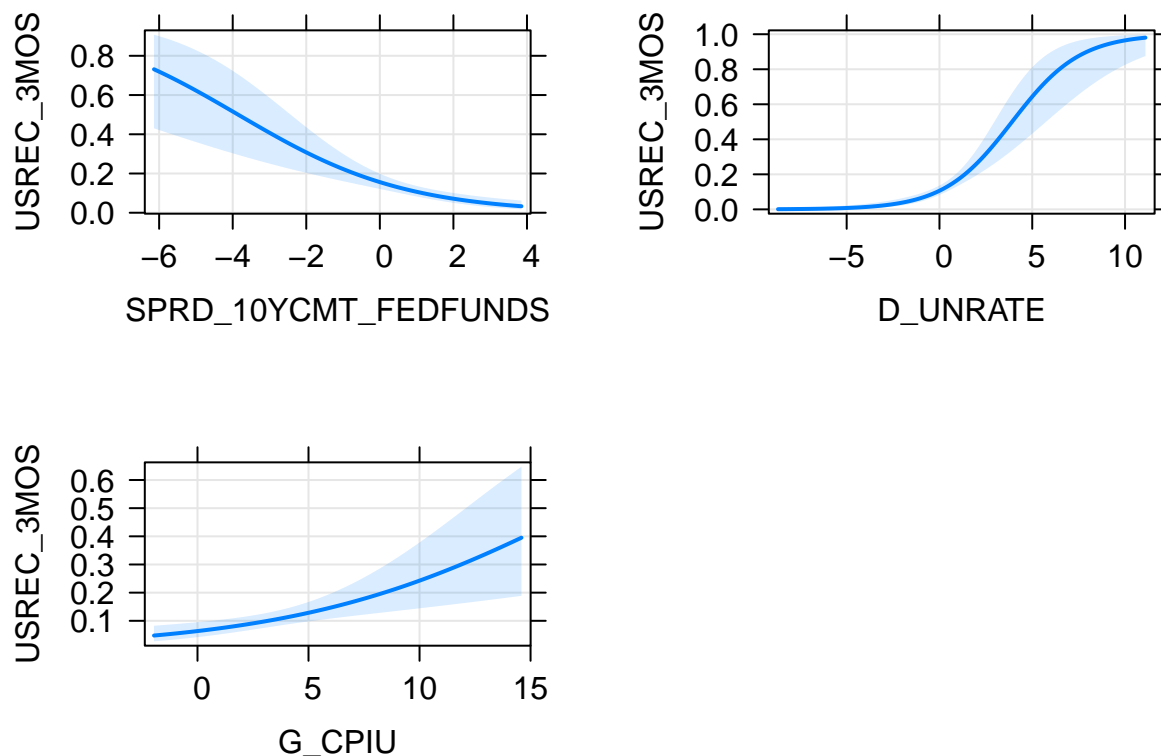
```r
logit_mod <- glm(USREC_3MOS ~ SPRD_10YCMT_FEDFUNDS +
                   D_UNRATE +
                   G_CPIU, data=train_data, family=binomial)

summary(logit_mod)
```

```
##
## Call:
## glm(formula = USREC_3MOS ~ SPRD_10YCMT_FEDFUNDS + D_UNRATE +
##     G_CPIU, family = binomial, data = train_data)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.7010  -0.4846  -0.3675  -0.2489   2.3973
##
## Coefficients:
##                       Estimate Std. Error z value Pr(>|z|)
## (Intercept)           -2.20841    0.26214  -8.424  < 2e-16 ***
## SPRD_10YCMT_FEDFUNDS  -0.43788    0.09475  -4.622 3.81e-06 ***
## D_UNRATE               0.54415    0.09383   5.799 6.66e-09 ***
## G_CPIU                 0.15521    0.04646   3.341 0.000835 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 552.23  on 652  degrees of freedom
## Residual deviance: 426.82  on 649  degrees of freedom
## AIC: 434.82
##
## Number of Fisher Scoring iterations: 5
```

## Effect Plot for Logistic Regression

```r
plot(predictorEffects(logit_mod, focal.levels=1000),
    main=NULL,
    axes = list(
      grid = TRUE,
      x = list(rug = FALSE),
      y = list(type = "response")
    ))
```

## Logit with Knots

```
logit_mod_knot <- glm(USREC_3MOS ~ SPRD_10YCMT_FEDFUNDS +
                        pmax(0,SPRD_10YCMT_FEDFUNDS) +
                        D_UNRATE +
                        pmax(0, D_UNRATE) +
                        G_CPIU +
                        pmax(0, G_CPIU),
                      data=train_data, family=binomial)

summary(logit_mod_knot)
```

```
##
## Call:
## glm(formula = USREC_3MOS ~ SPRD_10YCMT_FEDFUNDS + pmax(0, SPRD_10YCMT_FEDFUNDS) +
##     D_UNRATE + pmax(0, D_UNRATE) + G_CPIU + pmax(0, G_CPIU),
##     family = binomial, data = train_data)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.3256  -0.5214  -0.2297  -0.0339   3.7074
##
## Coefficients:
```

```
##                              Estimate Std. Error z value Pr(>|z|)
## (Intercept)                   -1.0396     0.3154  -3.297 0.000979 ***
## SPRD_10YCMT_FEDFUNDS          -0.2625     0.1962  -1.338 0.180997
## pmax(0, SPRD_10YCMT_FEDFUNDS) -0.2941     0.2805  -1.048 0.294494
## D_UNRATE                       5.2405     0.8986   5.832 5.47e-09 ***
## pmax(0, D_UNRATE)             -5.1232     0.9291  -5.514 3.51e-08 ***
## G_CPIU                        -0.4150     1.0358  -0.401 0.688657
## pmax(0, G_CPIU)                0.5723     1.0464   0.547 0.584431
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 552.23  on 652  degrees of freedom
## Residual deviance: 372.02  on 646  degrees of freedom
## AIC: 386.02
##
## Number of Fisher Scoring iterations: 8
```

## Effect Plot for Knots

```
plot(predictorEffects(logit_mod_knot, focal.levels=1000),
    main=NULL,
    axes = list(
      grid = TRUE,
      x = list(rug = FALSE),
      y = list(type = "response")
    ))
```
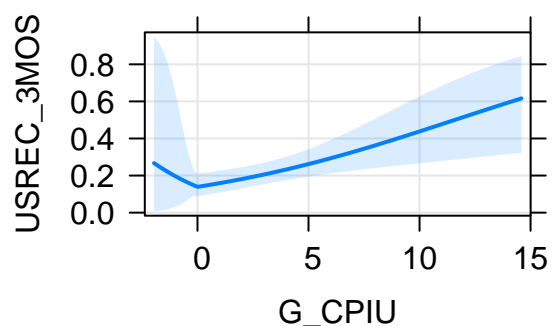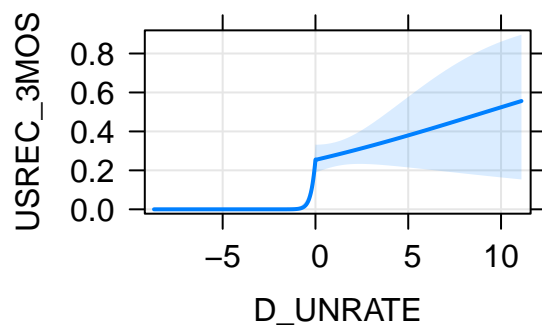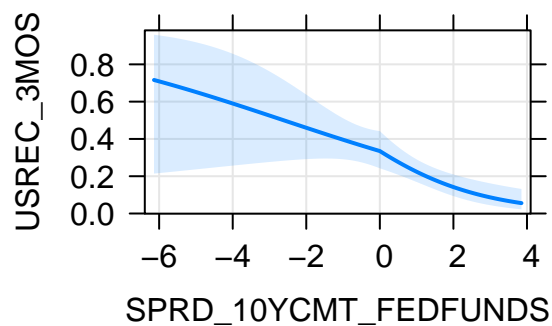
## Shape-Constrained GAM

```r
scam_mod <- scam(USREC_3MOS ~ s(SPRD_10YCMT_FEDFUNDS, bs="mpd") +
                    D_UNRATE  +
                    G_CPIU,
                 data=train_data, family=binomial())

summary(scam_mod)
```

```
##
## Family: binomial
## Link function: logit
##
## Formula:
## USREC_3MOS ~ s(SPRD_10YCMT_FEDFUNDS, bs = "mpd") + D_UNRATE +
##      G_CPIU
##
## Parametric coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.78007    1.04342   0.748  0.45470
## D_UNRATE     0.54794    0.09487   5.776 7.66e-09 ***
## G_CPIU       0.15715    0.04631   3.394  0.00069 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                          edf Ref.df Chi.sq p-value
## s(SPRD_10YCMT_FEDFUNDS) 1.312  1.537  9.994 0.00449 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.2018   Deviance explained = 22.8%
## UBRE score = -0.33425  Scale est. = 1         n = 653
```

```
plot(scam_mod,pages=1,se=FALSE)
```
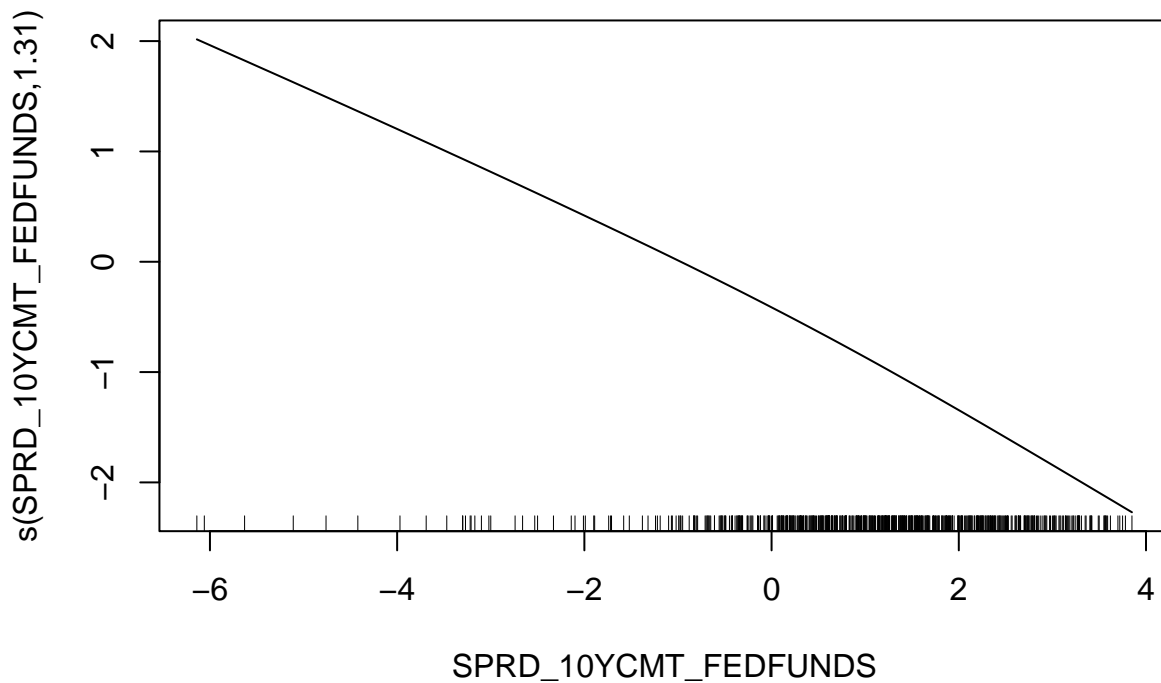


## Null Model

```
null_mod <- glm(USREC_3MOS ~ 1, data=train_data, family=binomial)

summary(null_mod)
```

```
##
## Call:
## glm(formula = USREC_3MOS ~ 1, family = binomial, data = train_data)
##
```

```
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.5703  -0.5703  -0.5703  -0.5703   1.9476
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -1.7340     0.1096  -15.82   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 552.23  on 652  degrees of freedom
## Residual deviance: 552.23  on 652  degrees of freedom
## AIC: 554.23
##
## Number of Fisher Scoring iterations: 4
```

**Performance Metrics**

```r
test_preds <- predict(logit_mod, newdata=test_data, type="response")
null_preds <- predict(null_mod, newdata=test_data, type="response")
knot_preds <- predict(logit_mod_knot, newdata=test_data, type="response")
scam_preds <- predict(scam_mod, newdata=test_data, type="response")


perf <- function(lst_preds, f_metric=caTools::colAUC, metricname="ROC-AUC"){
  map_dfr(lst_preds, function(x){
  f_metric(x, test_data$USREC_3MOS)
}) %>%
  pivot_longer(everything(), names_to="model", values_to=metricname) %>%
  knitr::kable()
}

myPreds <- list(logit_reg=test_preds, null_model=null_preds,
                knot_reg=knot_preds, scam_mod = scam_preds)

perf(myPreds, caTools::colAUC, "ROC-AUC")
```

| model      | ROC-AUC   |
|------------|-----------|
| logit_reg  | 0.8457554 |
| null_model | 0.5000000 |
| knot_reg   | 0.9027338 |
| scam_mod   | 0.8454676 |

```r
perf(myPreds, MLmetrics::LogLoss, "LogLoss")
```

| model | LogLoss |
|---|---|
| logit_reg | 0.3273263 |
| null_model | 0.4269384 |
| knot_reg | 0.2757863 |
| scam_mod | 0.3257167 |

## Probability of Recession (10/01/2022)

```
curr_data <- tail(full_data_wide_features, 1) %>%
  select(date, UNRATE, SPRD_10YCMT_FEDFUNDS, D_UNRATE, G_CPIU)

knitr::kable(curr_data)
```

| date | UNRATE | SPRD_10YCMT_FEDFUNDS | D_UNRATE | G_CPIU |
|---|---|---|---|---|
| 2022-10-01 | 3.7 | 0.9 | -0.9 | 7.763115 |

```
mods <- list(logistic_reg=logit_mod,
                          scam_mod=scam_mod,
                          knot_mod=logit_mod_knot,
                          baseline=null_mod)

output <- map_dfc(mods, function(x){
  predict(x, newdata=curr_data, type="response")
}) %>%
  pivot_longer(everything(), names_to = "model",
              values_to = "prob_rec")

output$prob_rec <- scales::percent(output$prob_rec)

knitr::kable(output)
```

| model | prob_rec |
|---|---|
| logistic_reg | 13.15% |
| scam_mod | 13.42% |
| knot_mod | 0.65% |
| baseline | 15.01% |

Relative to the historical baseline, the probability of a recession in the next 3 months is low or close to the historical risk of a recession.