

Probability of Recession

William Chiu

2022-11-13

Summary

Forecast the probability of a recession in the next 6 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_LEAD4 = lead(USREC, 4),
         USREC_LEAD5 = lead(USREC, 5),
         USREC_LEAD6 = lead(USREC, 6),
         FUTREC = pmax(USREC_LEAD1, USREC_LEAD2, USREC_LEAD3,
                       USREC_LEAD4, USREC_LEAD5, USREC_LEAD6)) %>%
  drop_na() %>%
  select(-USREC, -USREC_LEAD1, -USREC_LEAD2, -USREC_LEAD3,
        -USREC_LEAD4, -USREC_LEAD5, -USREC_LEAD6)
```

Split Train/Test

```
set.seed(111)

train_id <- sample.split(full_data_wide$FUTREC, 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. : 3.40	Min. : 26.71	Min. : 0.62
##	1st Qu.:1970-12-16	1st Qu.: 4.70	1st Qu.: 39.85	1st Qu.: 3.55
##	Median :1988-08-01	Median : 5.60	Median :119.00	Median : 5.10
##	Mean :1988-03-20	Mean : 5.91	Mean :124.76	Mean : 5.60
##	3rd Qu.:2005-04-16	3rd Qu.: 6.90	3rd Qu.:193.65	3rd Qu.: 7.39
##	Max. :2022-04-01	Max. :14.70	Max. :288.66	Max. :15.32
##	FEDFUNDS	SPRD_10YCMT_FEDFUNDS	D_UNRATE	G_CPIU
##	Min. : 0.050	Min. :-6.140	Min. :-8.70000	Min. :-1.959
##	1st Qu.: 1.750	1st Qu.: 0.340	1st Qu.: -0.70000	1st Qu.: 1.669
##	Median : 4.060	Median : 1.220	Median :-0.30000	Median : 2.845
##	Mean : 4.478	Mean : 1.122	Mean : 0.05008	Mean : 3.456
##	3rd Qu.: 6.110	3rd Qu.: 2.180	3rd Qu.: 0.40000	3rd Qu.: 4.292

```
## Max. :19.100 Max. : 3.850 Max. :11.10000 Max. :14.592
## FUTREC
## Min. :0.0000
## 1st Qu.:0.0000
## Median :0.0000
## Mean :0.1874
## 3rd Qu.:0.0000
## Max. :1.0000
```

Logistic Regression

```
logit_mod <- glm(FUTREC ~ SPRD_10YCMT_FEDFUNDS +
  D_UNRATE +
  G_CPIU, data=train_data, family=binomial)

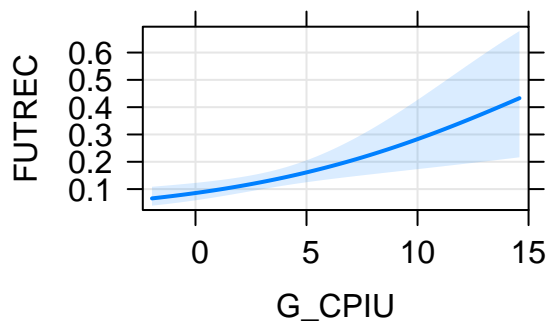
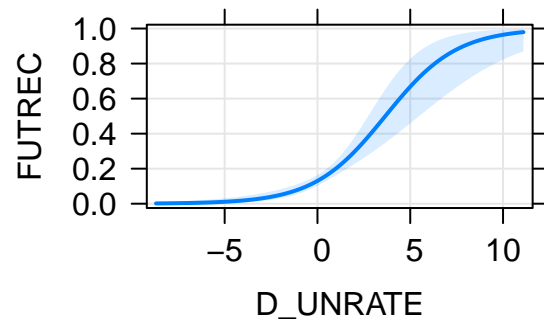
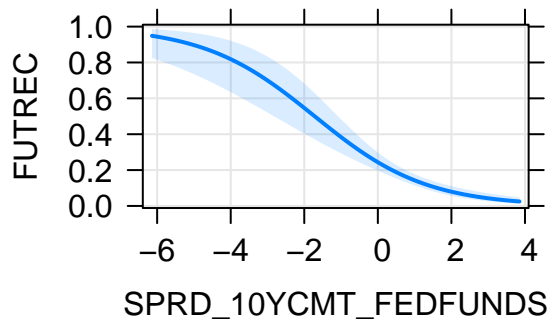
summary(logit_mod)
```

```
##
## Call:
## glm(formula = FUTREC ~ SPRD_10YCMT_FEDFUNDS + D_UNRATE + G_CPIU,
##      family = binomial, data = train_data)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -3.0820  -0.5476  -0.3695  -0.2369   2.2615
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.65543    0.24176  -6.847 7.52e-12 ***
## SPRD_10YCMT_FEDFUNDS -0.65927    0.10119  -6.515 7.26e-11 ***
## D_UNRATE         0.51944    0.09406   5.522 3.35e-08 ***
## G_CPIU           0.14389    0.04504   3.195  0.0014 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 628.13  on 650  degrees of freedom
## Residual deviance: 464.85  on 647  degrees of freedom
## AIC: 472.85
##
## Number of Fisher Scoring iterations: 5
```

Effect Plot for Logistic Regression

```
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(FUTREC ~ 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)

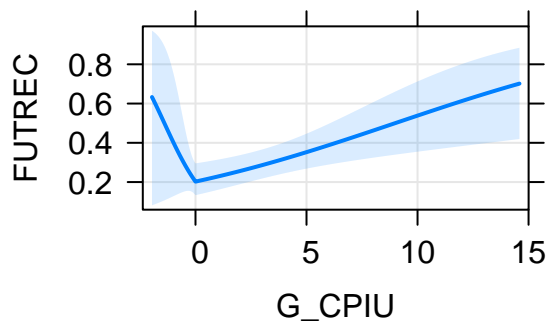
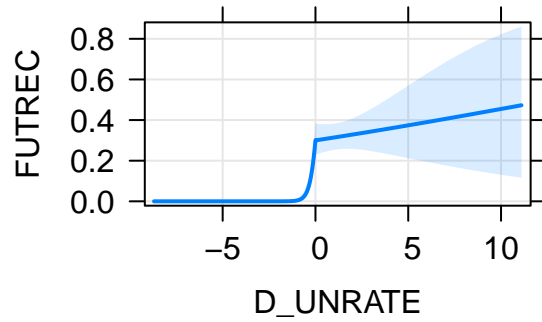
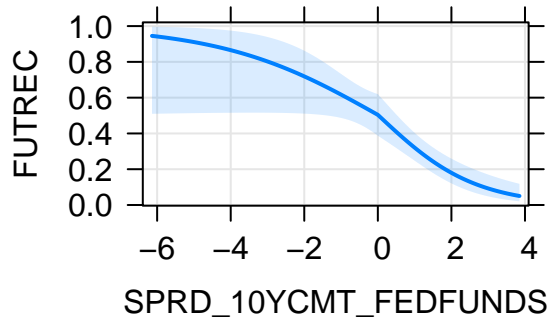
summary(logit_mod_knot)
```

```
##
## Call:
## glm(formula = FUTREC ~ 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
## -2.9192  -0.5329  -0.2528  -0.0382   3.4100
##
## Coefficients:
##                      Estimate Std. Error z value Pr(>|z|)
## (Intercept)          -0.5150    0.3010  -1.711   0.0871 .
## SPRD_10YCMT_FEDFUNDS  -0.4615    0.2458  -1.878   0.0604 .
## pmax(0, SPRD_10YCMT_FEDFUNDS) -0.3050    0.3297  -0.925   0.3549
## D_UNRATE              4.6484    0.7216   6.442 1.18e-10 ***
## pmax(0, D_UNRATE)      -4.5818    0.7544  -6.074 1.25e-09 ***
## G_CPIU                -0.9786    0.7865  -1.244   0.2134
## pmax(0, G_CPIU)         1.1312    0.7995   1.415   0.1571
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 628.13  on 650  degrees of freedom
## Residual deviance: 406.95  on 644  degrees of freedom
## AIC: 420.95
##
## Number of Fisher Scoring iterations: 7
```

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

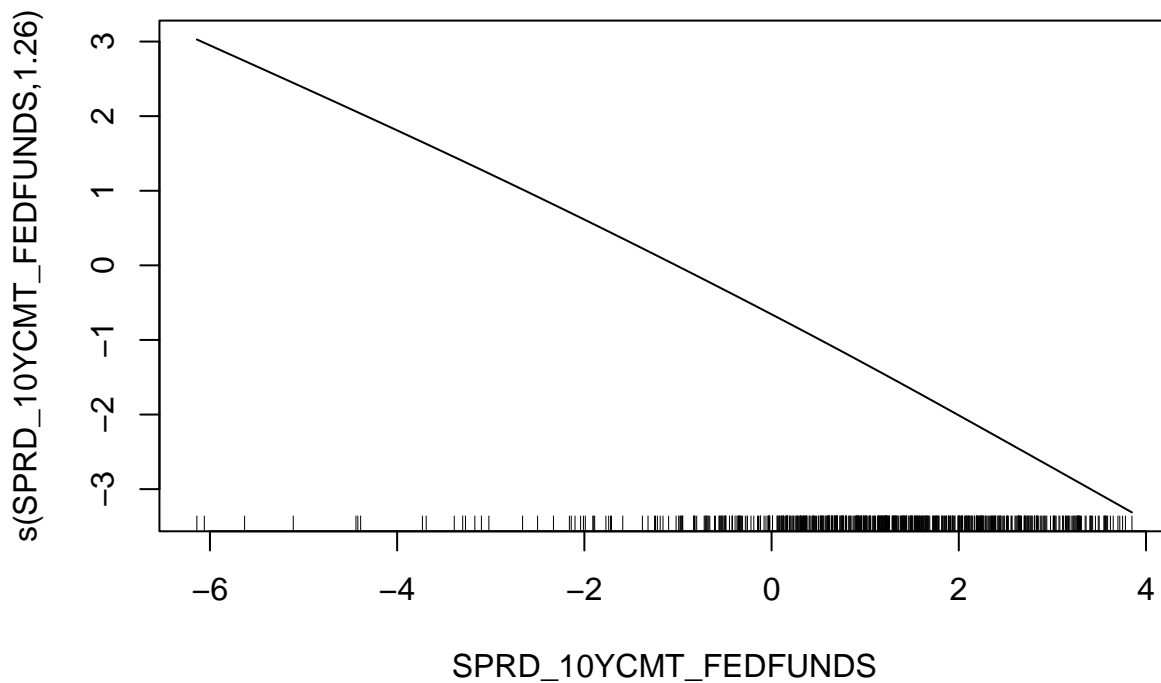
```
scam_mod <- scam(FUTREC ~ s(SPRD_10YCMT_FEDFUNDS, bs="mpd") +
  D_UNRATE +
  G_CPIU,
  data=train_data, family=binomial())

summary(scam_mod)
```

```
##
## Family: binomial
## Link function: logit
##
## Formula:
## FUTREC ~ s(SPRD_10YCMT_FEDFUNDS, bs = "mpd") + D_UNRATE + G_CPIU
##
## Parametric coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  2.84371    1.19200   2.386  0.01705 *
## D_UNRATE     0.52304    0.09499   5.507 3.66e-08 ***
## G_CPIU       0.14558    0.04521   3.220  0.00128 **
## ---
## 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) 1.265  1.474  16.27 0.000213 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.2706   Deviance explained = 26.1%
## UBRE score = -0.2738   Scale est. = 1         n = 651
```

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



Null Model

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

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

```
##      Min      1Q   Median      3Q      Max
## -0.6442 -0.6442 -0.6442 -0.6442  1.8300
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -1.4670     0.1004  -14.61  <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: 628.13  on 650  degrees of freedom
## Residual deviance: 628.13  on 650  degrees of freedom
## AIC: 630.13
##
## Number of Fisher Scoring iterations: 4
```

Performance Metrics

```
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$FUTREC)
  }) %>%
  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	0.8695015
null_model	0.5000000
knot_reg	0.8966276
scam_mod	0.8685239

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

model	LogLoss
logit_reg	0.3561798

model	LogLoss
null_model	0.4865150
knot_reg	0.3229384
scam_mod	0.3542811

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	16.81%
scam_mod	17.06%
knot_mod	1.47%
baseline	18.74%

Relative to the historical baseline, the probability of a recession in the next 6 months is low or close to the historical risk of a recession.