

Finance Quantitative

Modèle à un facteur

Patrick Hénaff

Version: 06 mars 2023

```
library(xts)
library(hornpa)
library(lubridate)
library(xtable)
library(quantmod)
library(PerformanceAnalytics)
library(TTR)
library(lubridate)
library(roll)
library(Hmisc)
library(nFactors)
library(kableExtra)
library(broom)
library(tidyverse)
library(ggplot2)
library(cowplot)

get.src.folder <- function() {
  path.expand("../GP/src")
}

get.data.folder <- function() {
  path.expand("../GP/data")
}

source(file.path(get.src.folder(), 'utils.R'))
source(file.path(get.src.folder(), 'FileUtils.R'))
```

Données

Séries de rendement quotidien pour 11 valeurs:

```
monthly.ret.file <- file.path(get.data.folder(), "monthly.ret.rda")
load(monthly.ret.file)
```

Pour l'indice de marché, on utilise VT, un ETF "World Market":

```

VT.series.file <- file.path(get.data.folder(), "ret.VT.rda")

if(!file.exists(VT.series.file)) {

sym <- "VT"
world.index <- Ad(getSymbols(sym, auto.assign=FALSE))
world.index.ret <- monthlyReturn(world.index)
colnames(world.index.ret) <- "Market"
save(world.index.ret, file=VT.series.file)
} else {
  load(VT.series.file)
}

```

Rendement moyen:

```

monthly.ret <- merge.xts(monthly.ret, world.index.ret, join="inner")
kable(colMeans(monthly.ret), "latex", escape=FALSE, col.names=c("$r$"), caption="Average monthly return")

```

Table 1: Average monthly return

	r
AAPL	0.0220532
AMZN	0.0271364
MSFT	0.0169185
F	0.0139604
SPY	0.0086184
QQQ	0.0126927
XOM	0.0012265
MMM	0.0090297
HD	0.0191698
PG	0.0080793
KO	0.0096675
Market	0.0063881

Matrice de covariance des rendements:

```

kable(cov(monthly.ret), "latex", booktabs=T) %>%
kable_styling(latex_options="scale_down")

```

taux sans risque

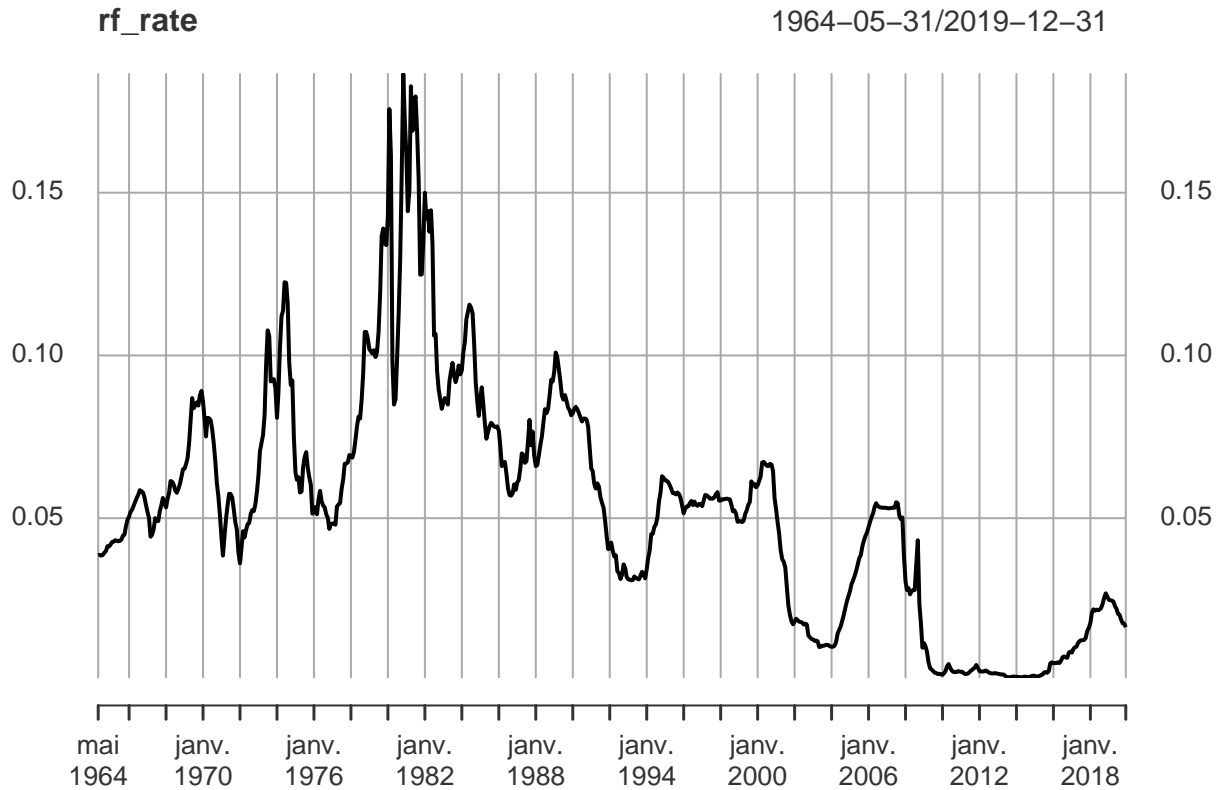
Le taux sans risque mensuel (annualisé) est obtenu de la Réserve Fédérale US.

	AAPL	AMZN	MSFT	F	SPY	QQQ	XOM	MMM	HD	PG	KO	Market
AAPL	0.0067861	0.0029132	0.0023909	0.0034726	0.0020525	0.0030696	0.0008125	0.0019703	0.0017385	0.0007716	0.0007773	0.0019879
AMZN	0.0029132	0.0081477	0.0025052	0.0026818	0.0019708	0.0029000	0.0008198	0.0013520	0.0018658	0.0001333	0.0011566	0.0020887
MSFT	0.0023909	0.0025052	0.0041486	0.0034082	0.0018237	0.0022291	0.0010236	0.0014625	0.0016284	0.0007682	0.0010500	0.0019091
F	0.0034726	0.0026818	0.0034082	0.0228940	0.0033899	0.0035843	0.0013655	0.0039663	0.0034734	0.0018252	0.0017233	0.0037993
SPY	0.0020525	0.0019708	0.0018237	0.0033899	0.0018541	0.0019954	0.0012216	0.0018248	0.0017008	0.0008786	0.0009489	0.0019549
QQQ	0.0030696	0.0029000	0.0022291	0.0035843	0.0019954	0.0025283	0.0009971	0.0018315	0.0018600	0.0007702	0.0008702	0.0020805
XOM	0.0008125	0.0008198	0.0010236	0.0013655	0.0012216	0.0009971	0.0024359	0.0015475	0.0011221	0.0006220	0.0007314	0.0012568
MMM	0.0019703	0.0013520	0.0014625	0.0039663	0.0018248	0.0018315	0.0015475	0.0033789	0.0018843	0.0010283	0.0008990	0.0018143
HD	0.0017385	0.0018658	0.0016284	0.0034734	0.0017008	0.0018600	0.0011221	0.0018843	0.0034615	0.0008112	0.0007124	0.0015536
PG	0.0007716	0.0001333	0.0007682	0.0018252	0.0008786	0.0007702	0.0006220	0.0010283	0.0008112	0.0018438	0.0008778	0.0008302
KO	0.0007773	0.0011566	0.0010500	0.0017233	0.0009489	0.0008702	0.0007314	0.0008990	0.0007124	0.0008778	0.0020062	0.0010466
Market	0.0019879	0.0020887	0.0019091	0.0037993	0.0019549	0.0020805	0.0012568	0.0018143	0.0015536	0.0008302	0.0010466	0.0023080

```

taux.sans.risque.csv <- file.path(get.data.folder(), "DP_LIVE_01032020211755676.csv")
tmp <- read.csv(taux.sans.risque.csv, header=TRUE, sep=";")[, c("TIME", "Value")]
dt <- ymd(paste(tmp$TIME, "-01", sep=""))-1
rf_rate <- xts(tmp$Value/100.0, dt)

```



```

##                AAPL                AMZN                MSFT                F                SPY
## 2008-06-30 -0.11290069 -0.10156825 -0.02860143 -0.292647079 -0.083575759
## 2008-07-31 -0.05070466  0.04104724 -0.06506746 -0.002079304 -0.008985578
## 2008-09-30 -0.32955848 -0.09961634 -0.02198624  0.165918802 -0.094173681
## 2008-10-31 -0.05340487 -0.21330401 -0.16335732 -0.578846139 -0.165186687
## 2008-12-31 -0.07899032  0.20093672 -0.03857552 -0.148698828  0.009796723
## 2009-03-31  0.17702424  0.13350827  0.13746188  0.314999474  0.083310627
##                QQQ                XOM                MMM                HD                PG
## 2008-06-30 -0.09615030 -0.007097684 -0.10275897 -0.136855504 -0.079333799
## 2008-07-31  0.00642039 -0.087370994  0.01149577  0.017506396  0.083531745
## 2008-09-30 -0.15576296 -0.029371125 -0.04594999 -0.037367436 -0.001146315

```

```
## 2008-10-31 -0.15471594 -0.045583125 -0.05870276 -0.088837478 -0.067977481
## 2008-12-31  0.02277380 -0.003992837 -0.14029598  0.006785848 -0.039316221
## 2009-03-31  0.10316511  0.002945407  0.09370851  0.141905895 -0.022420454
##
##           KO           Market           Rf
## 2008-06-30 -0.080159854 -0.0008065077  0.0023250000
## 2008-07-31 -0.009234303 -0.0268471854  0.0023250000
## 2008-09-30  0.030105361 -0.0910820596  0.0036000000
## 2008-10-31 -0.166793449 -0.2143686201  0.0019666667
## 2008-12-31 -0.034136711  0.0541143392  0.0008500000
## 2009-03-31  0.098898124  0.0883459620  0.0007416667
```

Estimation d'un modèle à un facteur

Choisir une période de 48 mois. A partir des exemples présentés en cours, estimer le modèle:

$$R_i(t) - R_f(t) = \alpha + \beta(R_M(t) - R_f(t)) + \epsilon(t)$$

en utilisant la fonction `lm`. Utilisez la fonction `kable` de `knitr` pour produire une présentation soignée des résultats.

```
nb.obs <- 48
Assets <- c("AAPL", "AMZN", "MSFT", "F", "XOM", "MMM", "HD", "PG", "KO")
r.set <- monthly.ret.2[1:nb.obs,]
r.set$SPY <- NULL
r.set$QQQ <- NULL
# Excess return
excess.r <- r.set[, c(Assets, "Market")]
for(i in seq_along(ncol(excess.r))) {
  excess.r[,i] <- excess.r[,i] - r.set$Rf
}

sigma2.M <- as.numeric(var(excess.r$Market))
r.M <- mean(excess.r$Market)

res <- data.frame(alpha=double(), beta=double(), sigma.e=double(), asset=character())

for(A in Assets) {
  tmp <- lm(paste(A, " ~ Market"), data=excess.r)
  alpha <- tmp$coefficients["(Intercept)"]
  beta <- tmp$coefficients["Market"]
  sigma.e <- glance(tmp)$sigma
  p.value <- tidy(tmp)$p.value[1]
  res <- rbind(res, list(alpha=alpha, beta=beta, sigma.e=sigma.e, p.value=p.value, asset=A), stringsAsFactors=FALSE)
}
rownames(res) <- res$asset
res$Mean <- apply(excess.r[, rownames(res)], 2, mean)
res$Sd <- apply(excess.r[, rownames(res)], 2, sd)

res$asset <- NULL
```

	α	β	σ_e	$Pr(> t)_\alpha$
AAPL	0.0070516	0.9410955	0.0818077	0.5599770
AMZN	0.0160117	0.9205405	0.0764936	0.1606683
MSFT	0.0007420	0.6855290	0.0555355	0.9278776
F	0.0097772	2.2831124	0.1917387	0.7299182
XOM	-0.0022001	0.3598236	0.0401517	0.7106808
MMM	0.0119003	0.7225732	0.0484812	0.1013119
HD	0.0194094	0.7118749	0.0516044	0.0137477
PG	0.0028729	0.2927801	0.0425845	0.6480080
KO	0.0009564	0.4006184	0.0447764	0.8849566

```
res_disp <- res[, c("alpha", "beta", "sigma.e", "p.value")]
colnames(res_disp) <- c("$\\alpha$", "$\\beta$", "$\\sigma_e$", "$Pr(>|t|)_\\alpha$")
kable(res_disp, "latex", booktabs=T, escape=FALSE) %>% kable_styling(latex_options="striped")
```

Sur cet intervalle, seul HD a un α significativement différent de 0.

L'étape suivante est d'étudier la stabilité dans le temps des α . Un indicateur de cherté relative fiable oscillerait autour de 0 de manière assez régulière: un écart par rapport au prix d'équilibre serait plus ou moins rapidement comblé. Par contre, un α qui resterait structurellement positif ou négatif ne peut plus être interprété comme un indicateur de cherté relative. Dans ce cas, α capture des facteurs de risque qui ne sont pas pris en compte par notre modèle de marché.

Pour étudier l'évolution de α dans le temps, on évalue le modèle sur des intervalles glissants, en se décalant d'un mois à chaque estimation. On obtient une série chronologique de coefficient α pour chaque titre, que l'on représente graphiquement avec la p-value correspondante. Dans les calculs qui suivent, le terme constant α n'est presque jamais significatif.

```
calc.alpha.beta <- function(r.data) {
  res <- data.frame(matrix(ncol=5,nrow=0, dimnames=list(NULL, c('alpha', 'beta', 'sigma.e', 'asset', 'd
  dt.calc <- as.Date(last(index(r.data))))
  for(A in Assets) {
    tmp <- lm(paste(A, " ~ Market"), data=r.data)
    alpha <- tmp$coefficients["(Intercept)"]
    beta <- tmp$coefficients["Market"]
    sigma.e <- glance(tmp)$sigma
    p.value <- tidy(tmp)$p.value[1]
    r.bar <- mean(r.data[,A])
    res <- rbind(res, list(alpha=alpha, beta=beta, sigma.e=sigma.e, p.value=p.value, r.bar=r.bar, asset=A
  }

  res
}

r.set2 <- monthly.ret.2
r.set2$SPY <- NULL
r.set2$QQQ <- NULL
# Excess return
excess.r <- r.set2[, c(Assets, "Market")]
for(i in seq_along(ncol(excess.r))) {
  excess.r[,i] <- excess.r[,i] - r.set2$Rf
}
```

```

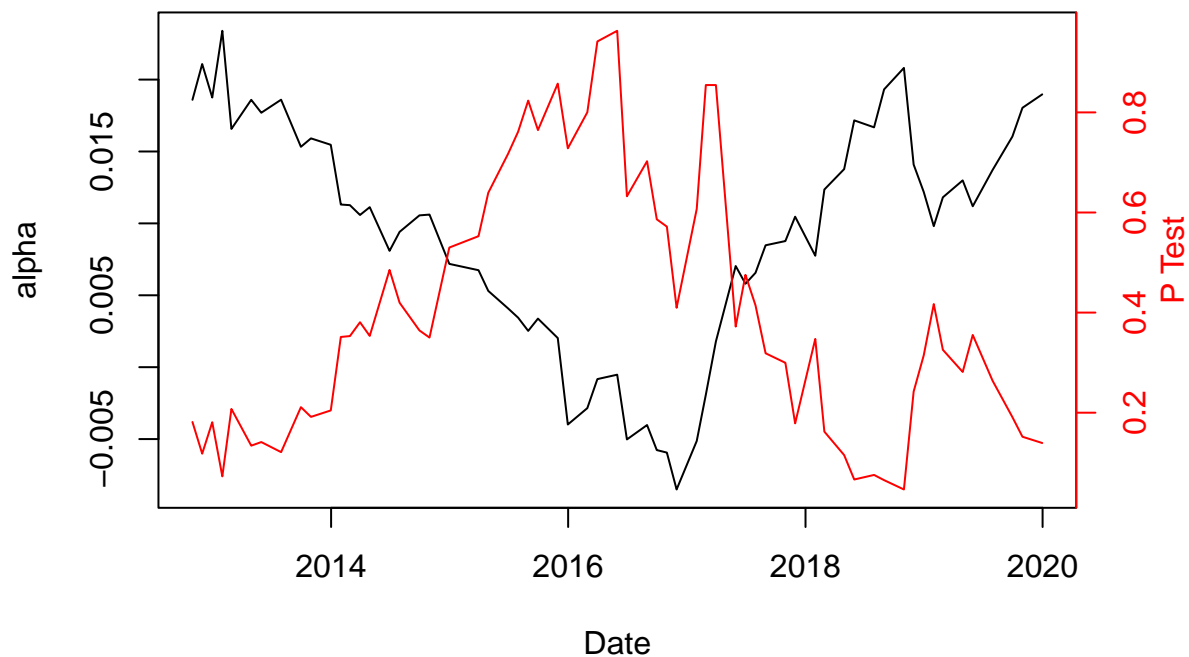
}

nb.obs <- 36
df.res <- data.frame(matrix(ncol=5,nrow=0, dimnames=list(NULL, c('alpha', 'beta', 'sigma.e', 'asset', 'p'),
for(idx.last in seq(nb.obs, nrow(excess.r))) {
  idx.first <- idx.last - nb.obs + 1
  res <- calc.alpha.beta(excess.r[idx.first:idx.last,])
  df.res <- rbind(df.res, res)
}

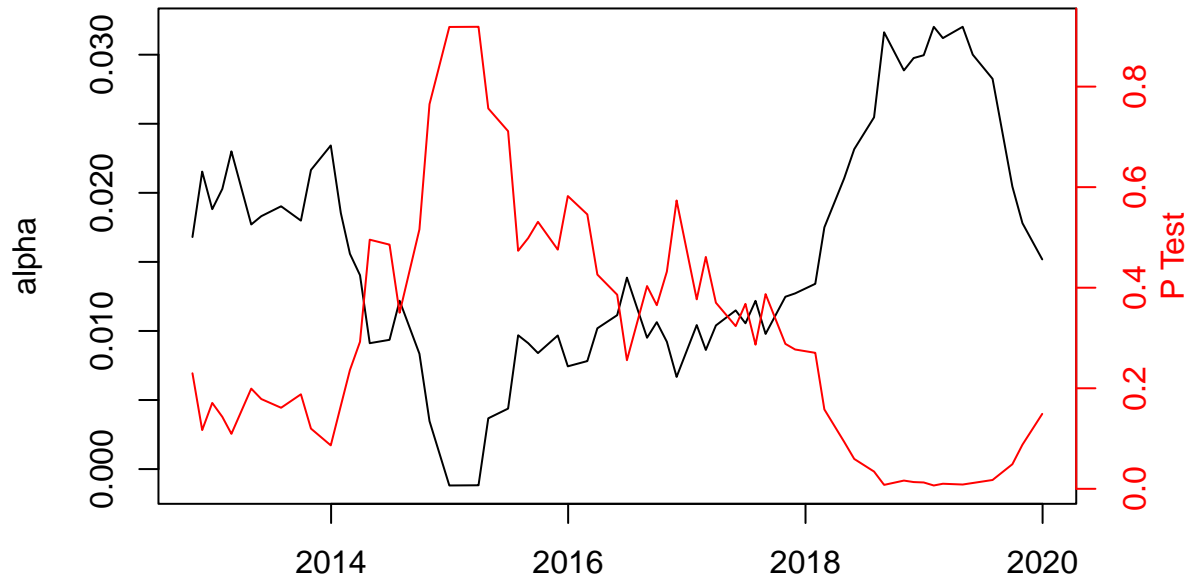
```

Série des coefficients α des régressions sur séries chronologiques et p values correspondantes.

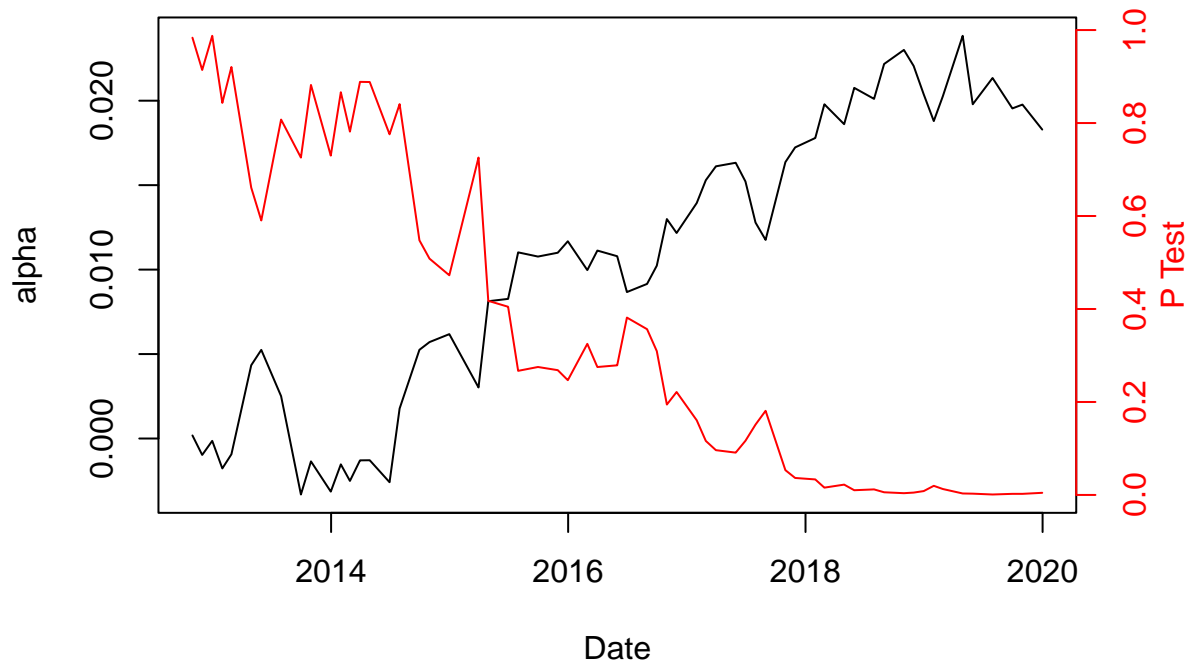
Modèle à un facteur AAPL



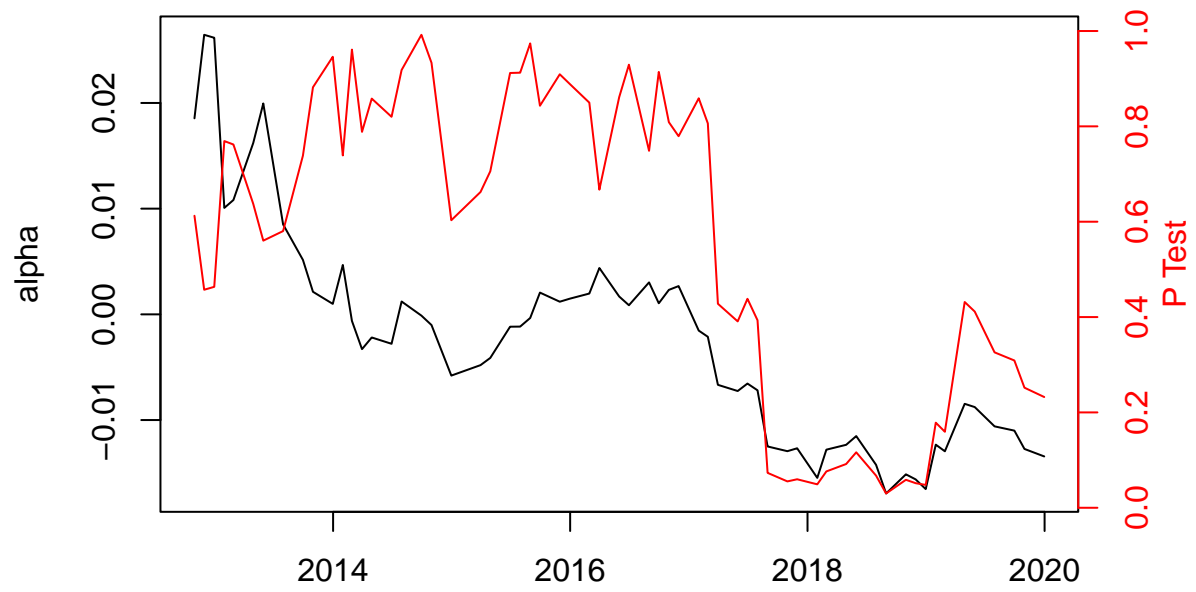
Modèle à un facteur AMZN



Modèle à un facteur MSFT



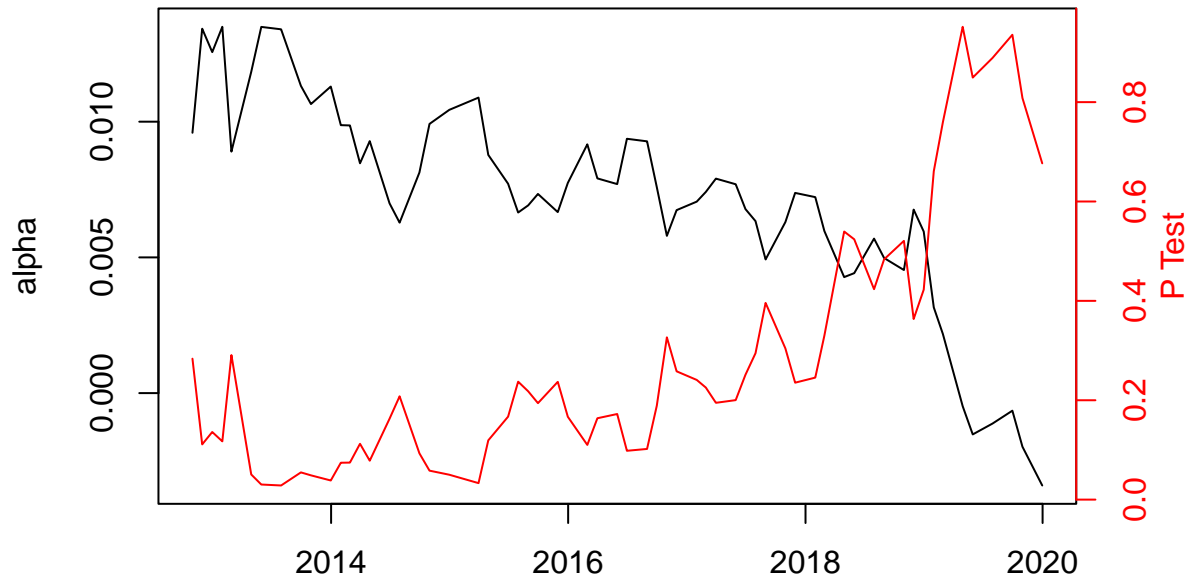
Modèle à un facteur F



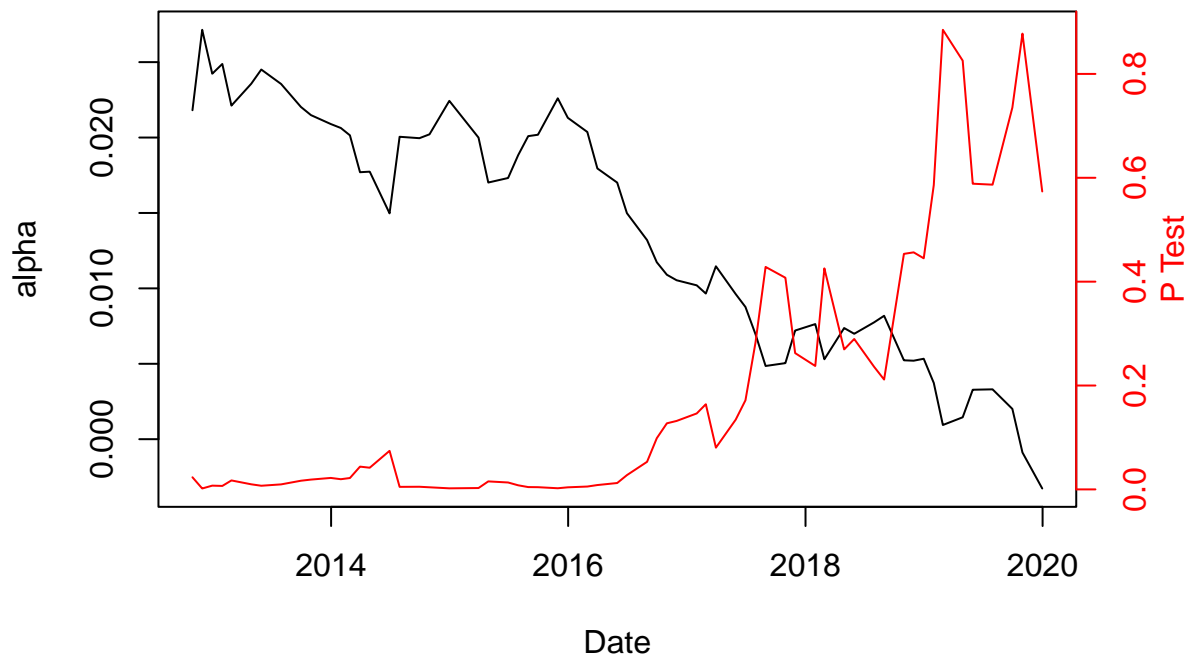
Modèle à un facteur XOM



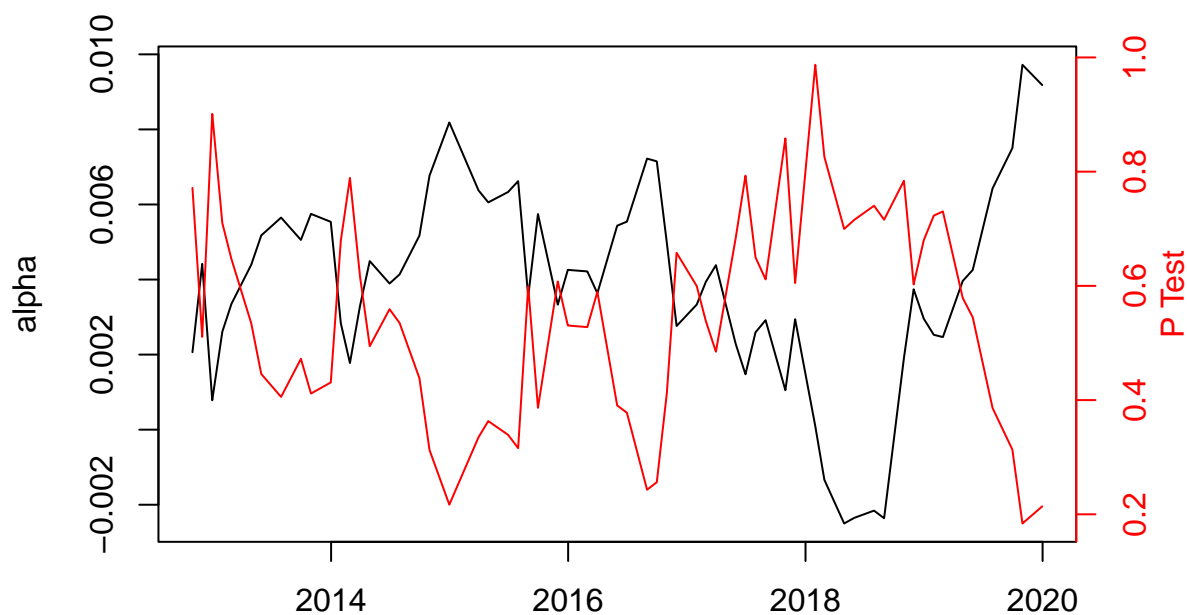
Modèle à un facteur MMM



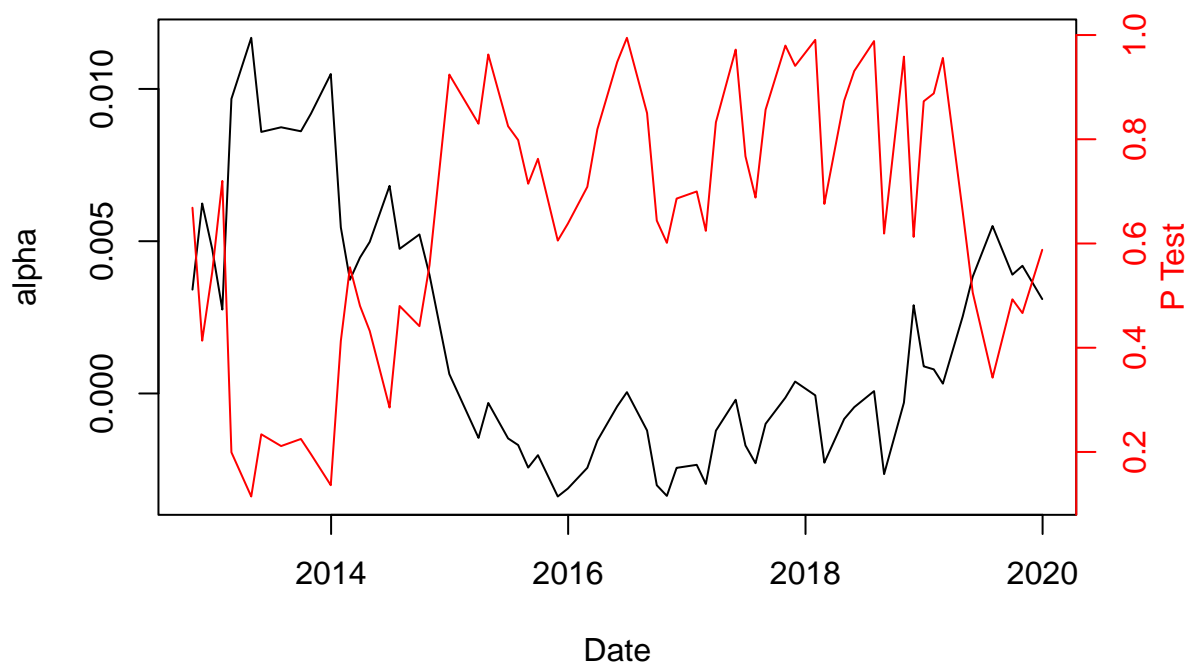
Modèle à un facteur HD



Modèle à un facteur PG



Modèle à un facteur KO



On observe que quelques titres (AMZN, PG) ont un alpha relativement stationnaire, globalement centré sur 0. Pour les autres titres, alpha semble présenter une tendance qui rend l'interprétation hasardeuse.

Calcul de la droite de marché

Calculons la droite de marché pour quelques dates, en utilisant les estimations de β faites plus haut.

```

ggplotRegression <- function (fit) {

  ggplot(fit$model, aes_string(x = names(fit$model)[2], y = names(fit$model)[1])) +
    geom_point() +
    stat_smooth(method = "lm", col = "red") +
    labs(title = paste("Adj R2 = ", signif(summary(fit)$adj.r.squared, 5),
                      " Intercept = ", signif(fit$coef[[1]], 5 ),
                      " Slope = ", signif(fit$coef[[2]], 5),
                      " P = ", signif(summary(fit)$coef[2,4], 5))) +
    geom_text(label=Assets, check_overlap = T)
}

```

```

# liste des dates de calcul
dt.calc.u <- as.Date(unique(df.res$dt.calc))

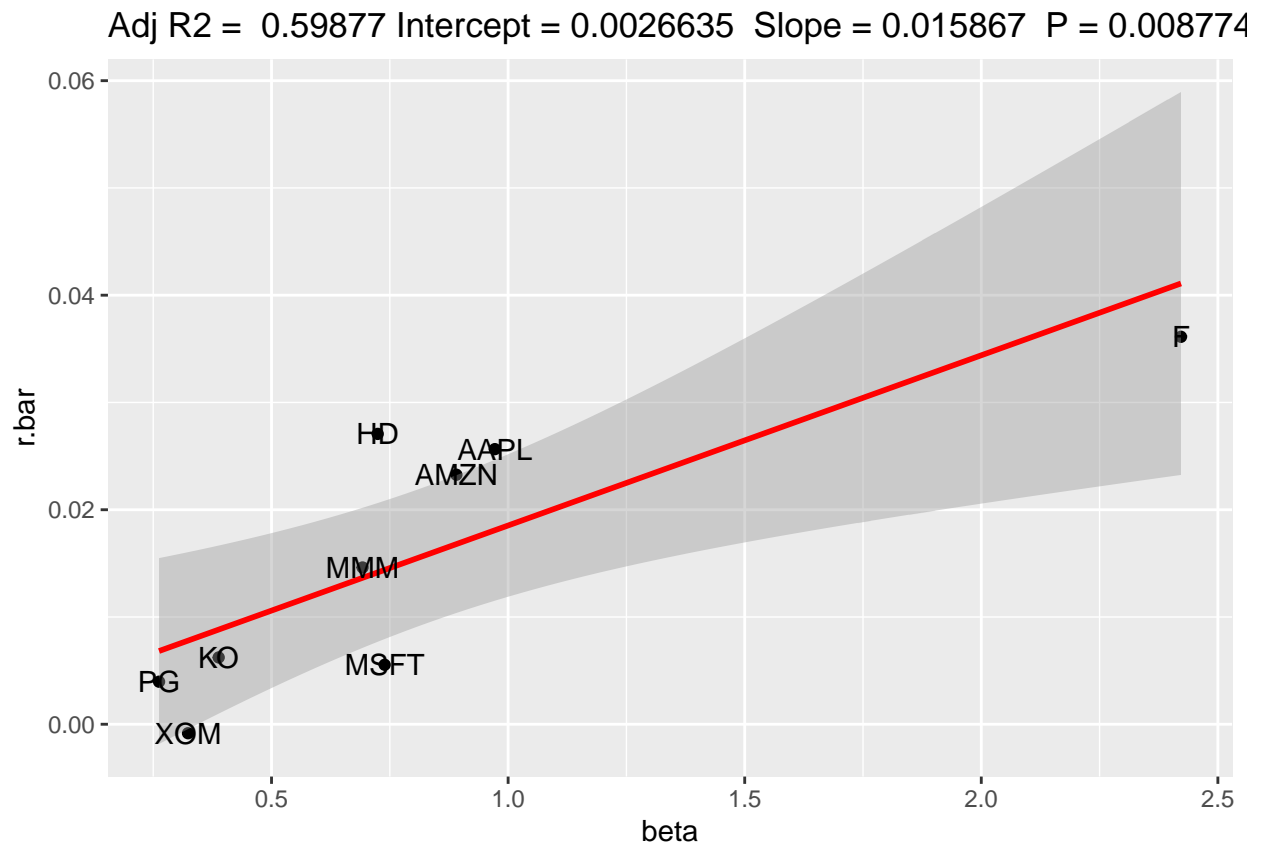
# selection de quelques dates
dt <- first(dt.calc.u)
df.one <- df.res[df.res$dt.calc == dt, ]
# regression Rdt sur beta
dmt <- lm(r.bar ~beta, data=df.one)
ggplotRegression(dmt)

```

Warning: 'aes_string()' was deprecated in ggplot2 3.0.0.

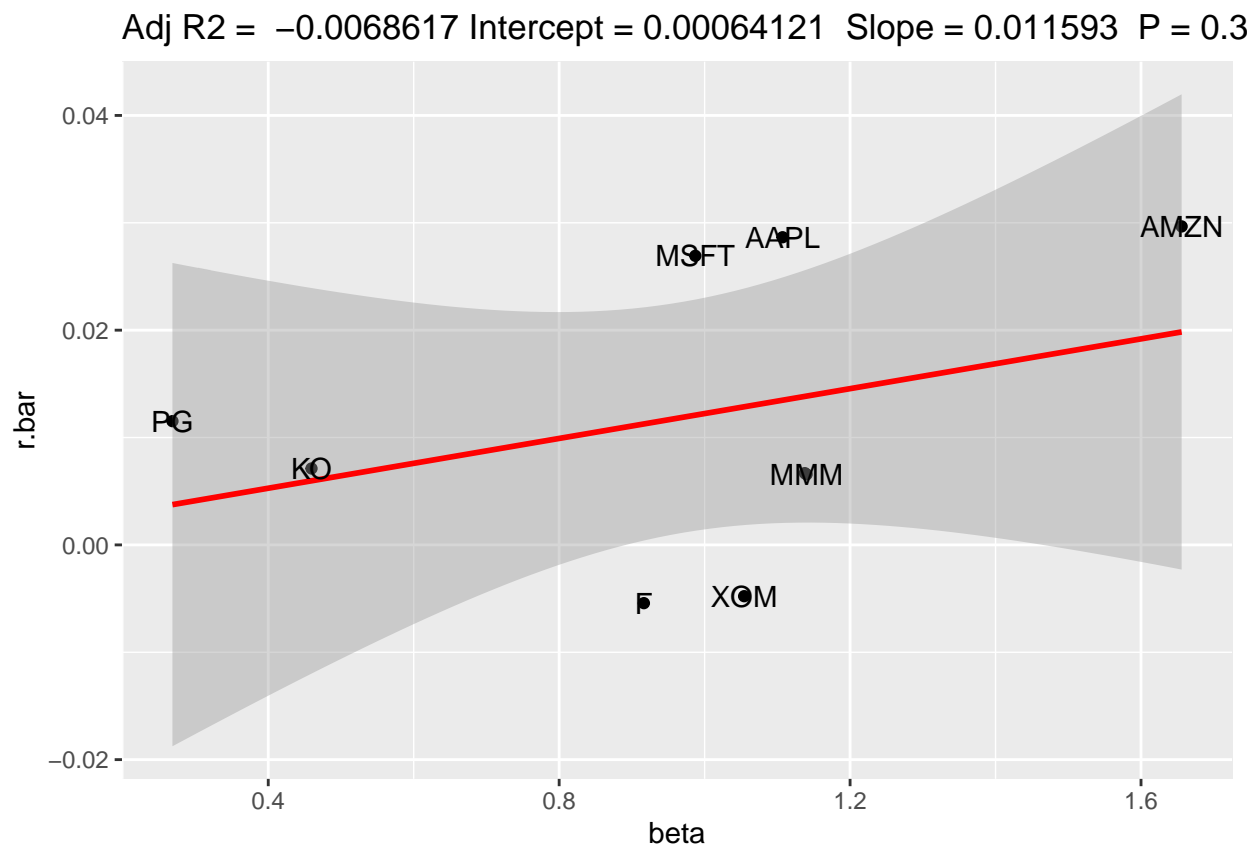
i Please use tidy evaluation ideoms with 'aes()'

'geom_smooth()' using formula = 'y ~ x'



```
dt <- last(dt.calc.u)
df.one <- df.res[df.res$dt.calc == dt, ]
# regression Rdt sur beta
dmt <- lm(r.bar ~beta, data=df.one)
ggplotRegression(dmt)
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```



Ces diagrammes permettent de visualiser les titres dans un espace β / rendement. L'intérêt principal de cette représentation est de pouvoir observer la place relative des titres:

- le β d'AMZN a augmenté significativement entre 2012 et 2019, bien plus que les autres GAFAM
- Conformément aux attentes, le β de KO et PG restent faibles et stables.

La qualité de l'ajustement statistique est par contre très variable.

On calcule enfin la droite de marché des titres pour chaque date de calcul, et on forme des séries chronologiques de résidus (écart entre l'excès de rendement et le rendement prédit par l'exposition au marché)

```
df.dmt <- NULL
col.names <- c(Assets, 'dt.calc')

for(dt in dt.calc.u) {
  df.one <- df.res[df.res$dt.calc == dt, ]
```

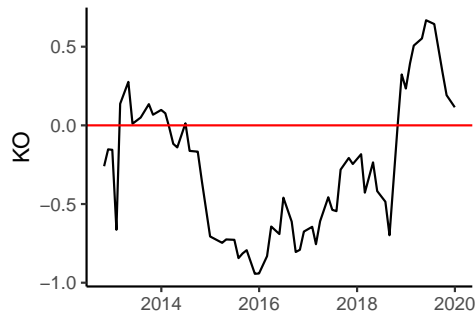
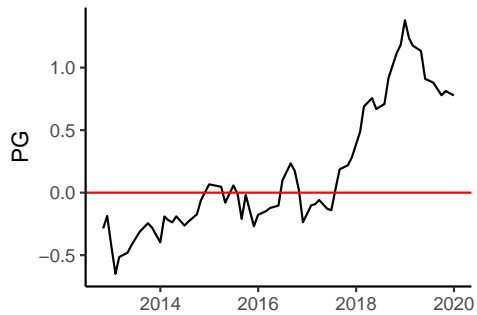
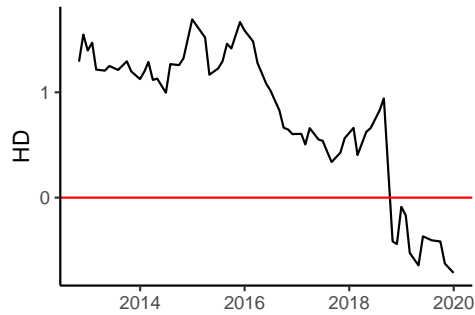
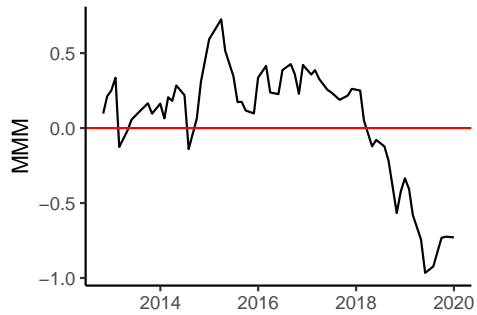
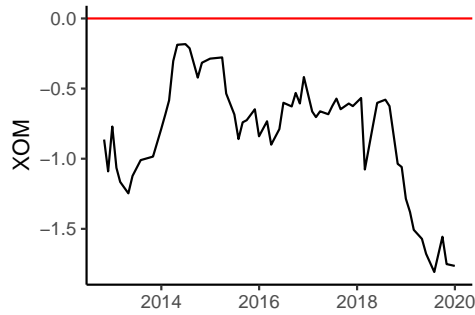
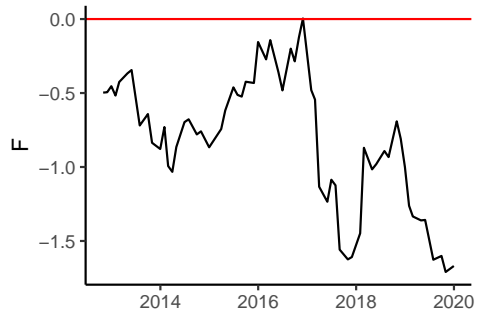
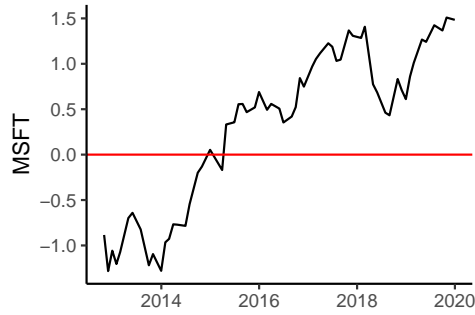
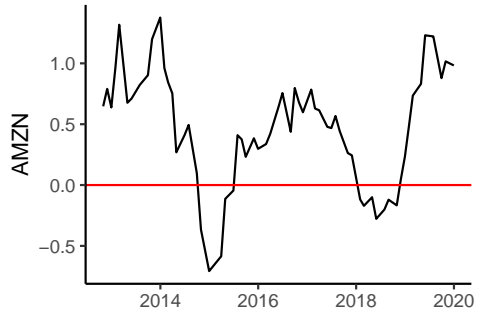
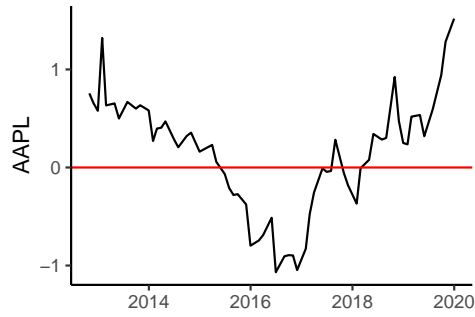
```

# regression Rdt sur beta
dmt <- lm(r.bar ~beta, data=df.one)

tmp <- c(dmt$residuals*100, dt)
names(tmp) = col.names
df.tmp <- data.frame(t(tmp))
if(is.null(df.dmt))
  df.dmt <- df.tmp
else
  df.dmt <- rbind(df.dmt, df.tmp)
}

```

Ecart à la droite de marché



Ces séries chronologiques montrent que l'écart de rendement par rapport à la droite de marché n'est en général pas stationnaire, ce qui suggère que des facteurs de risque importants ne sont pas pris en compte. En conclusion, l'aspect le plus pertinent du modèle à un facteur semble être la droite de marché des titres, qui permet une comparaison instantanée des titres dans l'espace β / rendement.