Magister en Finanzas Full Time

Universidad de Chile

Primavera 2022

Tarea 2

Curso: Inversiones

Profesor: Erwin Hansen, Ph.D.

Ayudante: Gabriel Cabrera.

La siguiente tarea tiene por objetivo principal familiarizar al alumno con la estimación del premio por riesgo de mercado y con la identificación de sus principales drivers usando datos para U.S. Los ejercicios que se realizarán involucran tanto evaluaciones estadísticas como evaluaciones económicas (portafolios) de las estimaciones realizadas. Los datos necesarios se encuentran en el archivo **PredictorData2021.xlsx**.

- La fecha de entrega es el miércoles 28 de septiembre del 2022 hasta las 23:59 hrs.
- La tarea debe realizarse en grupos **2-3 estudiantes**.
- Estos informes deben ser enviados al email **ehansen@fen.uchile.cl** con copia a **gcabrerag@fen.uchile.cl** con el asunto "KfW Inversiones Tarea 2 + Apellidos"
- Se espera recibir un informe de la tarea donde se presenten y discutan los resultados obtenidos, y el código utilizado con sus estimaciones.

Parte 1: Antes de ajustar cualquier modelo de predicción es necesario generar y definir tanto la variable de interés a predecir como los predictores.

a) Genere el logaritmo del exceso de retorno (*equity premium*) como:

$$r_t = log(1 + CRSP_t^{SPvw}) - log(1 + Rf_{t-1})$$

Donde $CRSP_t^{SPvw}$ es el *equity premium* incluyendo dividendos en t y Rf_{t-1} la tasa libre de riesgo (***rfree***) en $t-1$.

b) Basándose en Welch y Goyal (2008) construya y/o seleccione los siguientes predictores:

   i. **Log dividend-price ratio** (*log_dp*): Logaritmo de la suma móvil de 12 meses de los dividendos (***d12***) pagados por el *S&P 500 index* menos el logaritmo del *S&P 500 index* (***sp500***).

ii. **Log dividend yield** (*log_dy*): Logaritmo de la suma móvil de 12 meses de los dividendos (***d12***) pagados por el *S&P 500 index* menos el logaritmo del rezago del *S&P 500 index* (***sp500***).

iii. **Log earnings-price ratio** (*log_ep*): Logaritmo de la suma móvil de 12 meses de las ganancias (***e12***) del *S&P 500 index* menos el logaritmo del *S&P 500 index* (***sp500***).

iv. **Log dividend-payout ratio** (*log_de*): Logaritmo de la suma móvil de 12 meses de los dividendos (***d12***) pagados por el *S&P 500 index* menos el logaritmo de la suma móvil de 12 meses de las ganancias (***e12***) del *S&P 500 index*.

v. **Stock variance** (***svar***): Suma mensual de los retornos diarios al cuadrado del *S&P 500 index*.

vi. **Book-to-market ratio** (***bm***): Valor del ratio *book-to-market* del DJIA (*Dow Jones Industrial Average*).

vii. **Net equity expansion** (***ntis***): Relación entre la suma móvil de 12 meses de las emisiones netas de acciones cotizadas en la NYSE y la capitalización bursátil total anual de las acciones del NYSE.

viii. **Treasury bill rate** (***tbl***): Tasa de interés del *treasury bill* a tres meses (mercado secundario).

ix. **Long-term yield** (***tly***): Rendimiento a largo plazo de los bonos del gobierno.

x. **Long-term return** (***ltr***): Retorno a largo plazo de los bonos del gobierno.

xi. **Term spread** (*tms*): Rendimiento a largo plazo de los bonos del gobierno (***lty***) menos la tasa de interés del *treasury bill* a tres meses (***tbl***).

xii. **Default yield spread** (*dfy*): Diferencia entre el rendimiento de los bonos corporativos del tipo BAA (***baa***) y AAA (***aaa***).

xiii. **Default return spread** (*dfr*): Retorno a largo plazo de los bonos corporativos (***corpr***) menos el retorno a largo plazo de los bonos del gobierno (***ltr***).

xiv. **Inflation** (*infl_lag*): Rezago de la inflación (***infl***) calculado desde *Consumer Price Index* (*CPI, all urban consumers*).

Los nombres en paréntesis, **negrita** y *cursiva* es el nombre que tiene la variable en la base de datos. Los nombres en azul son predictores que deben ser creados. En total son 14 predictores.

c) Basándose en Neely et al. (2014) construya los siguientes indicadores técnicos:

i. **Media Móvil (MA)**:

$$S_{i,t} = \begin{cases} 1 & \text{si} \quad MA_{s,t} \geq MA_{l,t} \\ 0 & \text{si} \quad MA_{s,t} < MA_{l,t} \end{cases}$$

Donde

$$MA_{j,t} = (1/j)\sum_{i=0}^{j-1} P_{t-i} \quad \text{para} \quad j = s, l$$

Acá P_t es el nivel del *S&P 500 index* (***sp500***), y $s(l)$ el tamaño del MA *short* (*long*). Construya el indicador con $s = 1,2,3$ y $l = 9,12$.

ii. **Momentum (MOM)**:

$$S_{i,t} = \begin{cases} 1 & \text{si} \quad P_t \geq P_{t-m} \\ 0 & \text{si} \quad P_t < P_{t-m} \end{cases}$$

Donde m son los periodos atrás, que indica un *momentum* positivo si el precio actual es mayor al precio en m. Construya el indicador para $\mathrm{m} = 9,12$.

iii. **Volume (OBV)**:

$$OBV_t = \sum_{k=1}^{t} VOL_k D_k$$

Donde VOL_k es el volumen transado durante el periodo k y D_k una variable binaria igual a 1 si $P_k - P_{k-1} \geq 0$ y -1 en caso contrario. A partir de lo anterior, se forma la señal OBV_t como:

$$S_{i,t} = \begin{cases} 1 & \text{si} \quad MA_{s,t}^{OBV} \geq MA_{l,t}^{OBV} \\ 0 & \text{si} \quad MA_{s,t}^{OBV} < MA_{l,t}^{OBV} \end{cases}$$

Donde:

$$MA_{j,t}^{OBV} = (1/j)\sum_{i=0}^{j-1} OBV_{t-i} \quad \text{para} \quad j = s, l$$

Construya el indicador con $s = 1,2,3$ y $l = 9,12$.

La hoja "Tech" contiene las variables necesarias para generar los indicadores técnicos.

d) Realice una breve estadística descriptiva de los predictores y respuesta (*equity premium*) creados en (a), (b) y (c) desde 1968:12 hasta 2021:12. Debe incluir la media muestral, desviación estándar, *skewness*, *kurtosis* y correlación de primer orden.

Parte 2: Realice mediante una regresión bivariada por OLS la predicción del logaritmo del exceso de retorno (*equity premium*) fuera de muestra (*out-of-sample*) utilizando individualmente cada predictor creado en la **Parte 1**. Considerando lo anterior, el modelo que se debe estimar mediante OLS es:

$$r_{t+1} = \alpha_i + \beta_i q_{i,t}^{j} + \varepsilon_{i,t+1} \quad \text{para} \quad j = Econ, Tech$$

Donde r_{t+1} es el logaritmo del exceso de retorno en $\mathrm{t}+1$ y $q_{i,t}^{j}$ es el predictor $\mathrm{i}-$esimo en t para los predictores macroeconómicos ($\mathrm{j} = \mathrm{Econ}$) e indicadores técnicos ($j = Tech$). Debe utilizar una muestra *in-sample* inicial desde 1968:12 hasta 1989:12 (en la práctica comienza desde 1969:01, debido que el predictor está rezagado un periodo y se "pierde" la observación del periodo 1968:12). La predicción en la muestra *out-of-sample* debe ser utilizando una ventana recursiva/expandida (*recursive/expanded window*) a un mes.

a) Suponga que una muestra de T observaciones está disponible para r_t y $x_{i,t-1}$. Si dividimos la muestra total en *in-sample* utilizando las primeras n_1 observaciones y para *out-of-sample* $n_2 = T - n_1$ podemos calcular el siguiente estadístico (Campbell y Thompson 2008):

$$R_{os}^2 = 1 - \frac{(1/n_2)\sum_{s=1}^{n_2}\left(r_{n_1+s} - \hat{r}_{i,n_1+s}\right)^2}{(1/n_2)\sum_{s=1}^{n_2}\left(r_{n_1+s} - \overline{r}_{n_1+s}\right)^2} = 1 - \frac{MSFE_i}{MSFE_0}$$

La finalidad del estadístico R_{os}^2 es verificar si la predicción realizada por una predictor o modelo determinado ($MSFE_i$) es superior a otro predictor o modelo, este último es utilizado como *benchmark* ($MSFE_0$), de manera tal que cuando $R_{os}^2 > 0$, la regresión predictiva es más precisa que el *benchmark*.

Aplique el estadístico R_{os}^2 a las predicciones realizadas considerando como *benchmark* la media histórica (también estimada *out-of-sample* con utilizando el esquema de *recursive window*). ¿Qué predictores logran ganarle a la media histórica?

b) Verifique la significancia del R_{os}^2 utilizando el test *MSFE-adjusted* de Clark y West (2007). Para generar el test se debe definir primero:

$$\tilde{d}_{i,n_i+s} = \hat{u}_{0,n_1|s}^2 - \left[\hat{u}_{i,n_1|s}^2 - \left(\overline{r}_{n_1+s} - \hat{r}_{i,n_1+s}\right)^2\right]$$

Donde $\hat{u}^2_{0,n_1|s}$ es el error de predicción al cuadrado utilizando la media histórica y $\hat{u}^2_{i,n_1|s}$ el error de predicción realizada por una predictor o modelo. Luego se debe regresionar $\tilde{d}_{i,n_i+s}$ con respecto a una constante, siendo el estadístico t de la regresión anterior equivalente al *MSFE-adjusted*. Finalmente, el contraste de hipótesis que se debe testear es:

$$H_0: MSFE_0 \leq MSFE_i \equiv R^2_{os} \leq 0$$

$$H_a: MSFE_0 > MSFE_i \equiv R^2_{os} > 0$$

¿Qué predictores son estadísticamente significativos? ¿Existe consistencia con lo obtenido en (2.a)?

Parte 3: Para cada predicción generada[1] en (2.a), construya un portafolio asumiendo a un inversionista averso al riesgo que al final del periodo t, asignará óptimamente una porción w_t en *equities* y $1 - w_t$ en la tasa libre de riesgo durante el mes $t + 1$:

$$w_t = \left(\frac{1}{\gamma}\right)\left(\frac{\hat{r}_{t+1}}{\hat{\sigma}^2_{t+1}}\right)$$

Donde $\hat{r}_{t+1}$ es la predicción del exceso de retorno, $\hat{\sigma}^2_{t+1}$ es la predicción de la varianza y γ la aversión al riesgo. El retorno del portafolio en $t + 1$ es:

$$R_{p,t+1} = w_t r_{t+1} + R_{f,t+1}$$

Siguiendo a Campbell y Thompson (2008) utilice la varianza poblacional móvil de 60 meses como predictor de la varianza, una aversión al riesgo de 5 e imponga una restricción para que w_t se encuentre entre 0 y 1.5.

a) Calcule el ratio de Sharpe (*Sharpe*) para cada portafolio creado.

$$Sharpe_p = \frac{\frac{1}{T}\sum_{t=1}^{T}\left(R_{p,t} - R_{f,t}\right)}{\sqrt{\frac{1}{T-1}\sum_{t=1}^{T}\left(R_{p,t} - \overline{R}_p\right)^2}} = \frac{\hat{\mu}_p}{\sqrt{\hat{\sigma}^2_p}}$$

Donde $\hat{\mu}_p$ y $\hat{\sigma}^2_p$ es la media y varianza del portafolio. ¿Cuál es el valor anualizado? Interprete

b) Calcule el *Downside Risk* (*DR*) o *Semi-Standard Deviation* para cada portafolio creado.

[1] Debe incluir la media histórica como *benchmark*.

$$DR_p = \left[(1/T) \sum_{t=1}^{T} \min\left(R_{p,t}, 0\right)^2 \right]^{0.5}$$

Donde $R_{p,t}$ es el retorno del portafolio en t y T el total de las observaciones fuera de muestra (*out-of-sample*). ¿Cuál es el valor anualizado? Interprete.

c) Calcule el ratio de Sortino (*Sortino*) para cada portafolio creado.

$$Sortino_p = \frac{\frac{1}{T}\sum_{t=1}^{T} R_{p,t}}{DR_p} = \frac{\hat{\mu}_p^*}{DR_p}$$

Donde $\hat{\mu}_p^*$ es el retorno promedio del portafolio y DR_p el *Downside Risk* (*DR*) o *Semi-Standard Deviation* del portafolio. ¿Cuál es el valor anualizado? Interprete.

d) Calcule el *Maximum Drawdown* (*MDD*) para cada portafolio creado.

$$MDD_p = -1 \times \min \left[\frac{f\left(1 + R_{p,t}\right)}{g\left(f\left(1 + R_{p,t}\right)\right)} - 1 \right] \quad \text{para} \quad t = 1, \ldots, T$$

Donde $f(.)$ representa el producto acumulado y $g(.)$ el máximo acumulado. ¿Cuál es el valor anualizado? Interprete.

e) Para cada portafolio creado grafique los retornos acumulados (gráfico de línea) y las ponderaciones (w_t) correspondientes (gráfico de barra). Interprete.

Referencias

1) Campbell, J.Y. y Thompson, S.B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? The Review of Financial Studies, 21(4), pp.1509-1531.
2) Clark, T.E. y West, K.D. (2007). Approximately normal tests for equal predictive accuracy in nested models. Journal of econometrics, 138(1), pp.291-311.
3) Neely, C.J. et al. (2014). Forecasting the equity risk premium: the role of technical indicators. Management science, 60(7), pp.1772-1791.
4) Welch, I. y Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. The Review of Financial Studies, 21(4), pp.1455-1508.