Assignment 2

Yongseok Kim

March 23, 2020

1. (a) See Table 1 for the results.

   (b) See Table 2 for the results. In contrast to positive in-sample R^2 statistics, we obtain negative out-of-sample R^2 statistics, which indicates that the predictive regression has *greater* average mean-squared prediction error than the historical average return. This result is consistent with Welch and Goyal's (2008) claim: the predictive regression is lack of out-of-sample predictability.

   (c) See Table 3 for the results. Following Campbell and Thompson's (2008) modification, we obtain positive out-of-sample R^2 statistic for the dividend-price ratio, which indicates that the predictive regression has *lower* average mean-squared prediction error than the historical average return. We can conclude that there is an inprovement in the out-of-sample explanatory power of the dividend-price ratios, though I do not derive a statistical inference for significance here.

   (d) See Appendix for a code to construct our new predictor x. Until now, we estimate

$$\begin{aligned}\hat{\mathbb{E}}_t g_{t+1} &= \frac{1}{t}\sum_{s=1}^{t} g_s \\ &= \hat{\mathbb{E}}_{t-1} g_t + t^{-1}[g_t - \hat{\mathbb{E}}_{t-1} g_t]\end{aligned}$$

$$\begin{aligned}\widehat{\text{var}}_t(r_{t+1}) &= \frac{1}{t}\sum_{s=1}^{t}(r_s - \bar{r}_s)^2 \\ &= \widehat{\text{var}}_{t-1}(r_t) + t^{-1}[(r_t - \bar{r}_t)^2 - \widehat{\text{var}}_{t-1}(r_t)]\end{aligned}$$

   which can be justified as a least squared learning using a perceived law of motion

$$\begin{aligned} g_t &= \mu_g + \varepsilon_t \\ \varepsilon_t &=_{i.i.d} (0, \sigma_\varepsilon^2)\end{aligned}$$

Table 1: In-sample R^2 statistics

predictor	rsq
dp	0.038
xp	0.043

Table 2: Out-of-sample R^2 statistics
In-sample R^2 statistics are calculated using the out-of-sample period observations.

predictor	rsq_os	rsq_is
dp	-0.015	0.065
xp	-0.012	0.072

Table 3: Campbell and Thompson's (2008) modification

predictor	rsq_os
dp	0.020
xp	-0.012

$$r_t = \mu_r + \nu_t$$
$$\nu_t =_{i.i.d} (0, \sigma_\nu^2)$$

We can consider alternative procedures by modifying a learning or a perceived law of motion or both. As an alternative procedure, we may assume normality of unknown parameters and a prior distribution. Then, by the Bayes update, we can obtain posterior distribution of parameters at each t and consider moments from posterior as our conditional expectations at t.

(e) See Table 4 for the repeated calculations of part b) and c) for x_t. Note that the sign restriction is also reasonable in this predictive regression. The repeated calculation improves the forecasting performance of part b). Moreover, the performance of part c) is also improved compared to part b).

(f) See Table 5 for the results. At the first glimpse, the out-of-sample performance of the predictive regression looks bad, suggesting it may not be useful for real-time forecasting. From this exercise, however, we can learn that the out-of-sample performance of the predictive regression can be improved with theoretical restrictions.

2. (a) Note that

$$\begin{aligned}
\mathrm{var}_t^* R_{i,t+1} &= \mathbb{E}_t^* R_{i,t+1}^2 - (\mathbb{E}_t^* R_{i,t+1})^2 \\
&= R_{f,t+1}\mathbb{E}_t[M_{t+1}R_{i,t+1}^2] - R_{f,t+1}^2
\end{aligned}$$

$$\begin{aligned}
\mathrm{cov}_t(M_{t+1}R_{i,t+1}, R_{i,t+1}) &= \mathbb{E}_t[M_{t+1}R_{i,t+1}^2] - \mathbb{E}_t[M_{t+1}R_{i,t+1}]\mathbb{E}_t[R_{i,t+1}] \\
&= \mathbb{E}_t[M_{t+1}R_{i,t+1}^2] - \mathbb{E}_t[R_{i,t+1}]
\end{aligned}$$

Table 4: The adjusted version of the dividend yield

part	rsq_os	rsq_is
b)	-0.008	0.065
c)	0.007	

Table 5: Fully imposed theoretical restrictions

restriction	rsq_os
a = -1, b = 1	0.022

Then

$$\begin{aligned}\frac{\mathrm{var}_t^* R_{i,t+1}}{R_{f,t+1}} &= \mathbb{E}_t[R_{i,t+1}] - R_{f,t+1} + \mathbb{E}_t[M_{t+1}R_{i,t+1}^2] - \mathbb{E}_t[R_{i,t+1}] \\ &= \mathbb{E}_t[R_{i,t+1}] - R_{f,t+1} + \mathrm{cov}_t(M_{t+1}R_{i,t+1}, R_{i,t+1})\end{aligned}$$

If

$$\mathrm{cov}_t(M_{t+1}R_{i,t+1}, R_{i,t+1}) \leq 0$$

we obtain

$$\begin{aligned}R_{f,t+1}SVIX_{i,t+1}^2 &= \frac{\mathrm{var}_t^* R_{i,t+1}}{R_{f,t+1}} \\ &= \mathbb{E}_t[R_{i,t+1}] - R_{f,t+1} + \mathrm{cov}_t(M_{t+1}R_{i,t+1}, R_{i,t+1}) \\ &\leq \mathbb{E}_t[R_{i,t+1}] - R_{f,t+1}\end{aligned}$$

(b) i. If a risk-neutral distribution of return follows the conditional lognormal distribution, i.e. $\log R_{i,T}|I_t =_d \mathbb{N}(r_{f,T} - \frac{1}{2}\sigma_T^2, \sigma_T^2)$, we have

$$\log \frac{R_{i,T}}{R_{f,T}}|I_t =_d \mathbb{N}(-\frac{1}{2}\sigma_T^2, \sigma_T^2)$$

$$\log \mathbb{E}_t^*(\frac{R_{i,T}}{R_{f,T}}) = \mathbb{E}_t^*(\log \frac{R_{i,T}}{R_{f,T}}) + \frac{1}{2}\mathrm{var}_t^*(\log \frac{R_{i,T}}{R_{f,T}})$$

Then

$$\begin{aligned}VIX_{i,t,T}^2 &= \frac{2}{T-t}L_t^*(\frac{R_{i,T}}{R_{f,T}}) \\ &= \frac{2}{T-t}[\log \mathbb{E}_t^*(\frac{R_{i,T}}{R_{f,T}}) - \mathbb{E}_t^* \log \frac{R_{i,T}}{R_{f,T}}] \\ &= \frac{2}{T-t}[\frac{1}{2}\mathrm{var}_t^*(\log \frac{R_{i,T}}{R_{f,T}})] \\ &= \frac{1}{T-t}\sigma_T^2\end{aligned}$$

Moreover

$$\begin{aligned}SVIX_{i,t,T}^2 &= \frac{1}{T-t}\frac{1}{R_{f,T}^2}\mathrm{var}_t^*(R_{i,T}) \\ &= \frac{1}{T-t}\frac{1}{R_{f,T}^2}(\exp \sigma_T^2 - 1)\exp\left(2(r_{f,T} - \frac{1}{2}\sigma_T^2) + \sigma_T^2\right) \\ &= \frac{1}{T-t}(\exp \sigma_T^2 - 1)\end{aligned}$$

For sufficiently small $T-t$ and σ_T^2, we obtain

$$\exp \sigma_T^2 - 1 \approx \sigma_T^2$$

and

$$VIX_{i,t,T}^2 \approx SVIX_{i,t,T}^2$$

ii. First, by Stein's Lemma, we have

$$\begin{aligned}
\mathrm{cov}_t(e^{m_{t+1}}, r_{i,t+1}) &= \mathbb{E}[e^{m_{t+1}}]\mathrm{cov}_t(m_{t+1}, r_{i,t+1}) \\
&= \frac{1}{R_{f,t+1}}\mathrm{cov}_t(m_{t+1}, r_{i,t+1})
\end{aligned}$$

Then

$$\begin{aligned}
\mathbb{E}_t M_{t+1} r_{i,t+1} &= \mathrm{cov}_t(e^{m_{t+1}}, r_{i,t+1}) + \mathbb{E}_t M_{t+1} \mathbb{E}_t r_{i,t+1} \\
&= \frac{1}{R_{f,t+1}}[\mathrm{cov}_t(m_{t+1}, r_{i,t+1}) + \mathbb{E}_t r_{i,t+1}]
\end{aligned}$$

Second, from $\mathbb{E}_t M_{t+1} R_{i,t+1} = 1$, we have

$$\begin{aligned}
\log \mathbb{E}_t M_{t+1} R_{i,t+1} &= \mathbb{E}_t m_{t+1} + \mathbb{E} r_{i,t+1} + \frac{1}{2}\mathrm{var}_t(m_{t+1} + r_{i,t+1}) \\
&= \mathbb{E}_t r_{i,t+1} + \frac{1}{2}\mathrm{var}_t r_{i,t+1} + \mathbb{E}_t m_{t+1} + \frac{1}{2}\mathrm{var}_t m_{t+1} + \mathrm{cov}(m_{t+1}, r_{i,t+1}) \\
&= 0
\end{aligned}$$

where the first equality holds since $m_{t+1} + r_{i,t+1}$ is linear transformation of joint normal distribution. Moreover

$$\begin{aligned}
\log \mathbb{E}_t M_{t+1} &= \mathbb{E}_t m_{t+1} + \frac{1}{2}\mathrm{var}_t m_{t+1} \\
&= -r_{f,t+1}
\end{aligned}$$

Combining these results, we obtain

$$\begin{aligned}
VIX_{i,t,t+1}^2 &= 2L_t^*(\frac{R_{i,t+1}}{R_{f,t+1}}) \\
&= 2[\log \mathbb{E}_t^*(\frac{R_{i,t+1}}{R_{f,t+1}}) - \mathbb{E}_t^* \log \frac{R_{i,t+1}}{R_{f,t+1}}] \\
&= -2[R_{f,t+1}\mathbb{E}_t M_{t+1} r_{i,t+1} - r_{f,t+1}] \\
&= -2[\mathrm{cov}_t(m_{t+1}, r_{i,t+1}) + \mathbb{E}_t r_{i,t+1}] + 2r_{f,t+1} \\
&= -2[r_{f,t+1} - \frac{1}{2}\mathrm{var}_t r_{i,t+1}] + 2r_{f,t+1} \\
&= \mathrm{var}_t r_{i,t+1}
\end{aligned}$$

which is a desired expression. To obtain the expression for $SVIX_{i,t,t+1}^2$, note

$$\begin{aligned}
SVIX_{i,t,t+1}^2 &= \mathrm{var}_t^*(\frac{R_{i,t+1}}{R_{f,t+1}}) \\
&= \frac{1}{R_{f,t+1}^2}\mathbb{E}_t^*[R_{i,t+1}^2] - \mathbb{E}_t^*[\frac{R_{i,t+1}}{R_{f,t+1}}]^2 \\
&= \frac{1}{R_{f,t+1}}\mathbb{E}_t[M_{t+1}R_{i,t+1}^2] - 1
\end{aligned}$$

By the normality, we have

$$\begin{aligned}
\log \mathbb{E}_t[M_{t+1}R_{i,t+1}^2] &= \mathbb{E}_t[m_{t+1} + 2r_{i,t+1}] + \frac{1}{2}\mathrm{var}_t(m_{t+1} + 2r_{i,t+1}) \\
&= \mathbb{E}_t m_{t+1} + \frac{1}{2}\mathrm{var}_t m_{t+1} + 2\mathbb{E}_t r_{i,t+1} + 2\mathrm{var}_t r_{i,t+1} + 2\mathrm{cov}_t(m_{t+1}, r_{i,t+1}) \\
&= \mathrm{var}_t r_{i,t+1} - (\mathbb{E}_t m_{t+1} + \frac{1}{2}\mathrm{var}_t m_{t+1}) \\
&= \mathrm{var}_t r_{i,t+1} + r_{f,t+1}
\end{aligned}$$

Then

$$\begin{aligned}
SVIX_{i,t,t+1}^2 &= \frac{1}{R_{f,t+1}}\exp[\mathrm{var}_t r_{i,t+1} + r_{f,t+1}] - 1 \\
&= \exp[\mathrm{var}_t r_{i,t+1}] - 1
\end{aligned}$$

iii. From part b) ii), we obtain

$$SVIX_{i,t+1}^2 > VIX_{i,t+1}^2$$

under the lognormal assumption since

$$e^x - 1 > x$$

for any x. Hence, the empirical relationship suggests that the lognormal assumption is not supported in data.

(c) i. The portfolio choice problem for this agent becomes

$$\max_\theta \mathbb{E}_t[W_t(\sum \theta_i R_{i,t+1})]$$

$$\text{s.t. } \langle 1, \theta \rangle = 1$$

By the first-oder condition, we obtain

$$\mathbb{E}_t[W_t R_{i,t+1}] = \lambda$$

suggesting the constant SDF supports (infinitely many) solutions. Hence

$$\begin{aligned}
\mathrm{cov}_t(M_{t+1}R_{m,t+1}, R_{m,t+1}) &= \lambda^{-1}W_t\sigma_{m,t+1}^2 \\
&> 0
\end{aligned}$$

and NCC does not hold.

ii. We need two more assumptions. First, the agent is unconstrained. Second, a risk aversion parameter γ is greater than or equal to 1 for every state. Then, the portfolio choice problem for the agent becomes

$$\max_\theta \mathbb{E}_t u[W_t(\sum \theta_i R_{i,t+1})]$$

$$\text{s.t. } \langle 1, \theta \rangle = 1$$

By the first-order condition, we obtain

$$\mathbb{E}_t u'[W_t(\sum \theta_i R_{i,t+1})]R_{i,t+1} = \lambda$$

for any i. Since the optimal portfolio is the market portfolio, we have

$$\mathbb{E}_t \lambda^{-1} u'[W_t R_{m,t+1}] R_{i,t+1} = 1$$

which is a canonical asset pricing equation where the SDF is

$$M_t = \lambda^{-1} u'[W_t R_{m,t+1}]$$

Now, we want to show

$$\begin{aligned} \mathrm{cov}_t(M_{t+1} R_{m,t+1}, R_{m,t+1}) &= \lambda^{-1} \mathrm{cov}_t(u'[W_t R_{m,t+1}] R_{m,t+1}, R_{m,t+1}) \\ &\leq 0 \end{aligned}$$

Note that

$$\begin{aligned} \frac{\partial}{\partial R_{m,t+1}} u'[W_t R_{m,t+1}] R_{m,t+1} &= u''(W_t R_{m,t+1}) W_t R_{m,t+1} + u'[W_t R_{m,t+1}] \\ &= \underbrace{u'[W_t R_{m,t+1}]}_{\geq 0} \underbrace{(1-\gamma)}_{\leq 0} \\ &\leq 0 \end{aligned}$$

Hence

$$R_{m,t+1} \uparrow \Rightarrow u'[W_t R_{m,t+1}] R_{m,t+1} \downarrow$$

suggesting

$$\mathrm{cov}_t(M_{t+1} R_{m,t+1}, R_{m,t+1}) \leq 0$$

iii. For the brevity, I use lower cases to denote logaritm, i.e. $x = \log X$. First, we have

$$\begin{aligned} \log \mathbb{E}_t M_{t+1} R_{m,t+1} &= \mathbb{E}_t m_{t+1} + \mathbb{E}_t r_{m,t+1} + \frac{1}{2} \mathrm{var}_t(m_{t+1} + r_{m,t+1}) \\ &= \mathbb{E}_t r_{m,t+1} + \frac{1}{2} r_{m,t+1} + \mathbb{E}_t m_{t+1} + \frac{1}{2} \mathrm{var}_t m_{t+1} + \mathrm{cov}_t(m_{t+1}, r_{m,t+1}) \\ &= 0 \end{aligned}$$

Second,

$$\begin{aligned} \log \mathbb{E}_t M_{t+1} &= \mathbb{E}_t m_{t+1} + \frac{1}{2} \mathrm{var}_t m_{t+1} \\ &= -r_f \end{aligned}$$

Combining these results, we obtain

$$\mu_{m,t+1} - r_f + \mathrm{cov}_t(m_{t+1}, r_{m,t+1}) = 0$$

where

$$\begin{aligned} \mu_{m,t+1} &= \log \mathbb{E}_t R_{m,t+1} \\ &= \mathbb{E}_t r_{m,t+1} + \frac{1}{2} r_{m,t+1} \end{aligned}$$

Now, we have

$$\begin{aligned}\log \mathbb{E}_t M_{t+1} R_{m,t+1}^2 &= \mathbb{E}_t m_{t+1} + 2\mathbb{E}_t r_{m,t+1} + \frac{1}{2}\mathrm{var}_t(m_{t+1} + 2r_{m,t+1})\\ &= -r_f + 2\mathbb{E}_t r_{m,t+1} + 2\mathrm{var}_t r_{m,t+1} + 2\mathrm{cov}_t(m_{t+1}, r_{m,t+1})\\ &= \mu_{m,t+1} - r_f + \mathrm{cov}_t(m_{t+1}, r_{m,t+1}) + \mu_{m,t+1} + \mathrm{var}_t r_{m,t+1} + \mathrm{cov}_t(m_{t+1}, r_{m,t+1})\\ &= \mu_{m,t+1} + \mathrm{var}_t r_{m,t+1} + \mathrm{cov}_t(m_{t+1}, r_{m,t+1})\end{aligned}$$

where the first equality holds since linear transformation of joint normal distribution is normally distributed. Then

$$\mathbb{E}_t M_{t+1} R_{m,t+1}^2 \le \mathbb{E}_t R_{m,t+1}$$

$$\begin{aligned}\Leftrightarrow \log \mathbb{E}_t M_{t+1} R_{m,t+1}^2 - \log \mathbb{E}_t R_{m,t+1} &= \mathrm{var}_t r_{m,t+1} + \mathrm{cov}_t(m_{t+1}, r_{m,t+1})\\ &\le 0\end{aligned}$$

Hence, we have

$$\sqrt{\mathrm{var}_t r_{m,t+1}} \le \eta_t$$
$$\Leftrightarrow \text{NCC holds}$$

where

$$\eta_t = \frac{\mu_{m,t+1} - r_f}{\sqrt{\mathrm{var}_t r_{m,t+1}}}$$

Since those assumptions are quite standard in the literature and the assumptions on risk aversion and Sharpe ratio are empirically supported, we can conclude NCC imposes weak conditions.

(d) A satisfaction of the lower bound with equality implies

$$\mathrm{cov}_t(M_{t+1} R_{m,t+1}, R_{m,t+1}) \approx 0$$

which could be achieved when

$$\begin{aligned}\frac{\partial}{\partial R_{m,t+1}} u'[W_t R_{m,t+1}] R_{m,t+1} &= u'[W_t R_{m,t+1}](1-\gamma)\\ &\approx 0\end{aligned}$$

for any $W_t R_{m,t+1}$. Hence, based on the discussion in part c) ii), we can interpret this result as the relative risk aversion of the marginal investor who holds the market is close to 1.

(e) I would like to emphasize two advantages of SVIX approach. First, it is possible to construct dataset at high-frequency. Second, therefore, it is much more volatile than the conventional approach, which would be very helpful to understand high volatility of (short-run) equity premiums. However, one disadvantage of SVIX approach is that its computation heavily relies on option prices with different strikes. Therefore, if the option market is not working well, e.g. illiquid market or narrow range of strike prices, SVIX approach would not be feasible. Moreover, since it is a lower bound, a link between the lower bound and the actual equity premium depends on the level of $\mathrm{cov}_t(M_{t+1} R_{m,t+1}, R_{m,t+1})$ and its serial correlation with equity premium.

(f) From Campbell-Shiller approximation, we obtain the accounting identity

$$d_t - p_t = \kappa + \mathbb{E}_t \sum_{j=0}^{\infty} \rho^j (r_{t+1+j} - \Delta d_{t+1+j})$$

Since future dividend growth is empirically unforecastable, we can approximate

$$d_t - p_t = \kappa + \mathbb{E}_t \sum_{j=0}^{\infty} \rho^j r_{t+1+j}$$

suggesting that the dividend-price ratio is related to (appropriately discounted) long-run expected return. In contrast, $SVIX$ bound is related to short-run expected return, as we have seen above. Hence, we can reconcile two seemingly contrasting forecast by interpreting that during 1998-1999 the marginal investor expected short-run return to be high while long-run return to be low.

Appendix

For R codes used to produce above results, see `https://github.com/ysugk/Course-FIN625/tree/master/HW/code`.