Solutions to problems in *Asset Pricing*

John H. Cochrane*
Graduate School of Business
University of Chicago
1101 E. 58th St.
Chicago IL 60637
john.cochrane@gsb.uchicago.edu

March 26, 2001

This is a very preliminary draft; it's incomplete and I'm sure full of typos. Still, I welcome comments on any problems you find with these notes.

1 Problems for Chapter 1

1. a and b are trivial. For c,

$$\frac{c_2/c_1 d(c_1/c_2)}{dR/R} = -\frac{\frac{dc_1}{c_1} - \frac{dc_2}{c_2}}{\frac{dR}{R}}.$$

The first order conditions are

$$\begin{aligned} u'(c_1) &= \lambda \\ \beta u'(c_2) &= \frac{\lambda}{R}. \end{aligned}$$

Differentiating the first order conditions,

$$\begin{aligned} \gamma \frac{dc_1}{c_1} &= \frac{c_1 u''(c_1)}{u'(c_1)} \frac{dc_1}{c_1} = \frac{d\lambda}{\lambda} \\ \gamma \frac{dc_2}{c_2} &= \frac{c_2 u''(c_2)}{u'(c_2)} \frac{dc_2}{c_2} = \frac{d\lambda}{\lambda} - \frac{dR}{R} \end{aligned}$$

2. The expected return of the asset is the same as that of its mimicking portfolio, $proj(R|m)$

3.

   (a) We know there are a, b, such that $m = a + bR^{mv}$. Determine a,b, by pricing R^{mv} and the risk free rate R^f

$$\begin{aligned} 1 &= E(mR^{mv}) = E\left[(a + bR^{mv})(R^{mv})\right] \\ 1 &= E(mR^f) = \left[E(a + bR^{mv})R^f\right] \end{aligned}$$

$$\begin{aligned}
1 &= aE(R^{mv}) + bE\left(R^{mv2}\right) \\
1 &= aR^f + bE(R^{mv})R^f
\end{aligned}$$

$$\begin{aligned}
a &= \frac{E(R^{mv})R^f - E(R^{mv2})}{E(R^{mv})^2R^f - E(R^{mv2})R^f} = \frac{E(R^{mv2}) - E(R^{mv})R^f}{R^f var(R^{mv})} \\
&= \frac{var(R^{mv}) + \left(E(R^{mv}) - R^f\right)E(R^{mv})}{R^f var(R^{mv})} = \frac{1}{R^f}\left(1 + \frac{\left(E(R^{mv}) - R^f\right)E(R^{mv})}{var(R^{mv})}\right) \\
b &= \frac{E(R^{mv}) - R^f}{E(R^{mv})^2R^f - E(R^{mv2})R^f} = -\frac{1}{R^f}\frac{E(R^{mv}) - R^f}{var(R^{mv})}
\end{aligned}$$

$$\begin{aligned}
b &= -\frac{1}{R^f}\frac{E(R^{mv}) - R^f}{var(R^{mv})} \\
a &= \frac{1}{R^f} - bE(R^{mv}).
\end{aligned}$$

An easier way to do this is to parameterize the linear function by a mean and shock:

$$|\rho| = 1 : m = E(m) + a(R^{mv} - E(R^{mv}))$$

$$E(m) = 1/R^f : \; m = 1/R^f + a(R^{mv} - E(R^{mv}))$$

$$1 = E(mR^{mv}) : \; 1 = \frac{E(R^{mv})}{R^f} + a\sigma^2(R^{mv})$$

$$a = -\frac{E(R^{mv}) - R^f}{R^f\sigma^2(R^{mv})}$$

$$m = \frac{1}{R^f} - \frac{E(R^{mv}) - R^f}{R^f\sigma^2(R^{mv})}(R^{mv} - E(R^{mv}))$$

(b) We had

$$E(R^i) = R^f + \beta_{i,m}\lambda_m$$

We have

$$cov(R^i, a + bR^{mv}) = bcov(R^i, R^{mv}).$$

4. No. The Sharpe ratio bound applies to any excess return

$$\frac{E(R^i) - E(R^j)}{\sigma(R^i - R^j)} \le \frac{\sigma(m)}{E(m)} = \frac{E(R^{mv}) - R^f}{\sigma(R^{mv})}$$

5.

$$\begin{aligned}
\sigma\left[(c_{t+1}/c_t)^{-\gamma}\right] &= \sqrt{E\left(e^{-2\gamma\Delta\ln c_{t+1}}\right) - E\left(e^{-\gamma\Delta\ln c_{t+1}}\right)^2} \\
&= \sqrt{e^{-2\gamma E(\Delta\ln c_{t+1}) + 2\gamma^2\sigma^2(\Delta\ln c_{t+1})} - e^{-2\gamma E(\Delta\ln c_{t+1}) + \gamma^2\sigma^2(\Delta\ln c_{t+1})}} \\
&= e^{-\gamma E(\Delta\ln c_{t+1}) + \frac{1}{2}\gamma^2\sigma^2(\Delta\ln c_{t+1})}\sqrt{e^{\gamma^2\sigma^2(\Delta\ln c_{t+1})} - 1}
\end{aligned}$$

$$E\left[(c_{t+1}/c_t)^{-\gamma}\right] = E\left(e^{-\gamma\ln\Delta c_{t+1}}\right) = e^{-\gamma E(\Delta\ln c_{t+1}) + \frac{1}{2}\gamma^2\sigma^2(\Delta\ln c_{t+1})}.$$

Dividing, we get the first result. For the second result, use the approximation for small x that $e^x \approx 1 + x$.

6. You wouldn't put *all* your money in such an asset, but you might well put some of your money in such an asset if it provides insurance – if its beta is low. (Graph!)

7.

(a) Rather obviously, use the equation at t and $t+1$, i.e. start with

$$p_{t+1} = E_{t+1}\left(\beta\frac{u'(c_{t+2})}{u'(c_{t+1})}d_{t+2} + \beta^2\frac{u'(c_{t+3})}{u'(c_{t+1})}d_{t+3} + ...\right)$$

(b) Substitute recursively,

$$\begin{aligned} p_t &= E_t\left[\beta\frac{u'(c_{t+1})}{u'(c_t)}p_{t+1}\right] + E_t\left[\beta\frac{u'(c_{t+1})}{u'(c_t)}d_{t+1}\right] \\ &= E_t\left[\beta^2\frac{u'(c_{t+2})}{u'(c_t)}p_{t+2}\right] + E_t\left[\beta^2\frac{u'(c_{t+2})}{u'(c_t)}d_{t+2}\right] + E_t\left[\beta\frac{u'(c_{t+1})}{u'(c_t)}d_{t+1}\right] \\ &\quad ... \\ &= E_t\sum_{j=1}^{\infty}\beta^j\frac{u'(c_{t+j})}{u'(c_t)}d_{t+j} + \lim_{T\to\infty} E_t\left[\beta^T\frac{u'(c_{t+T})}{u'(c_t)}p_{t+T}\right] \end{aligned}$$

The last term is *not* automatically zero. For example, if $u'(c)$ is a constant, then $p_t = \beta^t$ or greater growth will lead to such a term. It also has an interesting economic interpretation. Even if there are no dividends, if the last term is present, it means the price today is driven entirely by the expectation that someone else will pay a higher price tomorrow. People think they see this behavior in "speculative bubbles" and some models of money work this way.

The absence of the last term *is* a first order condition for optimization of an infinitely-lived consumer. If $p_t < (>)\ E_t\sum_{j=1}^{\infty}\beta^j\frac{u'(c_{t+j})}{u'(c_t)}d_{t+j}$, he can buy (sell) more of the asset, eat the dividends as they come, and increase utility. This lowers c_t, increases c_{t+j}, until the condition is filled.

If markets are complete – if he can also buy and sell claims to the individual dividends – then he can do even more. For example, if $p_t >$, then he can sell the asset, buy claims to each dividend, pay the dividend stream of the asset with the claims, and make a sure, instant profit. He does not have to wait forever. (Advocates of bubbles point out that you have to wait a long time to eat the dividend stream, but they often forget the opportunities for immediate arbitrage that a bubble can induce. The plausibility of bubbles relies on incomplete markets.)

Bubble type solutions show up often in models with overlapping generations, no bequest motive, and incomplete markets. The OG gets rid of the individual first order condition that removes bubbles, and the incomplete markets gets rid of the arbitrage opportunity. The possibility of bubbles figures in the evaluation of volatility tests.

8.

$$\Lambda = e^{-\delta t}u_c(c,l)$$

$$d\Lambda = -\delta\Lambda dt + e^{-\delta t}\left[u_{cc}dc + u_{cl}dl + \frac{1}{2}u_{ccc}dc^2 + \frac{1}{2}u_{cll}dl^2 + u_{ccl}dcdl\right]$$

$$\frac{d\Lambda}{\Lambda} = -\delta dt + \left[\frac{u_{cc}}{u_c}dc + \frac{u_{cl}}{u_{cc}}dl + \frac{1}{2}\frac{u_{ccc}}{u_c}dc^2 + \frac{1}{2}\frac{u_{cll}}{u_c}dl^2 + \frac{u_{ccl}}{u_c}dcdl\right]$$

After multiplication by dP/P only the dc and dl terms will have anything left, so

$$\begin{aligned} E_t\left(\frac{dp}{p}\right)+\frac{D}{p}dt-r_t^f\,dt &= E_t\left(\frac{dp}{p}\frac{d\Lambda}{\Lambda}\right) \\ &= \frac{u_{cc}}{u_c}E_t\left(\frac{dp}{p}dc\right)+\frac{u_{cl}}{u_c}E_t\left(\frac{dp}{p}dl\right) \end{aligned}$$

or,

$$E_t(R^i)-R^f \approx \frac{u_{cc}}{u_c}cov_t(R^i,c)+\frac{u_{cl}}{u_c}cov_t(R^i,l)$$

this is your first view of a *multifactor model*, one with multiple betas or factors on the right hand side. Of course, there is nothing deep about multiple factors – the same model is expressed with the single Λ on the right hand side. But there may be more economic intuition in having the c and l separately rather than combining the two into Λ.

9.

$$\begin{aligned} 1 &= E(e^{\ln m+\ln R})>e^{E(\ln m)+E(\ln R)} \\ 0 &> E(\ln m)+E(\ln R) \\ -E(\ln m) &> E(\ln R) \end{aligned}$$

If you increase leverage α in $R=(1-\alpha)R^f+\alpha R^m$ you increase mean and volatility. If R can get anywhere near zero, $\ln R$ goes off to -∞. Thus, increasing α eventually leads to a decrease in $E\ln R$. For example, if returns are normal, then

$$\begin{aligned} E\left(R\right) &= e^{E(\ln R)+\frac{1}{2}\sigma^2(R)} \\ \ln E(R) &= E(\ln R)+\frac{1}{2}\sigma^2(R) \\ E(\ln R) &= \ln E(R)-\frac{1}{2}\sigma^2(R) \\ E(\ln R) &= \ln\left[\alpha E(R^m)+(1-\alpha)R^f\right]-\frac{1}{2}\alpha^2\sigma^2(R^m). \end{aligned}$$

As α increases, the second term eventually dominates.

2 Problems for Chapter 2

1.

(a)

$$\begin{aligned} p_t &= E_t\sum\beta^j\left(\frac{c_{t+j}}{c_t}\right)^{-\gamma}c_{t+j} \\ \frac{p_t}{c_t} &= E_t\sum\beta^j\left(\frac{c_{t+j}}{c_t}\right)^{1-\gamma}. \end{aligned}$$

If $\gamma=1$,

$$\frac{p}{c}=\beta/(1-\beta)=\frac{1}{\delta}$$

where $\beta=1/(1+\delta)$.

(b) If $\gamma < 1$, then a rise in c_{t+j} raises p_t. If $\gamma > 1$, however, a rise in c_{t+j} lowers p_t. Any piece of news has two possible effects: cashflows and discount rates. In this case the discount rate rises faster than the payoffs, so the price actually declines.

2.

(a) The first order conditions are

$$c_t - c^* = E_t\left[R\beta(c_{t+1} - c^*)\right]$$

with $R = 1 + r$, and hence

$$c_t = E_t\left(c_{t+1}\right).$$

Iterate the technology forward,

$$k_{t+2} = R\left(Rk_t + i_t\right) + i_{t+1} = R^2 k_t + Ri_t + i_{t+1}$$

$$\begin{aligned} k_{t+3} &= R^3 k_t + R^2 i_t + Ri_{t+1} + i_{t+2} \\ \frac{1}{R^3} k_{t+3} &= k_t + \frac{1}{R}\left[i_t + \frac{1}{R} i_{t+1} + \frac{1}{R^2} i_{t+2}\right] \\ \beta^3 k_{t+3} &= k_t + \beta\left[i_t + \beta i_{t+1} + \beta^2 i_{t+2}\right] \end{aligned}$$

Continuing and with the transversality condition $\lim_{T\to\infty} \beta^T k_{t+T} = 0$, and $i = e - c$

$$k_t + \sum_{j=0}^{\infty} \beta^{j+1} e_{t+j} = \sum_{j=0}^{\infty} \beta^{j+1} c_{t+j}$$

Taking expectations,

$$k_t + \sum_{j=0}^{\infty} \beta^{j+1} E_t e_{t+j} = \sum_{j=0}^{\infty} \beta^{j+1} E_t c_{t+j}.$$

Intuitively, the present value of future consumption must equal wealth plus the present value of future endowment (labor income).

The $j + 1$ comes from the timing, alas standard in the macro literature and national income accounts . If you adopt the more common finance timing convention

$$k_{t+1} = (1 + r)\left(k_t + i_t\right)$$

you get more natural present value formulas with β^j.

Now, substitute the first order condition in the budget constraint (production possibility frontier if you want the General Equilibrium interpretation)

$$\begin{aligned} k_t + \sum_{j=0}^{\infty} \beta^{j+1} E_t e_{t+j} &= \sum_{j=0}^{\infty} \beta^{j+1} c_t = \\ \beta \frac{1}{(1-\beta)} c_t &= \frac{1}{R} \frac{1}{(1 - \frac{1}{R})} c_t = \frac{1}{R-1} c_t = \frac{c_t}{r} \end{aligned}$$

$$c_t = rk_t + r\sum_{j=0}^{\infty} \beta^{j+1} E_t e_{t+j}.$$

Consumption equals the annuity value of wealth (capital) rk_t plus the present value of future labor income (endowment). This is the *permanent income hypothesis*. It is not a "partial equilibrium" result – it is a general equilibrium model with linear technology and an endowment income process.

Now to the random walk in consumption. Just quasi-first difference, and use $k_{t+1}-k_t = rk_t+i_t$,

$$\begin{aligned}
c_t &= rk_t + r\left(\beta e_t + \beta^2 E_t e_{t+1} + \beta^3 E_t e_{t+2} + ...\right) \\
c_{t-1} &= rk_{t-1} + r\left(\beta e_{t-1} + \beta^2 E_{t-1} e_t + \beta^3 E_{t-1} e_{t+1} + ...\right) \\
c_t - c_{t-1} &= r(k_t - k_{t-1}) + ... \\
c_t - c_{t-1} &= r(rk_{t-1} + e_{t-1} - c_{t-1}) + ... \\
c_t - c_{t-1} &= r\left[rk_{t-1} + e_{t-1} - rk_{t-1} - r\left(\beta e_{t-1} + \beta^2 E_{t-1} e_t + \beta^3 E_{t-1} e_{t+1} + ...\right)\right] + ... \\
c_t - c_{t-1} &= re_{t-1} + r\left(\beta e_t + \beta^2 E_t e_{t+1} + \beta^3 E_t e_{t+2} + ...\right) \\
&\quad - \left(r^2 + r\right)\left(\beta e_{t-1} + \beta^2 E_{t-1} e_t + \beta^3 E_{t-1} e_{t+1} + ...\right) \\
c_t - c_{t-1} &= re_{t-1} + r\left(\beta e_t + \beta^2 E_t e_{t+1} + \beta^3 E_t e_{t+2} + ...\right) - r\left(e_{t-1} + \beta E_{t-1} e_t + \beta^2 E_{t-1} e_{t+1} + ...\right) \\
c_t &= c_{t-1} + (E_t - E_{t-1})\, r\beta \sum_{j=0}^{\infty} \beta^j e_{t+j}.
\end{aligned}$$

Consumption is a *random walk.* Changes in consumption equal the *innovation* in the present value of future income.

Bob Hall (1979) noticed the random walk nature of consumption in this model, and suggested testing it by running regressions of Δc_t on any variable at time $t-1$. This paper was a watershed. It is the first "Euler equation" test of a model; note it does not require the full model solution tying the shocks in Δc_t to fundamental taste and technology shocks – the second term in our random walk equation. The Hansen-Singleton (1982) Euler equation tests generalize to non-quadratic utility, random asset returns for which it is impossible to fully solve the model.

Technical details: I have assumed no free disposal - you follow the first order conditions even if past the bliss point. If you can freely dispose of consumption, then you will always end up at the bliss point c^* sooner or later. (Thanks to Ashley Wang for pointing this out. Hansen and Sargent's treatments of this problem deal with the bliss point issue.)

By the way, the algebra is much easier if you use lag operators, i.e. write $c_t = rk_t + r\beta E_t\left[(1-\beta L^{-1})^{-1} e_t\right]$. But if you know how to do that, you've probably seen this model before.

(b)

$$c_t = rk_t + r\sum_{j=0}^{\infty} \beta^{j+1} E_t e_{t+j} = rk_t + r\beta \sum_{j=0}^{\infty} \beta^j \rho^j e_t = rk_t + \frac{r\beta}{1-\beta\rho} e_t.$$

$$c_t = c_{t-1} + (E_t - E_{t-1})\, r\beta \sum_{j=0}^{\infty} \beta^j e_{t+j} = c_{t-1} + r\beta \sum_{j=0}^{\infty} \beta^j \rho^j \varepsilon_t = c_{t-1} + \frac{r\beta}{1-\beta\rho}\varepsilon_t.$$

The top equation does look like a consumption function, but notice that the parameter relating consumption c to income e depends on the persistence of income e. It is not a "psychological law" or a constant of nature. If the government changes policy so that income is more unpredictable (i.e. it gets rid of the predictable part of recessions), then this coefficient declines dramatically. The income coefficient is not "policy-invariant." This is the basis of Bob Lucas (1974) dramatic deconstruction of Keynesian models based on consumption functions that were used for policy experiments.

In both equations, you see that consumption responds to "permanent income" and that as shocks get more "permanent" – as ρ rises – consumption moves more.

(c) R was the rate of return on technology. Despite the symbol, it is not (yet) the interest rate – the equilibrium rate of return on one-period claims to consumption. That remains to be proved. The logic is, first find c, then price things from the equilibrium consumption stream. To be precise and pedantic, call the risk free rate R^f, and

$$\frac{1}{R_t^f} = E_t\left(\beta\frac{u'(c_{t+1})}{u'(c_t)}\right) = \beta E_t\left(\frac{c_{t+1}-c^*}{c_t-c^*}\right) = \beta\left(\frac{c_t-c^*}{c_t-c^*}\right) = \beta = \frac{1}{R}$$

Now, the fun stuff. We can approach the price of the consumption stream by brute force,

$$\begin{aligned} p_t &= E_t\sum_{j=1}^{\infty} m_{t,t+j}c_{t+j} = E_t\sum_{j=1}^{\infty}\beta^j\frac{c^*-c_{t+j}}{c^*-c_t}c_{t+j} = E_t\sum_{j=1}^{\infty}\beta^j\frac{c^*c_{t+j}-c_{t+j}^2}{c^*-c_t} \\ &= \sum_{j=1}^{\infty}\beta^j\frac{c^*c_t - E_t\left(c_{t+j}^2\right)}{c^*-c_t} = \sum_{j=1}^{\infty}\beta^j\frac{c^*c_t - c_t^2 - var_t\left(c_{t+j}\right)}{c^*-c_t} \end{aligned}$$

$$\begin{aligned} c_{t+1} &= c_t + \frac{r\beta}{1-\beta\rho}\varepsilon_{t+1} \\ c_{t+2} &= c_t + \frac{r\beta}{1-\beta\rho}\left(\varepsilon_{t+1}+\varepsilon_{t+2}\right) \\ c_{t+j} &= c_t + \frac{r\beta}{1-\beta\rho}\left(\varepsilon_{t+1}+..+\varepsilon_{t+j}\right) \end{aligned}$$

$$\begin{aligned} E_t(c_{t+j}) &= c_t \text{ (of course)} \\ var_t(c_{t+j}) &= j\left(\frac{r\beta}{1-\beta\rho}\right)^2\sigma_\varepsilon^2 \end{aligned}$$

$$\begin{aligned} p_t &= \sum_{j=1}^{\infty}\beta^j\frac{c_t\left(c^*-c_t\right) - j\left(\frac{r\beta}{1-\beta\rho}\right)^2\sigma_\varepsilon^2}{c^*-c_t} \\ &= \sum_{j=1}^{\infty}\beta^j\left[c_t - \frac{j\left(\frac{r\beta}{1-\beta\rho}\right)^2\sigma_\varepsilon^2}{c^*-c_t}\right] \\ &= \left(\sum_{j=1}^{\infty}\beta^j\right)c_t - \left(\sum_{j=1}^{\infty}j\beta^j\right)\frac{\left(\frac{r\beta}{1-\beta\rho}\right)^2\sigma_\varepsilon^2}{c^*-c_t} \end{aligned}$$

$$\sum_{j=1}^{\infty}j\beta^j = \frac{\beta}{(\beta-1)^2}$$

$$
\begin{aligned}
p_t &= \frac{\beta}{1-\beta}c_t - \frac{\beta}{(1-\beta)^2}\frac{\left(\frac{r\beta}{1-\beta\rho}\right)^2\sigma_\varepsilon^2}{c^*-c_t} \\
&= \frac{\frac{1}{1+r}}{1-\frac{1}{1+r}}c_t - \frac{\frac{1}{1+r}}{\left(1-\frac{1}{1+r}\right)^2}\frac{\left(\frac{r\beta}{1-\beta\rho}\right)^2\sigma_\varepsilon^2}{c^*-c_t} \\
p_t &= \frac{1}{r}c_t - \frac{\beta}{(1-\beta\rho)^2}\frac{1}{c^*-c_t}\sigma_\varepsilon^2
\end{aligned}
$$

Wow. The first term is the risk-neutral price – the value of a perpetuity paying c. (Don't forget $E_t(c_{t+j}) = c_t$) The second term is a risk correction. It lowers the price. If σ_ε^2 is high – more risk –the price is lower. If ρ is high – more persistent consumption – the price is lower.

Now, the hard term – the effect of consumption. At the bliss point, the consumer is as happy as can be, and marginal utility falls to zero. Hence, the consumer is infinitely risk averse. ($u''(c)/u'(c)$ rises to infinity). There is no consumption you can give him to compensate for risk, since he's at the bliss point. No surprise that the price goes off to $-\infty$ here. As consumption rises towards the bliss point, the consumer gets more and more risk averse (u'' is constant, u' is falling), so the price declines. Above the bliss point, the consumer values consumption negatively, so the price is higher than the risk-neutral version.

This feature – that risk aversion rises as consumption rises – is obviously not a good one. Quadratic utility is best used as a local approximation. Find a c^* that gives a sensible risk aversion, and then make sure the model doesn't get too far away!

The question says price as a function of e and k. I'm curious how I ever got that, since it seems a much more natural function of c. c is a function of e and k, of course, but substituting that in does not seem very easy.

3. This is not only a historically important model, it introduces a very important method. Evaluating infinite sums as in the last problem is a huge pain. In most models, conditioning information is a function of only a few *state variables*, x_t. Everything you could want to know about the current state of the economy, and the conditional distribution of everything you could want to know in the future is contained in the state variables. Hence, prices (at least properly scaled) have to be a function of the state variables. Instead of solving for p in terms of a huge infinite sum, you can solve the *functional equation* $p(x) = E_t\left[m_{t,t+1}(x_t, x_{t+1})\left(p(x_{t+1}) + d_{t+1}\right)\right]$. Here we go...

   (a) From the basic first order condition,

$$
p_t^b = E_t\beta u'(c_{t+1})/u'(c_t) = E_t\beta\Delta c_{t+1}^{-\gamma}
$$

$$
\begin{aligned}
p^b(\Delta c_t &= h) = \beta\pi_{h\to h}h^{-\gamma} + \beta\pi_{h\to l}l^{-\gamma} \\
p^b(\Delta c_t &= l) = \beta\pi_{l\to h}h^{-\gamma} + \beta\pi_{l\to l}l^{-\gamma}
\end{aligned}
$$

$$
\begin{bmatrix} p^b(\Delta c_t = h) \\ p^b(\Delta c_t = l) \end{bmatrix} = \begin{bmatrix} \pi_{h\to h} & \pi_{h\to l} \\ \pi_{l\to h} & \pi_{l\to l} \end{bmatrix}\begin{bmatrix} \beta h^{-\gamma} \\ \beta l^{-\gamma} \end{bmatrix}
$$

$$
p^b = \pi x.
$$

   The riskfree rate is of course

$$
R^f = 1/p^b.
$$

(b) The consumption stream:

$$p_t = E_t\left[\beta\Delta c_{t+1}^{-\gamma}\left(p_{t+1}+c_{t+1}\right)\right]$$

$$\frac{p_t}{c_t} = \beta E_t\left[\Delta c_{t+1}^{1-\gamma}\left(\frac{p_{t+1}}{c_{t+1}}+1\right)\right]$$

Solve this as a functional equation, as explained above. Find p/c in the h state and in the l state (functions from two points to the real line are easy to determine– you just find the values at the two points.)

$$\frac{p}{c}(h) = \beta\pi_{h\to h}h^{1-\gamma}\left(\frac{p}{c}(h)+1\right)+\beta\pi_{h\to l}l^{1-\gamma}\left(\frac{p}{c}(l)+1\right)$$

$$\left[\begin{array}{c} p/c(h) \\ p/c(l) \end{array}\right] = \beta\left[\begin{array}{cc} \pi_{h\to h}h^{1-\gamma} & \pi_{h\to l}l^{1-\gamma} \\ \pi_{l\to h}h^{1-\gamma} & \pi_{l\to l}l^{1-\gamma} \end{array}\right]\left(\left[\begin{array}{c} 1 \\ 1 \end{array}\right]+\left[\begin{array}{c} p/c(h) \\ p/c(l) \end{array}\right]\right)$$

$$pc = \beta\pi^*(\mathbf{1}+pc)$$

$$pc = \left(1-\beta\pi^*\right)^{-1}\beta\pi^*\mathbf{1}$$

We can find returns from

$$R_{t+1} = \frac{\frac{p_{t+1}}{c_{t+1}}+1}{\frac{p_t}{c_t}}\frac{c_{t+1}}{c_t}.$$

Note when p/c is constant, R is just a constant times consumption growth. You need a very small p/c before R is much different from consumption growth.

Conditionally expected returns follow from the probabilities.

(c) Start with the calibration. It's most natural to take the two points to be equally above and below the mean, $h = 1.01+x, l = 1.01-x$ and equal probabilities. Then, you want

$$1/2(1.01+x)+1/2(1.01-x) = 1.01$$

$$1/2x^2+1/2x^2 = 0.01^2$$

i.e., $x = 0.01$.

Here are my results.

	To state	In state h	In state l
$\gamma = 0.5$			
bond price		0.985	0.985
R^f		1.5%	1.5%
p/c		196	196
R	h	2.52%	2.52%
	l	0.51%	0.51%
$\gamma = 5$			
bond price		0.943	0.943
R^f		6.01	6.01
p/c		19.96	19.96
R	h	7.11	7.11
	l	5.01	5.01

The major failing is the *equity premium.* The mean stock return is almost exactly the same as the riskfree rate. Also, stock returns are perfectly correlated with consumption growth. The standard deviation of stock returns is about 1%, not about 20%. The Sharpe ratio $\left[E(R)-R^f\right]/\sigma(R)$ is way too low.

(d) To get serial correlation in consumption growth, I tried π of the form

$$\pi=\begin{bmatrix} 1/2+\theta & 1/2-\theta \\ 1/2-\theta & 1/2+\theta \end{bmatrix}$$

Now,

$$E(dc_{t+1}|dc_t=h)=(1/2+\theta)*(g+x)+(1/2-\theta)*(g-x)=g+2\theta x$$
$$E(dc_{t+1}|dc_t=l)=(1/2-\theta)*(g+x)+(1/2+\theta)*(g-x)=g-2\theta x$$

Here are my results for a positive serial correlation.

		h	l
$\gamma=5,\theta=0.1$			
p^b		0.934	0.953
R^f		7.07	4.97
p/c		19.93	20.3
R	h	7.12	5.05
	l	6.99	4.92
$\rho(\Delta c_t,\Delta c_{t-1})$	0.21		

The main reason I put this in at this stage is to get variation in prices with the initial state. In the previous case, the world looks the same from any starting date, so there is no variation in prices (ex-ante). The interest rate and stock return are higher from the high state, because expected *future* consumption growth is higher. Higher return means lower price or p/c.

3 Problems for Chapter 4

1. The absence of arbitrage implies the LOOP, but not vice versa.

   NA→LOOP. Suppose the absence of arbitrage holds, but not the LOOP. Let $z=ax+by$. If $p(z)>ap(x)+bp(y)$, however, the portfolio $z-(ax+by)$ is an arbitrage. In discount factor language, if there is an $m>0$, then there is an m. The LOOP theorem specifies an m in $\underline{X}$, however. Given an m, we can construct an m in $\underline{X}$ by $x^*=proj(m|\underline{X})$.

   LOOP→NA. In discount factor language, imagine a complete market with a discount factor that is negative in some state of nature. This generates a set of prices that obey the law of one price, but leave arbitrage opportunities. The corresponding set of prices and payoffs are a counterexample in portfolio language.

2. The danger of applying the LOOP or no arbitrage in a sample is that you typically don't see all of the possible realizations in any finite sample. For example, a corporate 10 year bond and a government 10 year bond will have identical payoffs in any sample in which the corporation does not default, but the corporate bond will have a lower price. This looks like a violation of the law of one price. Hansen-Jagannathan bounds with positivity typically show "arbitrage bound" limits on the risk free rate, which come from samples in which one security dominates another. These arbitrage bounds disappear if one posits a distribution in which it is always possible for each security to underperform the other.

(a) R^{-1} is a discount factor. It is not necessarily in the payoff space, since that space is constructed of linear combinations of the assets. x^* is the unique discount factor in the payoff space, but not the only discount factor. Often, $R > 0$, i.e. for limited liability securities like stocks. In this case, R^{-1} is always positive, but so is $R/E(R^2)$. Securities do not have to be limited liability, so in general R^{-1} can be negative. The biggest trouble with this discount factor is that it can be infinite if $R = 0$ can happen, in which case the expectation may not be defined. (It may be out of the set of random variables with second moments).

(b) The first order conditions are

$$E\left(\frac{1}{\alpha' R} R\right) = \lambda.$$

Thus,

$$m = \frac{1}{\lambda \alpha' R}$$

is a discount factor. In general, you can't solve the first order conditions for α analytically.

Another more beautiful way to do this. We know that every payoff in X can be priced by a discount factor $m > 0$ State the problem as

$$\begin{aligned}\max E(\ln(R)) \text{ s.t. } 1 &= E(mR) \\ \max \sum \pi_i \ln(R_i) \text{ s.t. } 1 &= \sum \pi_i m_i R_i\end{aligned}$$

the first order conditions – choose R_i in each state i – are

$$\pi_i \frac{1}{R_i} = \lambda \pi_i m_i$$

$$\frac{1}{R_i} = \lambda m_i$$

Plugging this in to the constraint, you find $\lambda = 1$. Thus, we have proved: the inverse return of the portfolio that maximizes ln returns is equal to a discount factor.

(c) The latter approach is a quick way to do this in continuous time. A discount factor is a proces Λ that prices payoffs at any date. Thus, consider the "growth optimal trading strategy" – the value process V that maximizes

$$\max E\left[\ln\left(\frac{V_T}{V_0}\right)\right] \text{ s.t. } V_0\Lambda_0 = E_t\left[V_T\Lambda_T\right].$$

Just as before, we have

$$\frac{V_T}{V_0} = \frac{\Lambda_0}{\Lambda_T}.$$

This holds at any horizon, so V_t is a numeraire – a price process such that for any security priced by Λ,

$$\frac{p_t}{V_t} = E_t \int_{s=0}^{\infty} \frac{D_{t+s}}{V_{t+s}} ds.$$

4 Problems for Chapter 5

1. You have to find equations that express the right angles in the picture. Right angles means orthogonal with second moment norm, so we want to prove that the line from any R^e to its projection on R^{e*} lies at right angles to R^{e*}, $E\left[(R^e - proj(R^e|R^{e*})) \times R^{e*}\right] = 0$. Working on the latter expression,

$$\begin{aligned}
E\left[(R^e - proj(R^e|R^{e*})) \times R^{e*}\right] &= E\left[R^e R^{e*} - proj(R^e|R^{e*})R^{e*}\right] \\
&= E(R^e R^{e*}) - E(proj(R^e|R^{e*})) \\
&= E(R^e) - E\left(\frac{E(R^e R^{e*})}{E(R^{e*2})} R^{e*}\right) \\
&= E(R^e) - E\left(\frac{E(R^e)}{E(R^{e*})} R^{e*}\right) = 0.
\end{aligned}$$

   I used the properties $E(R^{e*}R^e) = 0$ and $E(R^{e*2}) = E(R^{e*})$.

2. Start with a nonstochastic economy. In this case, x^* is typically below the set of returns. x^* is a discount factor, so typically less than one. Returns are returns, hence typically greater than one. Precisely, in a nonstochastic economy,

$$x^* = 1/R^f$$

   If $R^f > 1$, then $x^* < 1 < R^f$. It's possible that $R^f < 1$, if consumption is declining drastically, but not typical.

   Now, let's do it in a stochastic economy. R^* is the return parallel to x^*,

$$R^* = \frac{x^*}{E(x^{*2})}; \;\; x^* = \frac{R^*}{E(R^{*2})}.$$

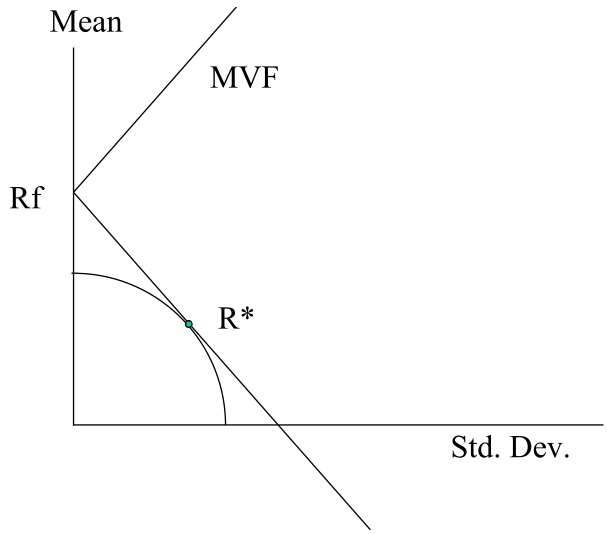
   so we just have to figure out if x^* is longer or shorter than R^*. Now, from the definitions,

$$|x^*|^2 = E(x^{*2}) = \frac{1}{E(R^{*2})} = \frac{1}{|R^*|^2}$$

$$|x^*|^2 \, |R^*|^2 = 1$$

   Thus, $|x^*| < |R^*|$ if $|R^*| > 1$ or if $|x^*| < 1$. This is very nice: In a nonstochastic economy $x^* R^f = 1$; in a stochastic economy this generalizes to $|x^*| \, |R^*| = 1$.

   So is the second moment of the return with smallest second moment greater or less than one? As you can see in the drawing below, this can happen if risk premia (slope of the mean-variance frontier) is high, and if the riskfree rate is low, not much more than 1.0.



There are lots of ways to continue from here, to see if typical numbers give one of these conditions. Easiest, by just looking at the frontier, I am able to show that

$$E(R^{*2}) = \frac{R^{f2}}{\left|\frac{E(R^e)}{\sigma(R^e)}\right|^2 + 1} = \frac{1.01}{1+0.25} < 1$$

(1% is about the average real interest rate and stocks have averaged roughly 9% mean and 16% standard deviation.) This means, we should actually expect x^* to lie somewhat above the return line.

Derivation:

$$E(R^2) = E(R)^2 + \sigma^2(R) = \left[R^f + E(R^e)\right]^2 + \sigma^2(R^e)$$

where $R^e = R - R^f$;

$$\begin{aligned}
&= R^{f2} + 2R^f E(R^e) + E(R^e)^2 + \sigma^2(R^e) \\
&= R^{f2} + 2R^f E(R^e) + E(R^e)^2 + E(R^e)^2 \left|\frac{\sigma(R^e)}{E(R^e)}\right|^2 \\
&= R^{f2} + 2R^f E(R^e) + E(R^e)^2 \left(1 + \left|\frac{\sigma(R^e)}{E(R^e)}\right|^2\right)
\end{aligned}$$

The minimum second moment return occurs (minimize over $E(R^e)$, holding Sharpe ratio constant) occurs at

$$E(R^e) = -\frac{R^f}{\left(1 + \left|\frac{\sigma(R^e)}{E(R^e)}\right|^2\right)}$$

and has value

$$E(R^{*2}) = R^{f2} - \frac{2R^{f2}}{\left(1 + \left|\frac{\sigma(R^e)}{E(R^e)}\right|^2\right)} + \frac{R^{f2}}{\left(1 + \left|\frac{\sigma(R^e)}{E(R^e)}\right|^2\right)^2}\left(1 + \left|\frac{\sigma(R^e)}{E(R^e)}\right|^2\right)$$

$$
\begin{aligned}
E(R^{*2}) &= R^{f2}\left(1-\frac{1}{\left(1+\left|\frac{\sigma(R^e)}{E(R^e)}\right|\right)^2}\right)\\
&= R^{f2}\left(\frac{\left|\frac{\sigma(R^e)}{E(R^e)}\right|^2}{1+\left|\frac{\sigma(R^e)}{E(R^e)}\right|^2}\right)=\frac{R^{f2}}{\left|\frac{E(R^e)}{\sigma(R^e)}\right|^2+1}
\end{aligned}
$$

A little more formally, or using some of the tools and representations of the course, form x^*

$$x^* = p'E(xx')^{-1}x$$

$$E(x^{*2}) = p'E(xx')^{-1}p$$

Take the risk free rate and the market excess return as the elements of x,

$$x = \begin{bmatrix} R^f \\ R^e \end{bmatrix};\ p = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$E(x^{*2}) = \begin{bmatrix} 1 & 0 \end{bmatrix}\begin{bmatrix} R^{f2} & R^f E(R^e) \\ & E(R^{e2}) \end{bmatrix}^{-1}\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$E(x^{*2}) = \frac{1}{R^{f2}E(R^{e2}) - R^{f2}E(R^e)^2}\begin{bmatrix} 1 & 0 \end{bmatrix}\begin{bmatrix} E(R^{e2}) & -R^f E(R^e) \\ -R^f E(R^e) & R^{f2} \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$E(x^{*2}) = \frac{1}{R^{f2}}\frac{E(R^{e2})}{\sigma^2(R^e)} = \frac{1}{R^{f2}}\frac{E(R^e)^2+\sigma^2(R^e)}{\sigma^2(R^e)} = \frac{1}{R^{f2}}\left[1+\frac{E(R^e)^2}{\sigma^2(R^e)}\right]$$

$$E(x^{*2}) = 1/E(R^{*2}) = \frac{1}{R^{f2}}\left[1+\frac{E(R^e)^2}{\sigma^2(R^e)}\right]$$

$$E(R^{*2}) = R^{f2}\frac{1}{1+\frac{E(R^e)^2}{\sigma^2(R^e)}}$$

To get some more intuition, price the risk free rate with x^*,

$$\frac{1}{R^f} = E(x^*1) = E\left[proj(x^*|1)1\right]$$

Again, in a nonstochastic economy this reduces to $x^* = 1/R^f$ and we learn that the *projection* of x^* on 1 should lie below the line of returns. But when there are large risk premia – high Sharpe ratios; large distortions between actual and risk neutral probabilities; large distortions between contingent claims prices and probabilities – the stochastic discount factor is very different from a constant, so x^* itself must lie above R as shown in the graph below. If you believe the size of the equity premium, we live in an economy with severe distortions from risk neutrality, enough to get x^* above the plane of returns.

3. We already had from (5.20)

$$R^f = R^* + R^f R^{e*}.$$

Rearranging,

$$R^{e*} = (R^f - R^*)/R^f.$$

This is a useful formula to show that you can construct R^{e*} from knowledge of R^* and R^f.

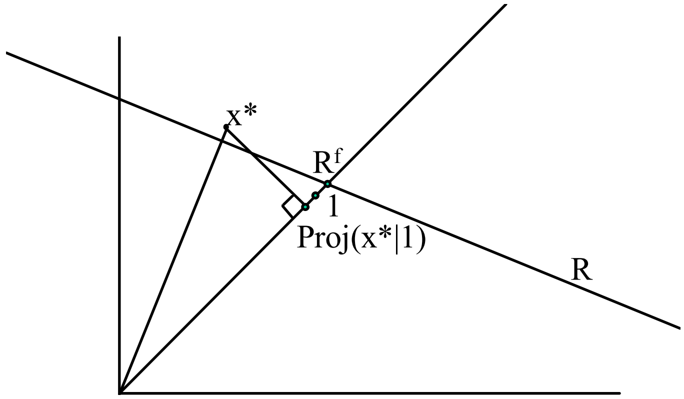


Figure 1:

4.

   (a) R^* and 1 are collinear, so R^{e*} collapses to zero. The frontier collapses to a point, $R^* = R^f$. In mean-variance space, all returns have the same mean, so the frontier collapses to a line, with $R^* = R^f$ at the leftmost point.

   (b) The projection of 1 on $\underline{R^e}$ is still zero, so R^* is still the mean variance frontier. When the payoffs are generated from x, $R^{e*} = E(x)'E(xx')^{-1}x$. But if consumers are risk neutral, $p(x) = E(mx) = E(x)$ for all assets, so $E(x) = 0$ for excess returns and $R^{e*} = 0$.

5. *No.* This is subtle. $R^* = x^*/p(x^*) = proj(m|\underline{X})/p\,[proj(m|\underline{X})]$. $\underline{R}$ is a set, but not a space, since it does not include zero, so you can't project on it. There is a rather general point in here. For example, you don't form factor-mimicking portfolios by projecting on $\underline{R}$, you project on $\underline{X}$ instead.

5 Problems for chapter 6

1. No. When the economy is risk-neutral, $R^{e*} = 0$. Thus, the frontier collapses to R^* alone. Of course, $m = x^* = 0 + \frac{1}{E(R^{*2})}R^*$ is a discount factor, and is linear in the mean-variance efficient return.

2. No. The β are different, so the λ must also be different. As a simple example, if the factor is 2 times a return, then the factor mimicking portfolio is the return, so β is cut in half and λ doubles.

3. m is in $\underline{X}$, so it is x^*. $R^* = x^*/p(x^*) = (a - bR^m)\,/\,\left(a/R^f - b\right)$.

4. They can stay the same. If you span the frontier by R^* and R^f, as you increase or decrease weight on R^*, you change the amount of one particular linear combination of factor mimicking portfolios, but not the relative weights of the factors.

5. As in the text, find y in $R^* + yR^{e*}$ to generate zero covariance.

$$\begin{aligned}
E\left[(R^* + wR^{e*})(R^* + yR^{e*})\right] &= E(R^* + wR^{e*})E(R^* + yR^{e*}) \\
E(R^{*2}) + wyE(R^{e*}) &= E(R^*)^2 + wyE(R^{e*})^2 + (w+y)E(R^*)E(R^{e*}) \\
var(R^*) + wy\,var(R^{e*}) &= (w+y)E(R^*)E(R^{e*})
\end{aligned}$$

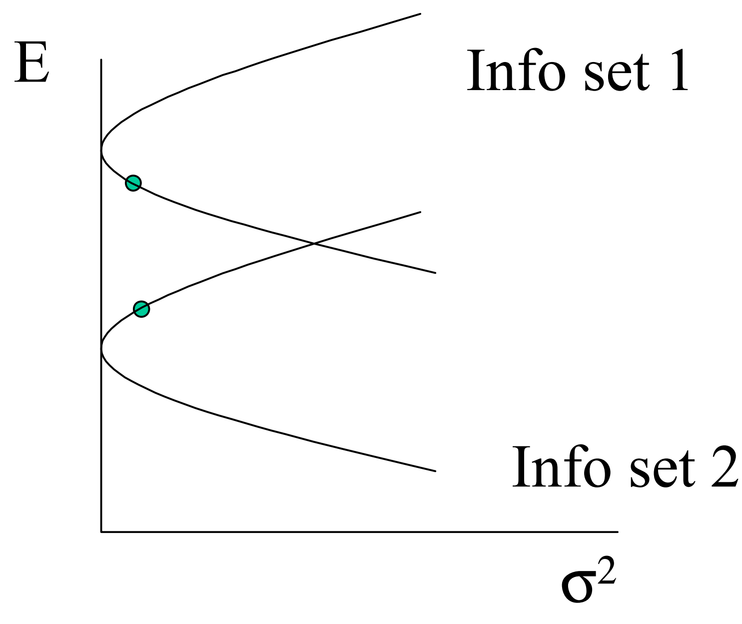
$$y = \frac{var(R^*) - wE(R^*)E(R^{e*})}{E(R^*)E(R^{e*}) - wvar(R^{e*})}$$

6. In a risk-neutral economy $R^{e*} = 0$, so the whole mean-variance frontier, including the various risk-free rate proxies, collapses to the minimum-variance point R^*. R^* is not a constant – no risk free rate is traded. In a risk neutral economy, R^* is the return closest to the unit vector.

6 Problems for Chapter 8

1. A risk free asset is on the conditional frontier. It is only on the unconditional frontier if it is constant. $R_t^f = R_{t+1}^* + R_t^f R_{t+1}^{e*}$. The conditional mean variance frontier is $R^* + w_t R^{e*}$ so the risk free rate is on the conditional mvf. The unconditional mvf is $R^* + wR^{e*}$. If the risk free rate is not constant, it is not of this form.

   Intuition: $var(R) = var(E_t(R)) + E(var_t(R))$. (Check it – it's a handy identity!) The risk free rate has none of the second term, but if not constant some of the first term. Other assets can get less $var(E_t(R))$ by accepting a little more $E(var_t(R))$ – for example the asset shown in the figure below. The frontier is a line in mean-standard deviation space, but a parabola in mean-variance space. Thus, near R^f you get so much change in conditional mean per unit of added conditional variance, that moving to an asset such as the one shown must get you towards a minimum unconditional variance portfolio, on the unconditional frontier. Analytically, the conditional mean-variance frontier is a line $E_t = R_r^f + a\sigma_t$. Thus, $\sigma_t^2 \left(\frac{E_t - R_r^f}{a}\right)^2 = \sigma_t^2 \; d\sigma_t^2/dE_t = 2\frac{E_t - R_r^f}{a}$. Thus, at $E_t = R^f$, $d\sigma_t^2/dE_t = 0$.



2. No. This is really the same question. Put another way, the unconditional mean-variance frontier will not intersect the vertical axis. This happens all the time. The 3 month T bill rate is (nominally) conditionally risk free. Yet a plot of the unconditional mean-variance frontier will not intersect the vertical axis since the T bill rate varies over time. If you're running $E(R) = \gamma + \beta\lambda$, the question is, is there is risk free rate to identify γ, or must you use a zero-beta rate. Even if there is a conditionally risk free rate, if it is not constant, the unconditional representation will need a zero-beta rate.

3. Neither, since you didn't include the managed portfolios. Unless, of course, everything is i.i.d. so there are no instruments.

7 Problems for Chapter 9

1.

$$\begin{aligned} \max Eu(c) \text{ st. } c &= (\alpha' R),\ \alpha' 1 = W \\ \frac{\partial}{\partial \alpha^i} &: E\left[u'(c)R^i\right] = \lambda \\ m_{t+1} &= \frac{u'(c_{t+1})}{\lambda_t} = \frac{c^* - c_{t+1}}{\lambda_t} = \frac{c^* - R^W_{t+1}W_t}{\lambda_t} \end{aligned}$$

2. This is a case where finite-dimensional intuition can be misleading. If there were a finite-dimensional state space, then

$$\min_{\{m\}} \text{ (or max) } p(x) = E(mx) \text{ s.t. } m \geq 0,\ p(f) = E(mf)$$

would have finite upper and lower bounds. In an infinite-dimensional state space – the typical case for factor pricing as applied to equities – it does not. Intuitively, there are no *arbitrage portfolios* of stocks – portfolios that dominate in *every state of nature*. Thus, the absence of negative prices for such portfolios doesn't help us at all. (Problems like this generate *arbitrage bounds* in option pricing problems. In option pricing, though, there are strictly dominating portfolios; a call option is better in every state of nature than the portfolio that holds the stock and borrows the strike payment.)

3. No, it has to be the risk free rate, or a zero-beta rate. Here's the potential confusion. With a risk-free rate, the CAPM is

$$E(R^i) = R^f + \beta_{i,m}\left[E(R^m) - R^f\right]$$

You can difference the left hand variables,

$$E(R^i - R^j) = \beta_{i-j,m}\left[E(R^m) - R^f\right]$$

Covariance is a linear operator, so $\beta_{i,m} - \beta_{j,m} = \beta_{i-j,m}$. The question is whether we can do this on the right hand side too. Can we write the CAPM in terms of an excess return on the right hand side? Can we write

$$E(R^{ei}) = \beta_{i,R^m - R^j} E(R^m - R^j)?$$

Once posed, you can see that the answer is no. Betas are not linear in the denominator. In discount factor language,

$$m = a - bR^m$$

can't be written

$$m = a - b(R^m - R^j)!$$

In fact, the CAPM is frequently tested with the T-bill rate as a "proxy" for the risk free rate. Though the riskiness of the T bill rate does not matter for the left hand side, it does for the right. In practice, this proxy is probably not a big deal

8 Problems for Chapter 11

1. Write the autocorrelation coefficient as

$$\rho_j = \frac{E(x_t - E(x_t))(x_{t-j} - E(x_t))}{E\left[(x_t - E(x_t))^2\right]} = \frac{E(x_t x_{t-j}) - [E(x_t)]^2}{E(x_t^2) - [E(x_t)]^2}$$

Line up the required moments as

$$E\left[\begin{array}{ccc} x_t x_{t-j} & x_t & x_t^2 \end{array}\right]'$$

The derivatives we need are

$$\begin{aligned}
\frac{\partial \rho_j}{\partial E\left[x_t x_{t-j}\right]} &= \frac{1}{E(x_t^2) - [E(x_t)]^2} = \frac{1}{\sigma^2(x)} \\
\frac{\partial \rho_j}{\partial E(x_t)} &= \frac{-2\left(E(x_t^2) - [E(x_t)]^2\right)E(x_t) + 2\left(E(x_t x_{t-j}) - [E(x_t)]^2\right)E(x_t)}{\left\{E(x_t^2) - [E(x_t)]^2\right\}^2} \\
&= 2E(x_t)\frac{E(x_t x_{t-j}) - E(x_t^2)}{\left\{E(x_t^2) - [E(x_t)]^2\right\}^2} = 2E(x)\frac{cov(x_t x_{t-j})}{\sigma^4(x)} \\
&= 2E(x)\frac{\rho_j}{\sigma^2(x)} \\
\frac{\partial \rho_j}{\partial E(x_t^2)} &= \frac{-\left(E(x_t x_{t-j}) - [E(x_t)]^2\right)}{\left\{E(x_t^2) - [E(x_t)]^2\right\}^2} = -\frac{cov(x_t x_{t-j})}{\sigma^4(x)} = -\frac{\rho_j}{\sigma^2(x)}
\end{aligned}$$

so

$$\frac{\partial \phi}{\partial b} = \frac{1}{\sigma^2(x)}\left[\begin{array}{c} 1 \\ 2E(x)\rho_j \\ -\rho_j \end{array}\right]$$

The S matrix is

$$S_j = \sum_{k=-\infty}^{\infty} E\left[\begin{array}{ccc} x_t x_{t-j} x_{t-k} x_{t-j-k} & x_t x_{t-j} x_{t-k} & x_t x_{t-j} x_{t-k}^2 \\ x_t x_{t-k} x_{t-k-j} & x_t x_{t-k} & x_t x_{t-k}^2 \\ x_t^2 x_{t-k} x_{t-k-j} & x_t^2 x_{t-k} & x_t^2 x_{t-k}^2 \end{array}\right]$$

The standard error is

$$var(corr_T) = \frac{1}{T}\left[\frac{\partial \phi}{\partial \mu}\right]' S \left[\frac{\partial \phi}{\partial \mu}\right]$$

This is straightforward to calculate, but I can't simplify it further.

If the series is i.i.d., then only the $k = 0$ terms survive,

$$S_j = E\left[\begin{array}{ccc} x_t^2 x_{t-j}^2 & x_t^2 x_{t-j} & x_t^3 x_{t-j} \\ x_t^2 x_{t-j} & x_t^2 & x_t^3 \\ x_t^3 x_{t-j} & x_t^3 & x_t^4 \end{array}\right] = \left[\begin{array}{ccc} E(x_t^2)^2 & 0 & 0 \\ 0 & E\left(x_t^2\right) & E\left(x_t^3\right) \\ 0 & E\left(x_t^3\right) & E\left(x_t^4\right) \end{array}\right],$$

and $\rho_j = 0$ in $\partial \phi / \partial \mu$. Thus,

$$\begin{aligned}
\sigma^2(\hat{\rho}_j) &= \frac{1}{T}\frac{1}{\sigma^4(x)}\left[\begin{array}{c} 1 \\ 0 \\ 0 \end{array}\right]'\left[\begin{array}{ccc} E(x^2)^2 & 0 & 0 \\ 0 & E\left(x^2\right) & E\left(x^3\right) \\ 0 & E\left(x^3\right) & E\left(x^4\right) \end{array}\right]\left[\begin{array}{c} 1 \\ 0 \\ 0 \end{array}\right] \\
&= \frac{1}{T}\frac{1}{\sigma^4(x)}\left[E\left(x^2\right)^2\right] = \frac{1}{T}\frac{\left[\sigma^2(x) + E(x)^2\right]^2}{\sigma^4(x)}
\end{aligned}$$

If, in addition, $E(x)=0$, we obtain the classic result: *If the series is i.i.d. mean zero, the standard error of the autocorrelation coefficient is* $1/\sqrt{T}$. Of course, you can calculate standard errors without mean zero, and without i.i.d.!

You can do the correlation coefficient from p.207 the same way:

$$corr(x_t, y_t) = \phi(\mu) = \frac{E(x_t, y_t) - E(x_t)E(y_t)}{\sqrt{E(x_t^2) - E(x_t)^2}\sqrt{E(y_t^2) - E(y_t)^2}}$$

$$\mu = E(u_t) = \left[E(x_t)\ E(x_t^2)\ E(y_t)\ E(y_t^2)\ E(x_t y_t)\right]'$$

$$\begin{aligned} var(corr_T) &= \frac{1}{T}\left[\frac{\partial\phi}{\partial\mu}\right]' S \left[\frac{\partial\phi}{\partial\mu}\right] \\ S &= \sum_{j=-\infty}^{\infty} cov(u_t, u'_{t-j}) \end{aligned}$$

We need the derivatives,

$$\frac{\partial\phi}{\delta\mu} = \begin{bmatrix} \frac{\partial\phi}{\partial E(x_t)} \\ \frac{\partial\phi}{\partial E(x_t^2)} \\ \frac{\partial\phi}{\partial E(y_t)} \\ \frac{\partial\phi}{\partial E(y_t^2)} \\ \frac{\partial\phi}{\partial E(x_t y_t)} \end{bmatrix} = \begin{bmatrix} \frac{-E(y_t)}{\sqrt{E(x_t^2)-E(x_t)^2}\sqrt{E(y_t^2)-E(y_t)^2}} + \frac{E(x_t,y_t)-E(x_t)E(y_t)}{\left(E(x_t^2)-E(x_t)^2\right)^{\frac{3}{2}}\sqrt{E(y_t^2)-E(y_t)^2}}E(x_t) \\ -\frac{1}{2}\frac{E(x_t,y_t)-E(x_t)E(y_t)}{\left(E(x_t^2)-E(x_t)^2\right)^{\frac{3}{2}}\sqrt{E(y_t^2)-E(y_t)^2}} \\ \frac{-E(x_t)}{\sqrt{E(x_t^2)-E(x_t)^2}\sqrt{E(y_t^2)-E(y_t)^2}} + \frac{E(x_t,y_t)-E(x_t)E(y_t)}{\sqrt{E(x_t^2)-E(x_t)^2}\left(E(y_t^2)-E(y_t)^2\right)^{\frac{3}{2}}}E(y_t) \\ -\frac{1}{2}\frac{E(x_t,y_t)-E(x_t)E(y_t)}{\sqrt{E(x_t^2)-E(x_t)^2}\left(E(y_t^2)-E(y_t)^2\right)^{\frac{3}{2}}} \\ \frac{1}{\sqrt{E(x_t^2)-E(x_t)^2}\sqrt{E(y_t^2)-E(y_t)^2}} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{\sigma_x}\left(\rho\frac{E(x_t)}{\sigma_x} - \frac{E(y_t)}{\sigma_y}\right) \\ -\frac{\rho}{2\sigma_x^2} \\ \frac{1}{\sigma_y}\left(\rho\frac{E(y_t)}{\sigma_y} - \frac{E(x_t)}{\sigma_x}\right) \\ -\frac{\rho}{2\sigma_y^2} \\ \frac{1}{\sigma_x\sigma_y} \end{bmatrix}$$

The S matrix is $[E(x_t)\ E(x_t^2)\ E(y_t)\ E(y_t^2)\ E(x_t y_t)]$

$$S = \sum_{j=-\infty}^{\infty} \begin{bmatrix} E(x_t x_{t-j}) & E(x_t x_{t-j}^2) & E(x_t y_{t-j}) & E(x_t y_{t-j}^2) & E(x_t x_{t-j} y_{t-j}) \\ E(x_t^2 x_{t-j}) & E(x_t^2 x_{t-j}) & E(x_t^2 y_{t-j}) & E(x_t^2 y_{t-j}^2) & E(x_t^2 x_{t-j} y_{t-j}) \\ E(y_t x_{t-j}) & E(y_t x_{t-j}) & E(y_t y_{t-j}) & E(y_t y_{t-j}^2) & E(y_t x_{t-j} y_{t-j}) \\ E(y_t^2 x_{t-j}) & E(y_t^2 x_{t-j}) & E(y_t^2 y_{t-j}) & E(y_t^2 y_{t-j}^2) & E(y_t^2 x_{t-j} y_{t-j}) \\ E(x_t y_t x_{t-j}) & E(x_t y_t x_{t-j}) & E(x_t y_t y_{t-j}) & E(x_t y_t y_{t-j}^2) & E(x_t y_t x_{t-j} y_{t-j}) \end{bmatrix}$$

I must have seen a pretty way to simplify this when I wrote the problem, but I can't seem to see one now. Still, it's easy enough to compute.

2, 3. The general formula:

$$var(\hat{\beta}) = \frac{1}{T}E(xx')^{-1}\sum_{j=-\infty}^{\infty} E(\varepsilon_t x_t x'_{t-j}\varepsilon_{t-j})E(x_t x'_t)$$

If you believe in homoskedasticity, then $E(\varepsilon_t x_t x'_{t-j}\varepsilon_{t-j}) = E(\varepsilon_t\varepsilon_{t-j})E(x_t x'_{t-j}) = \sigma_\varepsilon^2\rho_j E(x_t x'_{t-j})$. Then,

$$var(\hat{\beta}) = \frac{1}{T}\sigma_\varepsilon^2 E(xx')^{-1}\sum_{j=-\infty}^{\infty}\rho_j E(x_t x'_{t-j})E(x_t x'_t)$$

In retrospect, 2 and 3 seem the same. I think I meant in 2 to correct for *heteroskedasticity* but not *autocorrelation*, in which case the answer is

$$var(\hat{\beta}) = \frac{1}{T}E(xx')^{-1}E(\varepsilon_t^2 x_t x'_t)E(x_t x'_t)$$

4. Under the null that the model is true, a good asset pricing model should have

$$1 = E_t\left(m_{t+1}R_{t+1}\right).$$

Thus, the errors, $u_{t+1} = m_{t+1}R_{t+1} - 1$ should be unpredictable from anything z_t at time t, including $z_t = u_{t-1}$. Thus, $E(u_t u_{t-j}) = 0$. Yes, even if returns are predictable. The point of the asset pricing model is that even if $E(z_t R_{t+1}) \neq 0$, discounted returns should not be forecastable, so $E\left[z_t\left(m_{t+1}R_{t+1} - 1\right)\right] = 0$, including $z_t = u_t$. It's still true if the error is formed from an instrument or managed portfolio. If $f_{t+1} = z_t\left(m_{t+1}R_{t+1} - 1\right)$, the model predicts $E_t(f_{t+1}) = 0$, so $E(f_t f_{t+1}) = 0$.

9 Problems for Chapter 12

1. The test assets can be risky return differences, but the market excess return must be relative to a risk free rate proxy (which may be an estimated parameter). $E(R^i) - R^f = \beta_{i,m}\left(E(R^m) - R^f\right)$ implies $E(R^i - R^j) = \beta_{i-j,m}E(R^m - R^f)$ but not $E(R^i - R^j) = \beta_{i-j,m-j}\left(E(R^m - R^j)\right)$. Betas add in the left hand variable, but not in the right hand variable.

2. No. The GRS test requires factors that are returns.

3. Pricing errors can be correlated with betas with time-series regressions. Not with a cross-sectional OLS regression. Cross-sectional regressions set the right hand variable – betas – orthogonal to the error term – alphas. They can again be correlated with a GLS cross-sectional regression.

4. The cross-sectional regression with an intercept sets the average pricing error to zero. The pricing error of the equally weighted portfolio is, of course, the average pricing error. This regression does not necessarily pass through the origin or risk free rate.

10 Problems for chapter 13

1. The d matrix giving the derivative of moments with respect to parameters $(b', E(f)')$ is

$$d = \begin{bmatrix} -E\left(R^e\tilde{f}'\right) & E(R^e)b' \\ 0 & -I_K \end{bmatrix}$$

where $\tilde{f} = f - E(f)$ and K is the number of factors. The estimates $a_T g_T(b, E(f)) = 0$ give the OLS cross-sectional regression for b, and the sample mean for $E(f)$.

$$\begin{aligned} \hat{b} &= \left[C'C\right]^{-1}C'E_T(R^e) \qquad (1) \\ E(f) &= E_T(f). \end{aligned}$$

where

$$C \equiv E(R^e \tilde{f}')$$

denotes the covariance matrix of returns and factors. (It is best to reserve d for the d matrix, and the two are no longer equal.)

To find the standard errors, just plug in to the general GMM formulas. The general formula (**??**) is

$$cov \left(\begin{array}{c} \hat{b} \\ \hat{E}(f) \end{array} \right) = \frac{1}{T}(ad)^{-1} aSa'(ad)^{-1\prime}.$$

Filling in the pieces, the S matrix is

$$\begin{aligned} S &= \sum_{j=-\infty}^{\infty} E \left[\begin{array}{cc} u_t u'_{t-j} & u_t \tilde{f}'_{t-j} \\ \tilde{f}_t u'_{t-j} & \tilde{f}_t \tilde{f}'_{t-j} \end{array} \right] \\ u_t &\equiv R^e_t(1 - \tilde{f}'_t b). \end{aligned}$$

We can simplify S somewhat with the null hypothesis that pricing errors u_t should not be forecastable from $t-j$ information,

$$S = \left[\begin{array}{cc} E\left(u_t u'_t\right) & \sum_{j=0}^{\infty} E\left(u_t \tilde{f}'_{t+j}\right) \\ \sum_{j=0}^{\infty} E\left(\tilde{f}_{t+j} u'_t\right) & \sum_{j=-\infty}^{\infty} E\left(\tilde{f}_t \tilde{f}'_{t-j}\right) \end{array} \right].$$

However, the factors need not be unpredictable, and may comove with the pricing errors, so no further simplification is possible in general. The other terms are

$$ad = \left[\begin{array}{cc} -C' & 0 \\ 0 & -I_K \end{array} \right] \left[\begin{array}{cc} -C & E(R^e)b' \\ 0 & -I_K \end{array} \right] = \left[\begin{array}{cc} C'C & -C'E(R^e)b' \\ 0 & I_K \end{array} \right]$$

$$(ad)^{-1} = \left[\begin{array}{cc} (C'C)^{-1} & (C'C)^{-1}C'E(R^e)b' \\ 0 & I_K \end{array} \right]$$

$$(ad)^{-1}a = - \left[\begin{array}{cc} (C'C)^{-1}C' & (C'C)^{-1}C'E(R^e)b' \\ 0 & I_K \end{array} \right].$$

Under the null, asymptotically, we will have $E(R^e) = C'b$ so we can simplify the formula now with that substitution.

$$(ad)^{-1}a = - \left[\begin{array}{cc} (C'C)^{-1}C' & bb' \\ 0 & I_K \end{array} \right].$$

We're interested in the top left and bottom right elements of

$$\frac{1}{T}(ad)^{-1}a \left[\begin{array}{cc} S_{11} & S_{12} \\ S_{21} & S_{22} \end{array} \right] a'(ad)^{-1\prime}.$$

The bottom right element is thus

$$var(E_T(f)) = \frac{1}{T} S_{22}.$$

This is the standard formula for the variance of the sample mean. The top left element is

$$var(\hat{b}) = \frac{1}{T} \left((C'C)^{-1}C'S_{11}C(C'C)^{-1} + bb'S_{22}bb' + (C'C)^{-1}C'S_{12}bb' + bb'S_{21}(C'C)^{-1}C' \right)$$

This equation reminds us a great deal of the correction for cross-sectional regressions of average returns on betas. The first term is the same standard error we derived ignoring sampling variation in the sample mean, and looks like the usual formula for OLS regressions with standard errors corrected for covariation. The remaining terms add the effects of the fact that the sample mean must be estimated, as the extra terms in the Shanken formula correct for the fact that the betas had to be estimated.

Next, we want to test the pricing errors. Use the genral formula,

$$Tcov\left[g_T(\hat{b})\right] = \left(I - d(ad)^{-1}a\right) S \left(I - d(ad)^{-1}a\right)'.$$

We have $(ad)^{-1}a$. Then

$$\begin{aligned} I - d(ad)^{-1}a &= I - \left[\begin{array}{cc} C & -E(R^e)b' \\ 0 & I_K \end{array}\right]\left[\begin{array}{cc} (C'C)^{-1}C' & bb' \\ 0 & I_K \end{array}\right] \\ &= \left[\begin{array}{cc} I - C(C'C)^{-1}C' & \left(E(R^e) - Cb\right)b' \\ 0 & 0 \end{array}\right] \end{aligned}$$

Under the null, $E(R^e) - Cb = 0$, so the top right term vanishes. Since the $E(f)$ moment is zero in every sample, the last K diagonal elements of $cov(g_T)$ are zero. The top left part, in which we are interested, gives us

$$cov(\hat{\alpha}) = \frac{1}{T}\left(I - C(C'C)^{-1}C'\right) S_{11} \left(I - C(C'C)^{-1}C'\right)$$

Thus, the pricing error test statistic is not affected by the fact that the factor mean $E(f)$ is estimated. This is natural, since the $E(f)$ moments are set to zero in each sample.

2. If we really want to distinguish factor models based on factor risk premia λ, we can do it by using *single* regression betas in the expected return - beta model. However, though the new λ are useful for testing for the marginal importance of factors, they lose their interpretation as the expected returns on factor mimicking portfolios.

   To see this, define β^s_{ij} as the single regression coefficient of return i on a constant and factor j *alone.*

   $$\beta^s_{ij} = \frac{cov(r, f_j)}{var(f_j)}.$$

   The vector of such betas across all factors is

   $$\beta^s_i = diag\left[cov(\tilde{f}, \tilde{f}')\right]^{-1} cov\left(r, \tilde{f}\right)$$

   Picking up our proof that $m = f'b$ is equivalent to a beta model, we can multiply and divide by the diagonal of $cov(\tilde{f}\tilde{f}')$ rather than that covariance matrix itself to express the model in terms of single regression betas,

   $$E\left(R^i\right) = \alpha\left[1 - cov(R^i, \tilde{f}')\tilde{b}\right]$$

   $$= \alpha\left[1 - cov(R^i, \tilde{f}')diag\left[cov(\tilde{f}, \tilde{f}')\right]^{-1} diag\left[cov(\tilde{f}, \tilde{f}')\right]\tilde{b}\right]$$

   $$= \alpha\left[1 - \beta^{s\prime}_i diag\left[cov(\tilde{f}, f')\right] b\right]$$

$$E(R^i) = \alpha + \beta^{s\prime}\lambda^s.$$

Now the "factor risk premia" are defined as

$$\lambda^s \equiv -\alpha \; diag\left[cov(\tilde{f}, f')\right] b$$

Now $\lambda_j^s = 0$ if and only if $b_j = 0$, so a test on factor risk premia so defined is equivalent to a test whether factor j is marginally important. Surprisingly, the $\lambda^s s$ defined from *single* regression betas deliver a *multiple* regression test for the importance of a factor given all the others!

However, single regression based λ^s do not have the interpretation as the price of the factors in the interesting case that factors are correlated. A factor that is a return will have a single regression beta of one on itself, but will also have nonzero single regression betas on the other (correlated) factors. Therefore, the beta pricing model does not imply that the factor risk premium λ^s equals the expected excess return of the factor.

11 Problems for Chapter 14

1. Adding pricing errors to the time-series regression equation, we obtain

$$R_t^{ei} = \alpha_i + \beta_i'\lambda + \beta_i'\left[f_t - E(f_t)\right] + \varepsilon_t^i.$$

Stacking assets $i = 1, 2, ...N$ to a vector

$$R_t^e = \alpha + B\lambda + B\left[f_t - E(f_t)\right] + \varepsilon_t$$

where B denotes a $N \times K$ matrix of regression coefficients of the N assets on the K factors.

If we fit this model, maximum likelihood will give asset-by asset OLS estimates of the intercept $a = \alpha + B(\lambda - E(f_t))$ and slope coefficients B. It will not give separate estimates of α and λ. The most that the regression can hope to estimate is one intercept; if one chooses a higher value of λ, we can obtain the same error term with a lower value of α. The likelihood surface is flat over such choices of α and λ. One could do an ad-hoc second stage, minimizing (say) the sum of squared α to choose λ given B, $E(f_t)$ and a. This intuitively appealing procedure is exactly a cross-sectional regression. But it would be ad-hoc, not ML.

2. Instead of writing a regression, build up the ML for the CAPM a little more formally. Write the statistical model as just the assumption that individual returns and the market return are jointly normal,

$$\begin{bmatrix} R^e \\ R^{em} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} E(R^e) \\ E(R^{em}) \end{bmatrix}, \begin{bmatrix} \Sigma & cov(R^{em}, R^{e\prime}) \\ cov(R^{em}, R^e) & \sigma_m^2 \end{bmatrix} \right)$$

The model's restriction is

$$E(R^e) = \gamma cov(R^{em}, R^e).$$

Estimate γ and show that this is the same time-series estimator as we derived by presupposing a regression.

Answer: The likelihood function is

$$L = -\frac{1}{2}\sum_{t=1}^{T}\left(\begin{bmatrix} R^e \\ R^{em} \end{bmatrix} - \begin{bmatrix} E(R^e) \\ E(R^{em}) \end{bmatrix} \right)' \begin{bmatrix} \Sigma & cov(R^{em}, R^{e\prime}) \\ cov(R^{em}, R^e) & \sigma_m^2 \end{bmatrix}^{-1} \left(\begin{bmatrix} R^e \\ R^{em} \end{bmatrix} - \begin{bmatrix} E(R^e) \\ E(R^{em}) \end{bmatrix} \right).$$

Imposing the model

$$E(R^e) = \gamma cov(R^{em}, R^e)$$
$$E(R^{em}) = \gamma \sigma_m^2$$

the restricted likelihood function is

$$\begin{aligned} L &= -\frac{1}{2}\sum_{t=1}^{T}\left(\begin{bmatrix} R^e \\ R^{em} \end{bmatrix} - \gamma \begin{bmatrix} cov(R^{em}, R^e). \\ \sigma_m^2 \end{bmatrix}\right)' \begin{bmatrix} \Sigma & cov(R^{em}, R^{e\prime}) \\ cov(R^{em}, R^e) & \sigma_m^2 \end{bmatrix}^{-1} \times \\ &\quad \times \left(\begin{bmatrix} R^e \\ R^{em} \end{bmatrix} - \gamma \begin{bmatrix} cov(R^{em}, R^e) \\ \sigma_m^2 \end{bmatrix}\right). \end{aligned}$$

$$\frac{\partial L}{\partial \gamma} = \sum_{t=1}^{T} \begin{bmatrix} cov(R^{em}, R^e) \\ \sigma_m^2 \end{bmatrix}' \begin{bmatrix} \Sigma & cov(R^{em}, R^{e\prime}) \\ cov(R^{em}, R^e) & \sigma_m^2 \end{bmatrix}^{-1} \left(\begin{bmatrix} R^e \\ R^{em} \end{bmatrix} - \gamma \begin{bmatrix} cov(R^{em}, R^e) \\ \sigma_m^2 \end{bmatrix}\right) = 0.$$

$$\begin{bmatrix} cov(R^{em}, R^e) \\ \sigma_m^2 \end{bmatrix}' \begin{bmatrix} \Sigma & cov(R^{em}, R^{e\prime}) \\ cov(R^{em}, R^e) & \sigma_m^2 \end{bmatrix}^{-1} \left(\begin{bmatrix} E_T(R^e) \\ E_T(R^{em}) \end{bmatrix} - \gamma \begin{bmatrix} cov(R^{em}, R^e) \\ \sigma_m^2 \end{bmatrix}\right) = 0.$$

$$\frac{\begin{bmatrix} cov(R^{em}, R^e) \\ \sigma_m^2 \end{bmatrix}' \begin{bmatrix} \Sigma & cov(R^{em}, R^{e\prime}) \\ cov(R^{em}, R^e) & \sigma_m^2 \end{bmatrix}^{-1} \begin{bmatrix} E_T(R^e) \\ E_T(R^{em}) \end{bmatrix}}{\begin{bmatrix} cov(R^{em}, R^e) \\ \sigma_m^2 \end{bmatrix} \begin{bmatrix} \Sigma & cov(R^{em}, R^{e\prime}) \\ cov(R^{em}, R^e) & \sigma_m^2 \end{bmatrix}^{-1} \begin{bmatrix} cov(R^{em}, R^e) \\ \sigma_m^2 \end{bmatrix}} = \gamma = 0.$$

To make it easy, use only one test asset (you can use the partitioned matrix inverse formulas to do the same thing with many test assets)

$$\begin{bmatrix} \Sigma & cov(R^{em}, R^e) \\ cov(R^{em}, R^e) & \sigma_m^2 \end{bmatrix}^{-1} = \frac{\begin{bmatrix} \sigma_m^2 & -cov(R^{em}, R^e) \\ -cov(R^{em}, R^e) & \Sigma \end{bmatrix}}{\Sigma\sigma_m^2 - cov(R^{em}, R^e)^2}$$

Then,

$$\begin{aligned} \gamma &= \frac{\begin{bmatrix} cov(R^{em}, R^e) \\ \sigma_m^2 \end{bmatrix}' \begin{bmatrix} \sigma_m^2 & -cov(R^{em}, R^e) \\ -cov(R^{em}, R^e) & \Sigma \end{bmatrix} \begin{bmatrix} E_T(R^e) \\ E_T(R^{em}) \end{bmatrix}}{\begin{bmatrix} cov(R^{em}, R^e) \\ \sigma_m^2 \end{bmatrix} \begin{bmatrix} \sigma_m^2 & -cov(R^{em}, R^e) \\ -cov(R^{em}, R^e) & \Sigma \end{bmatrix} \begin{bmatrix} cov(R^{em}, R^e) \\ \sigma_m^2 \end{bmatrix}} \\ &= \frac{cov(R^{em}, R^e)\sigma_m^2 E_T(R^e) - cov(R^{em}, R^e)^2 E_T(R^{em}) - \sigma_m^2 cov(R^{em}, R^e) E_T(R^e) + \sigma_m^2 \Sigma E_T(R^{em})}{\sigma_m^2\left(\Sigma\sigma_m^2 - cov(R^{em}, R^e)cov(R^{em}, R^e)\right)} \\ &= \frac{-cov(R^{em}, R^e)^2 + \sigma_m^2\Sigma}{(\Sigma\sigma_m^2 - cov(R^{em}, R^e)^2)} \frac{E_T(R^{em})}{\sigma_m^2} = \frac{E_T(R^{em})}{\sigma_m^2} \end{aligned}$$

Look at the big picture: γ is estimated from the market return alone,

$$E_T(R^{em}) = \gamma \sigma_m^2$$

completely ignoring the rest of the model

$$E(R^e) = \gamma cov(R^{em}, R^e)$$

This is the time-series regression.

12 Problems for Chapter 17

1. You might want to exercise American puts early.

2. Retrace the steps in the integral derivation of the Black-Scholes formula and show that the dw does not affect the final result.

   A:

$$
\begin{aligned}
\ln S_T &= \ln S_0 + \left(\mu - \frac{\sigma^2}{2}\right) T + \sigma\sqrt{T}\varepsilon \\
\ln \Lambda_T &= \ln \Lambda_0 - \left(r + \frac{1}{2}\left(\frac{\mu - r}{\sigma}\right)^2 + \frac{1}{2}\sigma_w^2\right) T - \frac{\mu - r}{\sigma}\sqrt{T}\varepsilon - \sigma_w\sqrt{T}\delta
\end{aligned}
$$

$$
C_0 = \int_{S_T=X}^{\infty} \int \frac{\Lambda_T(\varepsilon)}{\Lambda_t} S_T(\varepsilon)\, df\,(\varepsilon)\, df(\delta) - \int_{S_T=X}^{\infty} \int \frac{\Lambda_T(\varepsilon)}{\Lambda_t} X\, df\,(\varepsilon)\, df(\delta).
$$

$$
\begin{aligned}
C_0 &= \int_{S_T=X}^{\infty} \int e^{-\left(r+\frac{1}{2}\left(\frac{\mu-r}{\sigma}\right)^2+\frac{1}{2}\sigma_w^2\right)T-\frac{\mu-r}{\sigma}\sqrt{T}\varepsilon-\sigma_w\sqrt{T}\delta} S_0 e^{\left(\mu-\frac{1}{2}\sigma^2\right)T+\sigma\sqrt{T}\varepsilon} f\,(\delta)\, d\delta f(\varepsilon) d\varepsilon \\
&\quad -X \int_{S_T=X}^{\infty} e^{-\left(r+\frac{1}{2}\left(\frac{\mu-r}{\sigma}\right)^2+\frac{1}{2}\sigma_w^2\right)T-\frac{\mu-r}{\sigma}\sqrt{T}\varepsilon-\sigma_w\sqrt{T}\delta} f(\delta) d\delta f\,(\varepsilon)\, d\varepsilon \\
&= S_0 \int_{S_T=X}^{\infty} \left[\int e^{-\frac{1}{2}\sigma_w^2 T-\sigma_w\sqrt{T}\delta} f(\delta) d\delta\right] e^{\left[\mu-r-\frac{1}{2}\left(\sigma^2+\left(\frac{\mu-r}{\sigma}\right)^2\right)\right]T+\left(\sigma-\frac{\mu-r}{\sigma}\right)\sqrt{T}\varepsilon} f(\varepsilon) d\varepsilon \\
&\quad -X \int_{S_T=X}^{\infty} \left[\int e^{-\frac{1}{2}\sigma_w^2 T-\sigma_w\sqrt{T}\delta} f(\delta) d\delta\right] e^{-\left(r+\frac{1}{2}\left(\frac{\mu-r}{\sigma}\right)^2\right)T-\frac{\mu-r}{\sigma}\sqrt{T}\varepsilon} f\,(\varepsilon)
\end{aligned}
$$

   We can bring the [] term out in front. Evaluating it,

$$
\begin{aligned}
& \int e^{-\frac{1}{2}\sigma_w^2 T-\sigma_w\sqrt{T}\delta} f(\delta) d\delta \\
=\ & \frac{1}{\sqrt{2\pi}} \int e^{-\frac{1}{2}\sigma_w^2 T-\sigma_w\sqrt{T}\delta-\frac{1}{2}\delta^2} d\delta \\
& \frac{1}{\sqrt{2\pi}} \int e^{-\frac{1}{2}\left(\delta+\sigma_w\sqrt{T}\right)^2} d\delta
\end{aligned}
$$

   That is the integral under a normal distribution with mean $\sigma_w\sqrt{T}$ and standard deviation 1, which is 1. Thus, we multiply both terms of the Black-Scholes formula by 1, which does not change them.

3. From (17.2), express the Black-Scholes discount factor as a function of the stock and bond price. This expression shows that the Black-Scholes model is equivalent to pricing an index option with the CAPM.

   Answer

$$
\begin{aligned}
\frac{d\Lambda}{\Lambda} &= -rdt - \frac{(\mu - r)}{\sigma^2}\left(\frac{dS}{S} - \mu dt\right) \\
&= \left[\frac{(\mu - r)\mu}{\sigma^2} - r\right] dt - \frac{(\mu - r)}{\sigma^2}\left(\frac{dS}{S}\right)
\end{aligned}
$$

$$
\ln \Lambda_t - \ln \Lambda_0 = \left[\frac{(\mu - r)\mu}{\sigma^2} - r\right] t - \frac{(\mu - r)}{\sigma^2}\left(\ln S_t - \ln S_0\right)
$$

13 Problems for chapter 18

14 Problems for Chapter 19

15 Problems for Chapter 20

1. Go back to the derivation on p.399. We start with

$$p_{t-1} - d_{t-1} = const + \sum_{j=0}^{\infty} \rho^j \left(\Delta d_{t+j} - r_{t+j}\right)$$

Then, take $E(\cdot|I_t) - E(\cdot|I_{t-1})$ of both sides. As long as p_t, d_t and r_t are in I_t, we can conclude

$$0 = [E(\cdot|I_t) - E(\cdot|I_{t-1})] \sum_{j=0}^{\infty} \rho^j \left(\Delta d_{t+j} - r_{t+j}\right)$$

and hence the identity

$$r_t - E(r_t|I_{t-1}) = \Delta d_t - E(\Delta d_t|I_{t-1}) + [E(\cdot|I_t) - E(\cdot|I_{t-1})] \sum_{j=1}^{\infty} \rho^j \left(\Delta d_{t+j} - r_{t+j}\right).$$

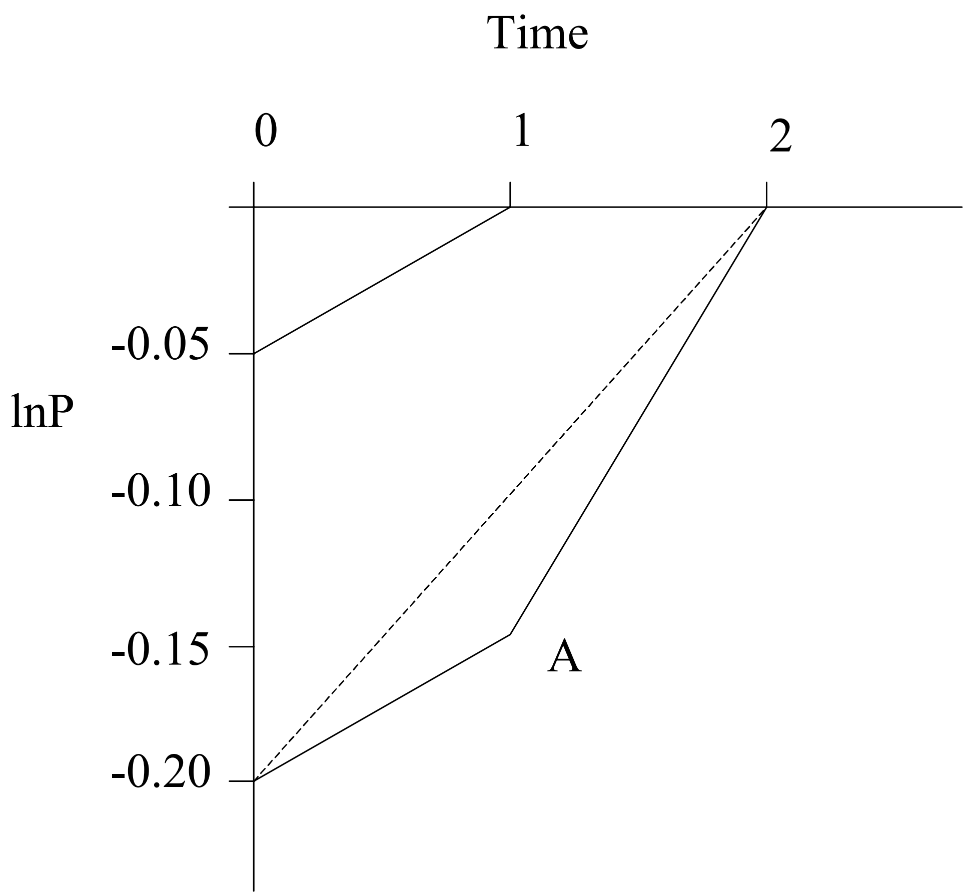
Hence,

   (a) The equation can work for information sets that are coarser than agent's, including the information set of a VAR. However,

   (b) The results will depend on the information set you choose. Adding more variables to a VAR can change the fraction of return variance attributed to the different components.

2. (Fama and Bliss 1987) See the term structure chapter for notation.

   The basic idea, is that there is a mechanical connection between the first period holding period return and the second period yield. The following picture considers what happens if the one year yield is 5% and the two year yield is 10%. As you can see, the expectations hypothesis – expected returns the same for the first year – mean that the one period yield must rise to 15% the next year. You can also see that as we change the first year holding period return, we *mechanically* change the second year return (one year yield in second year). Move point "A" up and down.



This holds for expected values as well as ex post values, or values forecast by a variables such as the forward - spot spread – the more goes to a change in yield $y_{t+1}^{(1)} - y_t^{(1)}$ the less goes to a holding period return $hpr_{t+1}^{(2)}$.

Now, to express the same idea formally.

$$p_t^{(2)} = -y_{t+1}^{(1)} - hpr_{t\to t+1}^{(2)}$$

that's cool in its own right. It says that the bond price is the discounted value of 1 dollar ($\ln(1) = 0$), discounted by the bond's returns. This should remind you of the Campbell Shiller identity that you can discount a stock price by its ex-post returns. You can see that given prices at t, the change in yield $y_{t+1}^{(1)}$ and the holding period return $hpr_{t\to t+1}^{(2)}$ are *mechanically* linked. We want to say something about forward rates, so let's get there

$$\begin{aligned}
p_t^{(1)} &= -y_t^{(1)} \\
f_t^{(1\to 2)} &= p_t^{(1)} - p_t^{(2)} = -y_t^{(1)} - p_t^{(2)} \\
f_t^{(1\to 2)} &= -y_t^{(1)} + y_{t+1}^{(1)} + hpr_{t\to t+1}^{(2)} \\
f_t^{(1\to 2)} - y_t^{(1)} &= \left(y_{t+1}^{(1)} - y_t^{(1)}\right) + \left(hpr_{t\to t+1}^{(2)} - y_t^{(1)}\right).
\end{aligned}$$

Aha! The left hand variable is the forward-spot spread in the Fama Bliss regression. The right hand term is the change in one year yield, and the holding period return on two year bonds. Now run a regression of both sides on the forward-spot spread. The left hand side is 1 – forward spot on forward spot. The right hand side gives the coefficient b_1 in

$$\left(y_{t+1}^{(1)} - y_t^{(1)}\right) = a + b_1 \left(f_t^{(1\to 2)} - y_t^{(1)}\right) + \varepsilon_{t+1}$$

and the coefficient b_2 in

$$\left(hpr_{t\to t+1}^{(2)} - y_t^{(1)}\right) = a_2 + b_2 \left(f_t^{(1\to 2)} - y_t^{(1)}\right) + \varepsilon_{t+1}$$

Thus,

$$1 = b_1 + b_2.$$

See Fama and Bliss (1987) for the general case – longer maturities.

$$r_t = \Delta d_t + \rho(p_t - d_t) - (p_{t-1} - d_{t-1})$$

$$r_t - E_{t-1}(r_t) = \Delta d_t - E_{t-1}(\Delta d_t) + \rho (E_t - E_{t-1})(p_t - d_t)$$

$$\begin{aligned} r_t - E_{t-1}(r_t) &= \Delta d_t - E_{t-1}(\Delta d_t) + \rho (E_t - E_{t-1}) \sum_{j=1}^{\infty} \rho^{j-1} (\Delta d_{t+j} - r_{t+j}) \\ & \Delta d_t - E_{t-1}(\Delta d_t) + (E_t - E_{t-1}) \sum_{j=1}^{\infty} \rho^{j} (\Delta d_{t+j} - r_{t+j}) \end{aligned}$$

3. Conceptually, what we're doing is simple. Imagine simulating out a huge number of data points from the VAR. Then, take only the return data, ignoring data on other variables. Run a regression of returns on lagged returns. We're looking for that regression. That's straightforward to do numerically for a specific example. The algebra is hard, because we're analytically deriving the result of this operation. You just have to go through and do for prices what we did for returns. The case is simple enough that you can follow the same procedure as in "How to find the univariate return representation" on p.418. (You don't have to follow the general procedure, constructing the spectral density and factoring it.) The equation for prices is, from (20.19)

$$\Delta p_{t+1} = (1-b)(d_t - p_t) + \left(\varepsilon_{dt+1} - \frac{1}{1-\rho b}\delta_{t+1}\right)$$

Thus, the system – analogous to the equations in the middle of p.419 – is

$$\begin{aligned} \Delta p_{t+1} &= (1-b)(d_t - p_t) + (\varepsilon_{dt+1} - \varepsilon_{dpt+1}) \\ d_{t+1} - p_{t+1} &= b(d_t - p_t) + \varepsilon_{dpt+1} \end{aligned}$$

(Yes, that's a typo on p.419 just above "Then, write returns." It should be p_{t+1} not p_t.) As for returns, find an expression with just price growth and shocks,

$$\begin{aligned} (1-bL)(d_t - p_t) &= \varepsilon_{dpt} \\ \Delta p_{t+1} &= \frac{1-b}{1-bL}\varepsilon_{dpt} + (\varepsilon_{dt+1} - \varepsilon_{dpt+1}) \\ (1-bL)\,\Delta p_{t+1} &= (1-b)\,\varepsilon_{dpt} + (1-bL)(\varepsilon_{dt+1} - \varepsilon_{dpt+1}) \\ (1-bL)\,\Delta p_{t+1} &= (\varepsilon_{dt+1} - \varepsilon_{dpt+1}) + (\varepsilon_{dpt} - b\varepsilon_{dt}) \end{aligned}$$

Again, the form suggests an ARMA(1,1)

$$\Delta p_t = \frac{1-\gamma L}{1-bL} v_t$$

Let's try it. As in the book, define $y_t = (1-bL)\Delta p_t$ so

$$\begin{aligned} E\left[y_t^2\right] &= \sigma_d^2 + \sigma_{dp}^2 - 2\sigma_{d,dp} + \sigma_{dp}^2 + b^2\sigma_d^2 - 2b\sigma_{d,dp} = (1+b^2)\sigma_d^2 + 2\sigma_{dp}^2 - 2(1+b)\sigma_{d,dp} \\ E(y_t y_{t-1}) &= -\sigma_{dp}^2 - b\sigma_d^2 - (1+b)\sigma_{d,dp} \end{aligned}$$

Matching the autocorrelations from the ARMA(1,1)

$$\begin{aligned} E(y^2) &= (1+\gamma^2)\sigma_v^2 \\ E(y_t y_{t-1}) &= -\gamma\sigma_v^2 \end{aligned}$$

Thus,

$$\begin{aligned} \frac{1+\gamma^2}{\gamma} &= \frac{(1+b^2)\sigma_d^2 + 2\sigma_{dp}^2 - 2(1+b)\sigma_{d,dp}}{\sigma_{dp}^2 + b\sigma_d^2 + (1+b)\sigma_{d,dp}} = 2q \\ \gamma &= q - \sqrt{q^2-1} \end{aligned}$$

(Looks like we have another typo on the top of p.420, the final - in the denominator should be +). Now, look through the cases on p.,415:

(a) No predictability. $\sigma_{dp} = 0$.

$$\begin{aligned} \frac{1+\gamma^2}{\gamma} &= \frac{(1+b^2)\sigma_d^2}{b\sigma_d^2} \\ \gamma &= b \end{aligned}$$

Again, the two roots cancel so $\Delta p_t = v_t$. Returns are i.i.d, dividend growth is i.i.d., so price growth is also i.i.d.

(b) Constant dividend growth $\sigma_d = 0$

$$\frac{1+\gamma^2}{\gamma} = \frac{2\sigma_{dp}^2}{\sigma_{dp}^2} = 2$$

$$\gamma = 1$$

$$\begin{aligned} \Delta p_t &= \frac{1-L}{1-bL} v_t \\ p_t &= \frac{1}{1-bL} v_t \end{aligned}$$

If there is no dividend growth shock, then *prices become stationary.* This isn't as surprising as it seems initially. Bond prices are stationary, because the only reason a bond price changes is that interest rates change. If we turn off dividend growth uncertainty, stocks become like bonds. You could see this in the VAR too. More subtly, this shows that when we turn off one source of noise, the univariate price shocks now reveal the underlying expected return shocks.

(c) Dividend growth uncorrelated with expected return shocks $\sigma_{d,dp} = 0$

$$\frac{1+\gamma^2}{\gamma} = \frac{(1+b^2)\sigma_d^2 + 2\sigma_{dp}^2}{\sigma_{dp}^2 + b\sigma_d^2} = \frac{1+b^2}{b}\frac{b\sigma_d^2}{\sigma_{dp}^2 + b\sigma_d^2} + \frac{1+1}{1}\frac{\sigma_{dp}^2}{\sigma_{dp}^2 + b\sigma_d^2}$$

Thus, following the logic on p.415, we see that γ is between b and 1 rather than between b and ρ. Using the numbers on the top of p.405,

$$\begin{aligned} q &= \frac{1}{2}\frac{(1+0.9^2)(0.1082)^2 + 2\,(0.125)^2}{(0.125)^2 + 0.9(0.1082)^2}. \\ \gamma &= q - \sqrt{q^2 - 1} = 0.93527 \end{aligned}$$

That's a little higher than the 0.928 that the book reports for returns.

4. Here we go again.

$$\begin{aligned} r_{t+1} &= ax_t + \varepsilon_{rt+1} \\ x_{t+1} &= bx_t + \varepsilon_{xt+1} \end{aligned}$$

$$r_{t+1} = \frac{a}{1-bL}\varepsilon_{xt} + \varepsilon_{rt+1}$$

$$(1-bL)\,r_{t+1} = a\varepsilon_{xt} - b\varepsilon_{rt} + \varepsilon_{rt+1}$$

guess

$$(1-bL)\,r_t = y_t = (1-\gamma L)v_t$$

$$\begin{aligned} E(y_t^2) &= \left(1+\gamma^2\right)\sigma_v^2 = a^2\sigma_x + \left(1+b^2\right)\sigma_r^2 - 2ab\sigma_{xr} \\ E(y_t y_{t-1}) &= -\gamma\sigma_v^2 = -b\sigma_r^2 + a\sigma_{xr} \end{aligned}$$

$$\frac{1+\gamma^2}{\gamma} = \frac{a^2\sigma_x^2 + (1+b^2)\,\sigma_r^2 - 2ab\sigma_{xr}}{b\sigma_r^2 - a\sigma_{xr}}$$

To generate uncorrelated returns, we need $\gamma = b$

$$\frac{1+b^2}{b} = \frac{a^2\sigma_x^2 + (1+b^2)\,\sigma_r^2 - 2ab\sigma_{xr}}{b\sigma_r^2 - a\sigma_{xr}}$$

solving for σ_{xr}

$$\left(\frac{1+b^2}{b}\right)\left(b\sigma_r^2 - a\sigma_{xr}\right) = a^2\sigma_x^2 + \left(1+b^2\right)\sigma_r^2 - 2ab\sigma_{xr}$$

$$\left(1+b^2\right)\sigma_r^2 - a\left(\frac{1+b^2}{b}\right)\sigma_{xr} = a^2\sigma_x^2 + \left(1+b^2\right)\sigma_r^2 - 2ab\sigma_{xr}$$

$$2ab\sigma_{xr} - a\left(\frac{1+b^2}{b}\right)\sigma_{xr} = a^2\sigma_x^2$$

$$\left[2b - \left(\frac{1+b^2}{b}\right)\right]\sigma_{xr} = a\sigma_x^2$$

$$\left(b-\frac{1}{b}\right)\sigma_{xr}=a\sigma_x^2$$

Substituting $\sigma_{xr}=\rho\sigma_x\sigma_r$,

$$\begin{aligned}\rho &= \frac{ab}{(b^2-1)}\frac{\sigma_x}{\sigma_r}\\ &= -\frac{ab}{(1+b)(1-b)}\frac{\sigma_x}{\sigma_r}\end{aligned}$$

As you see, you cannot get a zero correlation to do it. You need a negative correlation between expected return and actual return shocks to generate uncorrelated returns.

5. First, the long run variance of a stationary series must be zero. The definition of covariance stationary is that variances exist and variances and covariances are not functions of time. If a series is stationary, variances exist, so

$$\frac{1}{k}var(x_{t+k}-x_t)=\frac{1}{k}\left(var(x_{t+k})+var(x_t)-2cov(x_t,x_{t+k})\right)$$

The covariance is bounded by the variance, so the whole thing goes to zero as $k\rightarrow\infty$. If x has a unit root, $var(x)$ does not exist (infinite) so you can't do the first step of this.

Now, apply this logic to the stationary series x_t-y_t. and the same for y.

$$\begin{aligned}& \frac{1}{k}var\left[x_{t+k}-y_{t+k}-(x_t-y_t)\right]\\ = & \frac{1}{k}var\left[(x_{t+k}-x_t)-(y_{t+k}-y_t))\right]\\ = & \frac{1}{k}\left(var\left(x_{t+k}-x_t\right)+var\left(y_{t+k}-y_t\right)+2cov\left[\left(x_{t+k}-x_t\right),\left(y_{t+k}-y_t\right)\right]\right)\end{aligned}$$

Since each of x_t and y_t have unit roots, we know

$$\lim_{k\rightarrow\infty}\frac{1}{k}var(x_{t+k}-x_t)=v_x>0.$$

Thus, the only way for the above expression to go to zero is if the series are perfectly correlated in the long run – if $v_x=v_y$ and $cov\left[\left(x_{t+k}-x_t\right),\left(y_{t+k}-y_t\right)\right]=\sigma\left(x_{t+k}-x_t\right)\sigma\left(y_{t+k}-y_t\right)$

6.

$$\begin{aligned}dp_{t+1} &= bdp_t+\varepsilon_{dp_{t+1}}\\ r_{t+1} &= (1-\rho b)dp_t+\varepsilon_{rt+1}\end{aligned}$$

It's easiest to do all this sort of thing in vector form,

$$\begin{aligned}\begin{bmatrix} dp_{t+1}\\ r_{t+1}\end{bmatrix} &= \begin{bmatrix} b & 0\\ (1-\rho b) & 0\end{bmatrix}\begin{bmatrix} dp_t\\ r_t\end{bmatrix}+\begin{bmatrix}\varepsilon_{dpt+1}\\ \varepsilon_{rt+1}\end{bmatrix}\\ y_{t+1} &= Ay_t+\varepsilon_{t+1}\end{aligned}$$

Then

$$\begin{aligned}y_{t+2} &= Ay_{t+1}+\varepsilon_{t+2}\\ &= A^2y_t+\varepsilon_{t+2}+A\varepsilon_{t+1}\\ y_{t+3} &= A^3y_t+\varepsilon_{t+3}+A\varepsilon_{t+2}+A^2\varepsilon_{t+1}\end{aligned}$$

and the long horizon regression is

$$(y_{t+1}+y_{t+2}+y_{t+3}+...y_{t+N}) = (A+A^2+..+A^N)y_t+(\varepsilon_{t+N}+(I+A)\varepsilon_{t+N-1}+...+(I+A+..+A^{N-1})\varepsilon_{t+1})$$

We can do the coefficients analytically.

$$\begin{aligned}
A &= \begin{bmatrix} b & 0 \\ (1-\rho b) & 0 \end{bmatrix} \\
A^2 &= \begin{bmatrix} b^2 & 0 \\ (1-\rho b)\, b & 0 \end{bmatrix} \\
A^3 &= \begin{bmatrix} b^3 & 0 \\ (1-\rho b)\, b^2 & 0 \end{bmatrix} \\
A^j &= \begin{bmatrix} b^j & 0 \\ (1-\rho b)\, b^j & 0 \end{bmatrix}
\end{aligned}$$

Thus, since

$$\sum_{j=1}^{N} b^j = \frac{b-b^N+1}{1-b}$$

the long horizon return regression coefficient is

$$r_{t\to t+N} = \frac{b-b^{N+1}}{1-b}\,(1-\rho b)\,dp_t + e_{r,t+k}$$

As you can see, these rise close to linearly at first, but eventually approach a limit

$$\frac{b}{1-b}\,(1-\rho b)\,.$$

To do the R^2 we need to evaluate the error covariance matrix, the bottom right element of

$$\Sigma + (I+A)\Sigma(I+A)' + (I+A+A^2)\Sigma((I+A+A^2)'.$$

I didn't get anywhere with analytical manipulation of this. To see the R^2 rise with horizon, you have to compute $\sigma^2(x\beta)/\sigma^2(\varepsilon)$ numerically using this formula. (See Hodrick 1992 for lots of calculations like this.)

16 Problems for Chapter 21

1. Suppose habit accumulation is linear, and there is a constant riskfree rate or linear technology equal to the discount rate, $R^f = 1/\delta$. The consumer's problem is then

$$\max \sum_{t=0}^{\infty} \delta^t \frac{(C_t - X_t)^{1-\gamma}}{1-\gamma} \quad s.t. \sum_t \delta^t C_t = \sum_t \delta^t e_t + W_0; \quad X_t = \theta \sum_{j=1}^{\infty} \phi^j C_{t-j}$$

where e_t is a stochastic endowment. In an internal habit specification, the consumer considers all the effects that current consumption has on future utility through X_{t+j}. In an external habit specification, the consumer ignores such terms. Show that the two specifications give *identical* asset pricing predictions in this simple model, by showing that internal-habit marginal utility is proportional to external-habit marginal utility, state by state.

A: The first order conditions are

$$MU_t = E_t\left[MU_{t+1}\right]$$

where MU denotes marginal utility. In the external case, marginal utility is simply

$$MU_t = (C_t - X_t)^{-\gamma}. \tag{2}$$

In the internal case, marginal utility is

$$MU_t = (C_t - X_t)^{-\gamma} - \theta \sum_{j=1}^{\infty} \delta^j \phi^j E_t\left(C_{t+j} - X_{t+j}\right)^{-\gamma} \tag{3}$$

The sum measures the habit-forming effect of consumption. Now, guess the same solution as for the external case,

$$(C_t - X_t)^{-\gamma} = E_t\left[(C_{t+1} - X_{t+1})^{-\gamma}\right]. \tag{4}$$

and plug in to (3). We find that the internal marginal utility is simply proportional to marginal utility (2) in the external case,

$$MU_t = \left(1 - \frac{\theta\delta\phi}{1-\delta\phi}\right)(C_t - X_t)^{-\gamma}. \tag{5}$$

Since this expression satisfies the first order condition $MU_t = E_t MU_{t+1}$, we confirm the guess (4). Ratios of marginal utility are the same, so allocations and asset prices are completely unaffected by internal vs. external habit in this example.

2. Many models predict too much variation in the conditional mean discount factor, or too much interest rate variation. This problem guides you through a simple example. Introduce a simple form of external habit formation,

$$u = (C_t - \theta C_{t-1})^{1-\gamma}$$

and suppose consumption growth C_{t+1}/C_t is i.i.d. Show that there interest rates still vary despite i.i.d. consumption growth.

A:

$$\begin{aligned}
m_{t+1} &= \frac{(C_{t+1} - \theta C_t)^{-\gamma}}{(C_t - \theta C_{t-1})^{-\gamma}} \\
m_{t+1} &= \frac{(C_{t+1}/C_t - \theta)^{-\gamma}}{(C_t/C_{t-1} - \theta)^{-\gamma}}\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} \\
E_t\left(m_{t+1}\right) &= E\left[(C_{t+1}/C_t - \theta)^{-\gamma}\right]\left(\frac{C_t}{C_t - \theta C_{t-1}}\right)^{-\gamma} \\
&= E\left[\left(\frac{C_{t+1}}{C_t} - \theta\right)^{-\gamma}\right]\left(1 - \theta\frac{C_{t-1}}{C_t}\right)^{\gamma}
\end{aligned}$$

The first term is constant, but the second varies as consumption varies.

17 Problems for the Appendix

1. Find the diffusion followed by the log price,
$$y = \ln(p).$$
A: Applying Ito's lemma,
$$dy = \frac{1}{p}dp - \frac{1}{2}\frac{dp^2}{p^2} = \left(\mu - \frac{1}{2}\sigma^2\right)dt + \sigma dz.$$
This is *not*
$$d(\ln(p)) = \frac{dp}{p}.$$
You have to include the second order terms. It matters whether you specify
$$\frac{dp}{p} = \mu dt + \sigma dz$$
or
$$d\ln p = \mu dt + \sigma dz.$$
The two μ terms are not the same; you have to add or subtract $1/2\sigma^2$ to go from one to the other.

2. Find the diffusion followed by xy.
A: Usually, we write
$$d(xy) = xdy + ydx$$
But this expression comes from the usual first order expansions. When x and y are diffusions, we have to keep second order terms. Write $f(x,y) = xy$. $\partial f/\partial x = y$, $\partial f/\partial y = x$, $\partial^2 f/dy^2 = 0$, $\partial^2 f/\partial x^2 = 0$, $\partial^2 f/\partial y \partial x = 1$, so
$$d(xy) = xdy + ydx + dydx.$$
We used this fact in expanding $d(\Lambda p)$.

3. Suppose $y = f(x,t)$ Find the diffusion representaiton for y. (Follow the obvious multivariate extension of Ito's lemma.)
A: Recognizing ahead of time that terms dt^2 and $dtdz$ will drop,
$$dy = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dx + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}dx^2$$
$$dy = \left(\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}\mu_x + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\sigma_x^2\right)dt + \frac{\partial f}{\partial x}\sigma_x dz$$

4. Suppose $y = f(x,w)$, with both x, w diffusions. Find the diffusion representation for y. Denote the correlation between dz_x and dz_w by ρ.
A: First do a second order expansion,
$$dy = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial w}dw + \frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}dx^2 + \frac{\partial^2 f}{\partial w^2}dw^2 + 2\frac{\partial^2 f}{\partial x \partial w}dxdw\right)$$
Then, get rid of terms dt^2 and $dzdt$, and organize the result into dt and dz terms
$$dy = \left(\frac{\partial f}{\partial x}\mu_x + \frac{\partial f}{\partial w}\mu_w + \frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}\sigma_x^2 + \frac{\partial^2 f}{\partial w^2}\sigma_w^2 + 2\frac{\partial^2 f}{\partial x \partial w}\sigma_x\sigma_w\rho\right)\right)dt$$
$$+\frac{\partial f}{\partial x}\sigma_x dz_x + \frac{\partial f}{\partial w}\sigma_w dz_w$$