Teaching Notes #3
(Addendum)

Incomplete Information and Learning: Portfolio Allocation

Pietro Veronesi

University of Chicago

Booth School of Business

This Version: February 13, 2015

Introduction

- As usual, let's define a probability space $(\Omega, P, \mathcal{F})$.
- So far, we have assumed exogenously the existence of a Brownian motion $\{\mathbf{B}_t\}$ defined on $(\Omega, P, \mathcal{F})$ along with the standard (P-augmented) filtration $\{\mathcal{F}_t\}$.
- The filtration $\{\mathcal{F}_t\}$ represents the flow of information hitting the agents in the economy, who little by little get to know the elements in Ω (recall the simple examples in TN1).
- However, we may ask the following questions:
  - If we were to obtain the filtration $\{\mathcal{F}_t\}$ as the result of investors estimating some fundamental process from signals on the economy, how would it "move" over time?
  - How does investors' lack of information about stock market average returns affect their portfolio allocation?
  - Is it possible that imperfect information increases "risk" and thus decreases the allocation to stocks compared to the benchmark case?
- Before we attack these questions, we must find out what is the optimal rule for filtering (i.e. learning).

A Filtering Result

- The following is from Liptser and Shiryayev (1977).
- Suppose

$$d\mathbf{z}_t = [\mathbf{a}_{z0} + \mathbf{a}_{zz}\mathbf{z}_t + \mathbf{a}_{zs}\mathbf{s}_t]\,dt + \mathbf{b}_z d\mathbf{B}_t$$
$$d\mathbf{s}_t = [\mathbf{a}_{s0} + \mathbf{a}_{sz}\mathbf{z}_t + \mathbf{a}_{ss}\mathbf{s}_t]\,dt + \mathbf{b}_s d\mathbf{B}_t$$

- where

1. $\{\mathbf{z}_t\}$ is $(n \times 1)$ vector of state variables;
2. $\{\mathbf{s}_t\}$ is $(m \times 1)$ vector of signals;
3. $\{\mathbf{B}_t\}$ is a $d-$dimensional standard BM defined on a probability space $(\Omega, P, \mathcal{F})$;
4. $\mathbf{a}_{z0}$, $\mathbf{a}_{s0}$, $\mathbf{a}_{zz}$, $\mathbf{a}_{sz}$, $\mathbf{a}_{zs}$, $\mathbf{a}_{ss}$, $\mathbf{b}_z$ and $\mathbf{b}_s$ are matrices of constants of the appropriate dimension.

- Suppose that the prior at time $t = 0$ on $\mathbf{z}_0$ is

$$\mathbf{z}_0 \sim \mathcal{N}(\widehat{\mathbf{z}}_0, \widehat{\mathbf{q}}_0)$$

- where $\widehat{\mathbf{q}}_0$ is a $n \times n$ variance covariance matrix.

- Suppose that agents only observe $\mathbf{s}_t$ and let $\{\mathcal{F}_t\}$ the filtration generated by $\{\mathbf{s}_t\}$, that is, the $\mathcal{F}_t$ be smallest sigma algebra generated by $\{\mathbf{s}_t\}$.
- Then, investors have a *normal* posterior distribution on $\mathbf{z}_t$

$$\mathbf{z}_t|_{\mathcal{F}_t} \sim \mathcal{N}\left(\widehat{\mathbf{z}}_t, \widehat{\mathbf{q}}_t\right)$$

- where $\widehat{\mathbf{z}}_t = E\left[\mathbf{z}_t|\mathcal{F}_t\right]$ satisfies the SDE

$$d\widehat{\mathbf{z}}_t = \left[\mathbf{a}_{z0} + \mathbf{a}_{zz}\widehat{\mathbf{z}}_t + \mathbf{a}_{zs}\mathbf{s}_t\right]dt + \left[\widehat{\mathbf{q}}_t\mathbf{a}_{sz}^{\prime} + \mathbf{b}_z\mathbf{b}_s^{\prime}\right]\left(\mathbf{b}_s\mathbf{b}_s^{\prime}\right)^{-\frac{1}{2}} d\widetilde{\mathbf{B}}_t$$

- and where

$$\widehat{\mathbf{q}}_t = E\left[\left(\mathbf{z} - \widehat{\mathbf{z}}\right)\left(\mathbf{z} - \widehat{\mathbf{z}}\right)^{\prime}|\mathcal{F}_t\right]$$

- is positive definite symmetric matrix given by the solution to the Riccati equation

$$\frac{d\widehat{\mathbf{q}}_t}{dt} = \mathbf{a}_{zz}\widehat{\mathbf{q}}_t + \widehat{\mathbf{q}}_t\mathbf{a}_{zz}^{\prime} + \mathbf{b}_z\mathbf{b}_z^{\prime} - \left[\widehat{\mathbf{q}}_t\mathbf{a}_{sz}^{\prime} + \mathbf{b}_z\mathbf{b}_s^{\prime}\right]\left(\mathbf{b}_s\mathbf{b}_s^{\prime}\right)^{-1}\left[\mathbf{a}_{sz}\widehat{\mathbf{q}}_t + \left(\mathbf{b}_z\mathbf{b}_s^{\prime}\right)^{\prime}\right] \qquad (1)$$

- In addition,

$$\begin{aligned} d\widetilde{\mathbf{B}}_t &= \left(\mathbf{b}_s\mathbf{b}_s'\right)^{-\frac{1}{2}}\left(d\mathbf{s}_t - E\left[d\mathbf{s}_t|\mathcal{F}_t\right]\right) \\ &= \left(\mathbf{b}_s\mathbf{b}_s'\right)^{-\frac{1}{2}}\left(d\mathbf{s}_t - \left[\mathbf{a}_{s0} + \mathbf{a}_{sz}\widehat{\mathbf{z}}_t + \mathbf{a}_{ss}\mathbf{s}_t\right]dt\right) \end{aligned}$$

- is a $d-$dimensional standard Brownian motion with respect to the new filtration $\{\mathcal{F}_t\}$.
- A few properties to notice:

  1. Since the prior at time $t = 0$ was normally distributed, because of the linearity of both the signals and the state variable processes, also the posterior is normally distributed.
  2. The normal posterior distribution of $\mathbf{z}_t$ has a stochastic mean $\widehat{\mathbf{z}}_t$ (which evolves according to some diffusion) but *deterministic* variance covariance matrix $\widehat{\mathbf{q}}_t$. The latter does not depend on shocks $\left\{d\widetilde{\mathbf{B}}_t\right\}$.
     - Indeed, one can solve for the steady state $\overline{\mathbf{q}}$ by setting the LHS of (1) equal to zero:

$$\frac{d\widehat{\mathbf{q}}_t}{dt} = 0$$

3. The posterior mean $\widehat{\mathbf{z}}_t$ evolves according to the same linear process that was postulated at the beginning, but with different variance. The latter changes:

$$from \quad \mathbf{b}_z \quad to \quad \left[\widehat{\mathbf{q}}_t \mathbf{a}_{zz}^{/} + \mathbf{b}_z \mathbf{b}_s'\right] (\mathbf{b}_s \mathbf{b}_s')^{-\frac{1}{2}}$$

4. There is a new standard BM defined, which is a normalized expectation error.

- It turns out that if the prior at time $t = 0$ on $\mathbf{z}_0$ is *not* normally distributed, then the posterior is not normally distributed either. However, it can be fully characterized by two variables, representing the mean and the variance (covariance) of the distribution. See Lyptser and Shiryayev (1977) and Detemple (1992).

Example

- Consider the case where there is a single state variable and a single signal

$$dz_t = [a_{z0} + a_{zz} z_t]\, dt + b_z dB_{1,t}$$
$$ds_t = z_t dt + b_s dB_{2,t}$$

- Assume that $dB_{1,t} dB_{2,t} = 0$.
- From the result above we have that the posterior distribution on z_t is normal, with mean

$$d\widehat{z}_t = [a_{z0} + a_{zz}\widehat{z}_t]\, dt + [\widehat{q}_t a_{zz}]\, b_s^{-1} d\widetilde{B}_{2,t}$$

- where $\widehat{q}_t = E_t\left[(z_t - \widehat{z}_t)^2\right]$ is the mean square error and it follows the deterministic process

$$\frac{d\widehat{q}_t}{dt} = 2a_{zz}\widehat{q}_t + b_z^2 - \left(\frac{\widehat{q}_t}{b_s}\right)^2$$

- Finally, the innovation process is

$$d\widetilde{B}_{2,t} = b_s^{-1}\left(ds_t - E\left[ds_t | \mathcal{F}_t\right]\right) = b_s^{-1}\left(ds_t - \widehat{z}_t dt\right)$$

- One can find the stationary mean square error by setting $\frac{d\widehat{q}_t}{dt} = 0$, that is, find the q^* that satisfies

$$0 = 2a_{zz}q^* + b_z^2 - \left(\frac{q^*}{b_s}\right)^2$$

- Given this solution, the only part that is time varying in the posterior distribution is the mean

$$d\widehat{z}_t = \left[a_{z0} + a_{zz}\widehat{z}_t\right]dt + \left[q^* a_{zz}\right] b_s^{-1} d\widetilde{B}_{2,t}$$

The Portfolio Allocation Problem

- Consider the same portfolio allocation problem with time-varying average returns discussed in TN1 (Addendum.)
- Let $(\beta, \mathbf{S})$ be processes of the form

$$d\beta_t = r_t \beta_t dt \text{ with } \beta_0 > 0 \quad (2)$$

$$dS_t^i = \mu_t^i S_t^i dt + S_t^i \boldsymbol{\sigma}_t^i d\mathbf{B}_{1,t} \text{ with } S_0^i > 0 \quad (3)$$

- where r_t, $\boldsymbol{\mu}_t = \left(\mu_t^1, ..., \mu_t^d\right)$ and $\boldsymbol{\sigma}^i$ are bounded, adapted processes.
- For concreteness, assume that $\mu_t = \left(\mu_t^1, ..., \mu_t^d\right)$ follow a continuous time, VAR process

$$d\boldsymbol{\mu}_t = (\mathbf{A}_0 + \mathbf{A}_1 \boldsymbol{\mu}_t)\, dt + \boldsymbol{\Sigma} d\mathbf{B}_{2,t} \quad (4)$$

- Notice that differently from TN 1 (addendum), the set of Brownian motions moving $\boldsymbol{\mu}_t$ is different from the ones moving S_t.
- Instead, assume $r_t = r$ is constant and that $\boldsymbol{\sigma}_t^i = \boldsymbol{\sigma}^i$ is also constant, for every i.
- Assume again that $\{\boldsymbol{\sigma}\}$ is invertible.

Learning

- Differently from TN 1 (Addendum), we assume that $\boldsymbol{\mu}_t$ is *not observable*. That is, investors do not have full information about the stock returns average returns.
- However, they can learn about $\boldsymbol{\mu}_t$ by observing realized returns: If a stock consistently produces high returns, for instance, it is perhaps that his average return μ_t^i is high.
- What do investors know?
- We assume that they know the parameters $\mathbf{A}_0$, $\mathbf{A}_1$, $\boldsymbol{\Sigma}$
  - This is of course not a good assumption. But it tremendously simplify the model, and the effects are interesting.
  - Note that the classic case of constant but unobservable average returns $\boldsymbol{\mu}$ is a special case where $\mathbf{A}_0 = \mathbf{A}_1 = \boldsymbol{\Sigma} = \mathbf{0}$. In this case, there are no parameters that we assume known.
- The maximization problem of our investor is the standard one

$$\max_{c,\boldsymbol{\theta}} E_0 \left[\int_0^T e^{-\rho t} \frac{C_t^{1-\gamma}}{1-\gamma} dt \right] \tag{5}$$

- subject to the budget equation

$$\tilde{\boldsymbol{\theta}}_t \tilde{\mathbf{S}}_t = w + \int_0^t \tilde{\boldsymbol{\theta}}_u d\tilde{\mathbf{S}}_u - \int_0^t C_u du$$

- Note that clearly, the maximization of the investor must be performed conditional on his/her information set.

Learning about Average Returns

- Before moving to the optimal allocation problem, we first apply the learning results.
- This implies that we first solve the filtering problem, and then we solve the portfolio problem.
- This particular sequence of solutions to optimal problems (learning is the result of an optimal filtering problem) is not obvious.
- This result goes under the name of *Separation Theorems*. Feldmann (2004) contains a brief discussion of these problems, and the fact that they hold trivially in economics problem, as preferences do not depend on beliefs directly, but only through the choice variable C_t.
- Define for simplicity the "return" process

$$R_t^i = \int_0^t \left(S_u^i\right)^{-1} dS_u^i$$

- That is, the one defined by (for intuition)

$$dR_t^i = \frac{dS_t^i}{S_t^i}$$

- Stack all the processes in (3) on top of each other

$$d\mathbf{R}_t = \boldsymbol{\mu}_t dt + \boldsymbol{\sigma} d\mathbf{B}_{1,t} \tag{6}$$

- We can apply the optimal filtering result to (6) and (4).
- Let at $t = 0$

$$\boldsymbol{\mu}_0 \sim N\left(\widehat{\boldsymbol{\mu}}_0, \widehat{\mathbf{q}}_0\right)$$

- and for every $t > 0$ denote

$$\begin{aligned} \widehat{\boldsymbol{\mu}}_t &= E\left[\boldsymbol{\mu} | \mathcal{F}_t\right] \\ \widehat{\mathbf{q}}_t &= E\left[\left(\boldsymbol{\mu}_t - \widehat{\boldsymbol{\mu}}_t\right)\left(\boldsymbol{\mu}_t - \widehat{\boldsymbol{\mu}}_t\right)' | \mathcal{F}_t\right] \end{aligned}$$

- Then, we have the following system

$$d\mathbf{R}_t = \widehat{\boldsymbol{\mu}}_t dt + \boldsymbol{\sigma} d\widehat{\mathbf{B}}_t \tag{7}$$

$$d\widehat{\boldsymbol{\mu}}_t = \left(\mathbf{A}_0 + \mathbf{A}_1 \widehat{\boldsymbol{\mu}}_t\right) dt + \widehat{\boldsymbol{\Sigma}}_t d\widehat{\mathbf{B}}_t \tag{8}$$

- where

$$\widehat{\boldsymbol{\Sigma}}_t = \widehat{\mathbf{q}}_t \left(\boldsymbol{\sigma}'\right)^{-1} \tag{9}$$

$$\frac{d\widehat{\mathbf{q}}_t}{dt} = \mathbf{A}_1 \widehat{\mathbf{q}}_t + \widehat{\mathbf{q}}_t \mathbf{A}_1' + \boldsymbol{\Sigma}\boldsymbol{\Sigma}' - \widehat{\mathbf{q}}_t \left(\boldsymbol{\sigma}\boldsymbol{\sigma}'\right)^{-1} \widehat{\mathbf{q}}_t \tag{10}$$

- and the innovation process is given by

$$d\widehat{\mathbf{B}}_t = \boldsymbol{\sigma}^{-1}\left[d\mathbf{R}_t - E_t\left(d\mathbf{R}_t\right)\right] \tag{11}$$

- **Comments:**

  1. We are back in complete markets: Conditional of investors' information, the set of BMs that drive returns $d\mathbf{R}_t$ is the same that drive expected return $\widetilde{\boldsymbol{\mu}}_t$.
     - The reason is that the information filtration $\{\mathcal{F}_t\}$ is generated by the return process $d\mathbf{R}_t$.
     - Thus, expected returns will depend on the observation of $d\mathbf{R}_t$ only: if we observe high returns we change our posterior to on expected future returns. That is, expected returns and realized returns become perfectly correlated.

  2. The only difference from (7) and (8) and the problem discussed in TN 1 (addendum) is the fact that the volatility of $\widehat{\boldsymbol{\mu}}_t$ depends on t.
     - However, this volatility declines deterministically.
     - Thus, the methodology used in TN 1 (addendum) works here too, once we are careful to remember that $\widehat{\Sigma}_t$ is a function of time.

  3. The volatility $\widehat{\Sigma}_t$ converges to its steady state $\widehat{\Sigma}^*$. Once it does so, we are exactly in the same situation as in TN 1 (addendum). We can directly use that technique to solve our problem.

4. Learning has a bite: It has a prediction about the correlation between returns and expected returns.
   - They are *p*ositively correlated: A positive innovation in returns increases expected return.
   - The hedging demand will go in the right direction here: Bad news on returns are "twice bad news". You lost money, and now you expect to gain even less in the future. Thus, since you expect this, you buy less of the stock.
   - (This is opposite of what we found in our calibration in TN 1 (addendum), where we used the "predictability" intuition: negative returns increases the dividend price ratio, which predicts higher returns. That is, realized returns and expected returns were assumed negatively correlated in TN 1 (addendum).)

The Allocation to the Market Portfolio

- Consider again the case discussed in TN 1 (addendum).
- Assume there is only one stock (the market) and so that

$$\begin{aligned} dR_t &= \mu_t dt + \sigma dB_{1,t} \\ d\mu_t &= (a_0 + a_1\mu_t)dt + \Sigma dB_{2,t} \end{aligned}$$

- The filtering result then implies that conditional on investors information filtration $\{\mathcal{F}_t\}$ we have

$$\begin{aligned} dR_t &= \widehat{\mu}_t dt + \sigma d\widehat{B}_t \\ d\widehat{\mu}_t &= (a_0 + a_1\widehat{\mu}_t)dt + \widehat{\Sigma}_t d\widehat{B}_t \end{aligned}$$

- where

$$\widehat{\Sigma}_t = \widehat{q}_t \sigma^{-1}$$

- $\widehat{q}_t$ follows the Riccati equation

$$\frac{d\widehat{q}_t}{dt} = 2a_1\widehat{q}_t + \Sigma^2 - \left(\frac{\widehat{q}_t}{\sigma}\right)^2 \tag{12}$$

- and finally, $\widehat{B}_t$ is given by the innovation process

$$d\widehat{B}_t = \sigma^{-1}\left(dR_t - \widehat{\mu}_t dt\right)$$

- Note: $\widehat{B}_t$ denotes here the BM under the information filtration of agents, and not the risk neutral BM (as in TN 1). Later, we will denote the BM under the risk neutral measure as B_t^Q.
- In principle, we could assume that investors have been around long enough so that $\widehat{q}_t$ has converged to its long term value $\widehat{q}^*$, determined by

$$0 = 2a_1\widehat{q}^* + \Sigma^2 - \left(\frac{\widehat{q}^*}{\sigma}\right)^2$$

- In this case, we obtain the same setting as in TN 1 (addendum) and thus we just have to use the same formulas
- However, leaving $\widehat{\Sigma}_t$ as a function of time is as simple to solve.

Portfolio Holdings

- It may be useful to recall the methodology to obtain the portfolio holdings.

1. Define the market price of risk

$$\nu_t = \frac{\widehat{\mu}_t - r}{\sigma}$$

- which then follows the process

$$\begin{aligned} d\nu_t &= \frac{1}{\sigma} d\widehat{\mu} = \left(\frac{a_0}{\sigma} + a_1 \frac{\widehat{\mu}_t}{\sigma} \right) dt + \frac{\widehat{\Sigma}_t}{\sigma} d\widehat{B}_t \\ &= \left(\tilde{a}_0 + a_1 \nu_t \right) dt + \frac{\widehat{\Sigma}_t}{\sigma} d\widehat{B}_t \end{aligned}$$

- where

$$\tilde{a}_0 = \frac{a_0}{\sigma} + \frac{a_1 r}{\sigma}$$

2. Use the optimality condition

$$C_t = e^{-\frac{\rho}{\gamma}t}\lambda^{-\frac{1}{\gamma}}\pi_t^{-\frac{1}{\gamma}}$$

- to obtain the dynamics of the optimal consumption plan.
- Here π_t is the state price density given by

$$\pi_t = \exp\left(-rt - \int_0^t \nu_\tau^2 d\tau - \int_0^t \nu_\tau d\widehat{B}_\tau\right)$$

- So, the dynamics of optimal consumption plan is as follows: let $c_t = \log(C_t)$

$$dc_t = \left(\frac{1}{\gamma}(r-\rho) + \frac{1}{2\gamma}\nu_t^2\right) + \frac{1}{\gamma}\nu_t d\widehat{B}_t$$

3. Move to the risk neutral measure Q by defining the new BM

$$dB_t^Q = d\widehat{B}_t + \nu_t dt$$

- and translate all the process in the risk neutral measure Q (that is, substitute each $d\widehat{B}_t$ with $dB_t^Q - \nu_t dt$)

$$\begin{aligned} dc_t &= \left(\frac{1}{\gamma}(r-\rho) - \frac{1}{2\gamma}\nu_t^2\right) + \frac{1}{\gamma}\nu_t dB_t^Q \\ d\nu_t &= \left[\tilde{a}_0 + \left(a_1 - \frac{\widehat{\Sigma}_t}{\sigma}\right)\nu_t\right]dt + \frac{\widehat{\Sigma}_t}{\sigma}dB_t^Q \end{aligned}$$

4. Find the portfolio weights: Note that the wealth process must be always equal to the discounted value of future consumption, and thus

$$\begin{aligned} W_t &= E_t^Q \left[\int_t^T e^{-r(\tau - t)} C_\tau d\tau \right] \\ &= C_t E_t^Q \left[\int_t^T e^{-r(\tau - t) + (c_\tau - c_t)} d\tau \right] \\ &= C_t F(\nu_t, t) \end{aligned}$$

- where the last equality stems from the Markovian structure of the system (ν_t) (recall that $\widehat{q}_t$ is just a function of time, and so it is perfectly predictable).
- Ito's Lemma applied to W_t implies then

$$dW_t = \mu_W dt + \sigma_W dB_t^Q \tag{13}$$

- where

$$
\begin{aligned}
\mu_W &= FC_t\left(\frac{1}{\gamma}(r-\rho)+\frac{1}{2\gamma}\left(\frac{1}{\gamma}-1\right)\nu_t^2\right)+C_t\frac{\partial F}{\partial t} \\
&\quad +C_t\frac{\partial F}{\partial \nu}\left\{\left[\tilde{a}_0+\left(a_1-\frac{\widehat{\Sigma}_t}{\sigma}\right)\nu_t\right]\right\} \\
&\quad +\frac{1}{2}C_t\frac{\partial^2 F}{\partial \nu^2}\left(\frac{\widehat{\Sigma}_t}{\sigma}\right)^2+\frac{\partial F}{\partial \nu}\frac{1}{\gamma}\nu_t C_t\frac{\widehat{\Sigma}_t}{\sigma} \\
\sigma_W &= FC_t\frac{1}{\gamma}\nu_t+C_t\frac{\partial F}{\partial \nu}\frac{\widehat{\Sigma}_t}{\sigma}
\end{aligned}
$$

5. The budget equation under the original measure is

$$
dW_t=\left(W_t\left(\vartheta_t(\widehat{\mu}_t-r)+r\right)-C_t\right)dt+W_t\vartheta_t\sigma d\widehat{B}_t \qquad (14)
$$

- but under the risk neutral measure is given by

$$
\begin{aligned}
dW_t &= \left(W_t\left(\vartheta_t(\widehat{\mu}_t-r)+r\right)-C_t\right)dt+W_t\vartheta_t\sigma dB_t^Q-\vartheta_t\sigma\nu_t dt \\
&= \left(W_t r-C_t\right)dt+W_t\vartheta_t\sigma dB_t^Q \qquad (15)
\end{aligned}
$$

6. Equating volatilities of (13) and (15) provides the formula for portfolio weights:

$$W_t \vartheta_t \sigma = F C_t \frac{1}{\gamma} \nu_t + C_t \frac{\partial F}{\partial \nu} \frac{\widehat{\Sigma}_t}{\sigma}$$

- That is (recall $W_t = F C_t$)

$$\vartheta_t = \frac{1}{\gamma} \frac{\widehat{\mu}_t - r}{\sigma^2} + \frac{1}{F} \frac{\partial F}{\partial \nu} \frac{\widehat{\Sigma}_t}{\sigma^2}$$

7. Equating drifts of (13) and (15) provides the formula for F (that is, the PDE). In fact:

$$(W_t r - C_t) = \mu_W$$

- yields (recall that $W = CF$ and thus delete C on both sides):

$$\begin{aligned} Fr - 1 \; &= \; F \left(\frac{1}{\gamma}(r - \rho) + \frac{1}{2\gamma} \left(\frac{1}{\gamma} - 1 \right) \nu_t^2 \right) + \frac{\partial F}{\partial t} \\ &+ \frac{\partial F}{\partial \nu} \left\{ \tilde{a}_0 + \left(a_1 + \left(\frac{1}{\gamma} - 1 \right) \frac{\widehat{\Sigma}_t}{\sigma} \right) \nu_t \right\} + \frac{1}{2} \frac{\partial^2 F}{\partial \nu^2} \left(\frac{\widehat{\Sigma}_t}{\sigma} \right)^2 \end{aligned} \tag{16}$$

- This is the same as equation (17) in TN 1 (addendum).

Solving the PDE

- If we could solve (16) we are done.
- However, solving for this PDE is difficult (although numerically can be done). So, we use the trick we used last time.
- We recall that we can write

$$
\begin{aligned}
W_t &= C_t F(\nu_t, t; T) = E_t^Q \left[\int_t^T e^{-r(\tau - t)} C_\tau d\tau \right] \\
&= \int_t^T E_t^Q \left[e^{-r(\tau - t)} C_\tau \right] d\tau \\
&= \int_t^T C_t f(\nu_t, t; \tau) d\tau
\end{aligned}
\tag{17}
$$

- where the last equality comes from the homogeneity property of the consumption process:

$$
f(\nu_t, t; \tau) = E_t^Q \left[e^{-r(\tau - t) + (c_\tau - c_t)} \right] \tag{18}
$$

- Thus, in short

$$F(\nu_t, t; T) = \int_t^T f(\nu_t, t; \tau) d\tau \tag{19}$$

- where $f(\nu_t, t; \tau)$ satisfies the boundary condition

$$f(\nu_\tau, \tau; \tau) = 1 \tag{20}$$

- We can substitute (19) into the PDE (16). In fact, note that:[1]

$$\begin{aligned}
\frac{\partial F}{\partial t} &= -1 + \int_t^T \frac{\partial f}{\partial t} d\tau \\
\frac{\partial F}{\partial \nu} &= \int_t^T \frac{\partial f}{\partial \nu} d\tau \\
\frac{\partial^2 F}{\partial \nu^2} &= \int_t^T \frac{\partial^2 f}{\partial \nu^2} d\tau
\end{aligned}$$

[1] Recall the rules of differentiation for integral functions: If you have $F(x,t) = int_{a(t)}^{b(t)} f(x,t,\tau) d\tau$ we have that the chain rule has, for instance

$$\frac{\partial F}{\partial t} = b'(t) f(x,t,b(t)) - a'(t) f(x,t,a(t)) + \int_{a(t)}^{b(t)} \frac{\partial f(x,t,\tau)}{\partial t} d\tau$$

- Substitute into the PDE (16) to obtain

$$
\begin{aligned}
r\int_t^T f(\tau)d\tau &= \int_t^T f(\tau)d\tau\left(\frac{1}{\gamma}(r-\rho)+\frac{1}{2\gamma}\left(\frac{1}{\gamma}-1\right)\nu_t^2\right)+\int_t^T\frac{\partial f}{\partial t}d\tau \\
&\quad +\int_t^T\frac{\partial f}{\partial \nu}d\tau\left\{\tilde{a}_0+\left(a_1+\left(\frac{1}{\gamma}-1\right)\frac{\widehat{\Sigma}_t}{\sigma}\right)\nu_t\right\} \\
&\quad +\frac{1}{2}\int_t^T\frac{\partial^2 f}{\partial \nu^2}d\tau\left(\frac{\widehat{\Sigma}_t}{\sigma}\right)^2
\end{aligned}
$$

- which is clearly satisfied if for every t and τ we have the following PDE satisfied:

$$
\begin{aligned}
rf &= f\left(\frac{1}{\gamma}(r-\rho)+\frac{1}{2\gamma}\left(\frac{1}{\gamma}-1\right)\nu_t^2\right)+\frac{\partial f}{\partial t} \\
&\quad +\frac{\partial f}{\partial \nu}\left\{\left[\tilde{a}_0+\left(a_1+\left(\frac{1}{\gamma}-1\right)\frac{\widehat{\Sigma}_t}{\sigma}\right)\nu_t\right]\right\}+\frac{1}{2}\frac{\partial^2 f}{\partial \nu^2}\left(\frac{\widehat{\Sigma}_t}{\sigma}\right)^2
\end{aligned}
\tag{21}
$$

- We use again the method of undetermined coefficient to solve for this PDE. Conjecture

$$
f(\nu,t;\tau)=\exp\left(b_0(t;\tau)+b_1(t;\tau)\nu+b_2(t;\tau)\nu^2\right) \tag{22}
$$

- This yields

$$\frac{\partial f}{\partial t} = f\left(b_0' + b_1'\nu + b_2'\nu^2\right); \frac{\partial f}{\partial \nu} = f(b_1 + 2b_2\nu); \frac{\partial^2 f}{\partial \nu^2} = f(2b_2 + (b_1 + 2b_2\nu)^2)$$

- Substitute into PDE

$$\begin{aligned} 0 = & \frac{1}{\gamma}(r-\rho) - r + \frac{1}{2\gamma}\left(\frac{1}{\gamma} - 1\right)\nu_t^2 + \left(b_0' + b_1'\nu + b_2'\nu^2\right) \\ & + (b_1 + 2b_2\nu)\left\{\tilde{a}_0 + \left(a_1 + \left(\frac{1}{\gamma} - 1\right)\frac{\widehat{\Sigma}_t}{\sigma}\right)\nu_t\right\} \\ & + \frac{1}{2}(2b_2 + (b_1 + 2b_2\nu)^2)\left(\frac{\widehat{\Sigma}_t}{\sigma}\right)^2 \end{aligned}$$

- Pull together coefficients in ν and obtain the system of ODEs

$$
\begin{aligned}
b_2' + \left(a_1 + \left(\frac{1}{\gamma} - 1\right)\frac{\widehat{\Sigma}_t}{\sigma}\right) 2b_2 + 2b_2^2\left(\frac{\widehat{\Sigma}_t}{\sigma}\right)^2 + \frac{1}{2\gamma}\left(\frac{1}{\gamma} - 1\right) &= 0 \\
b_1' + b_1\left(a_1 + \left(\frac{1}{\gamma} - 1\right)\frac{\widehat{\Sigma}_t}{\sigma}\right) + 2\tilde{a}_0 b_2 + 2b_1 b_2\left(\frac{\widehat{\Sigma}_t}{\sigma}\right)^2 &= 0 \\
b_0' + \frac{1}{\gamma}(r - \rho) - r + b_1\tilde{a}_0 + \frac{1}{2}(2b_2 + b_1^2)\left(\frac{\widehat{\Sigma}_t}{\sigma}\right)^2 &= 0
\end{aligned}
$$

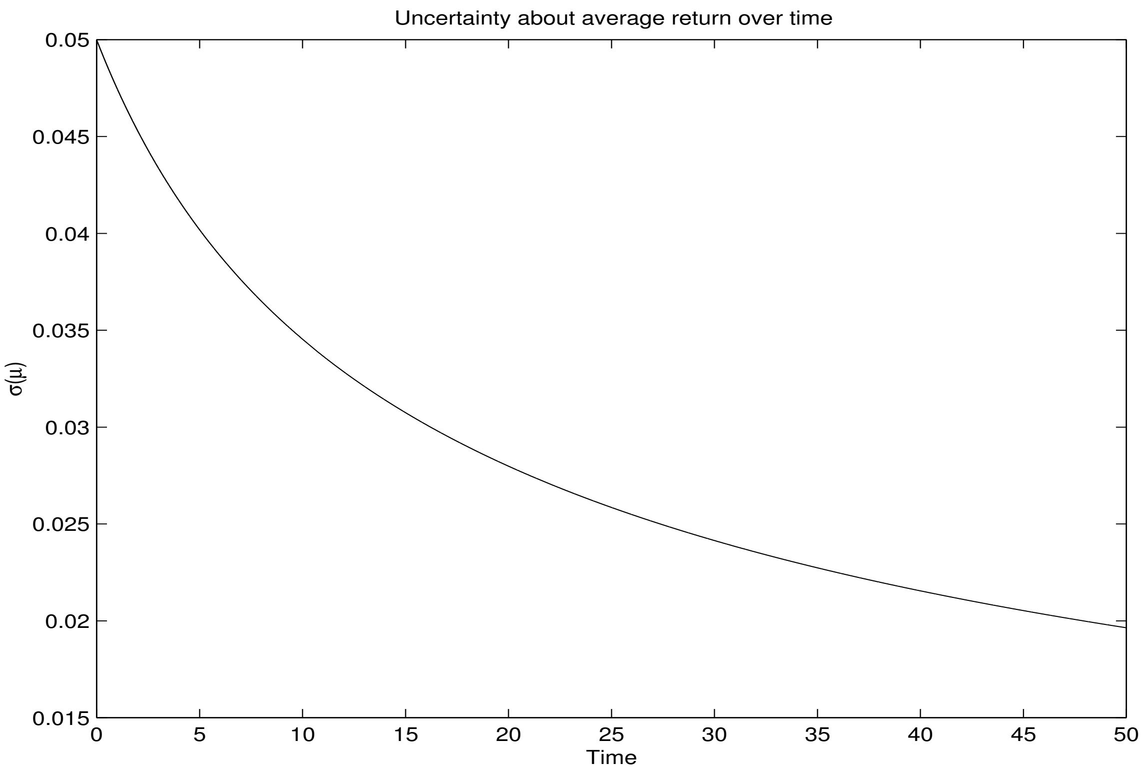
- These are the same as in TN 1 (addendum), except that now Σ_t depends on time t.
- The (numerical) solution strategy is straightforward:
  - First, iterate forward the equation for $\widehat{q}_t$ in equation (12) from 0 to T and obtain the time series of $\widehat{\Sigma}_t = \widehat{q}_t/\sigma$.
  - Second, given $\widehat{\Sigma}_t$ for every t, solve the ODEs, starting with b_2, then b_1 and then b_0.
- Once we compute $f(\nu_t, t; \tau)$ for every τ from (22), we can compute the demand from the usual formula

$$
\begin{aligned}
\text{Myopic Demand} &= \frac{1}{\gamma}\frac{\mu_t - r}{\sigma^2} \\
\text{Hedging Demand} &= \frac{1}{F}\frac{\partial F}{\partial \nu}\frac{\widehat{\Sigma}_t}{\sigma^2}
\end{aligned}
$$

Calibration

- Here I consider a special case of the setting above, namely, the one where μ_t is in fact constant, but unobservable.
- That is, investors believe that average market returns are constant. They just ignore what such constant is
- In the above formulas, we just need to set $a_0 = a_1 = \Sigma = 0$.
- In this case, the uncertainty $\sqrt{\widehat{q}_t}$ about μ decreases over time to converge to 0 as $t \to \infty$
- However, the learning speed is not fast, because returns are noisy.
- Figure 1 reports this quantity over time, assuming a large initial uncertainty at time $t = 0$ equals 5%

Figure 1: Uncertainty



- The most important effect of learning is that hedging demand this time is negative.
- The intuition, recall, is that bad news are twice bad news here: not only you get a negative return to stock, but now you expected even lower returns for the future. Thus, investors' optimally reduce their holding of stocks. This mechanism was first observed by Brennan (1998, European Finance Review), but then analyzed by many others.
- Figure 2 shows the hedging demand and Figure 3 total demand for three different value of initial uncertainty $\sqrt{\widehat{q}_0}$

- The parameter used are the same as in TN 1 (addendum), namely

Parameter Choice	
Rate of time preference ρ	0.0624
Risk free rate r	0.0168
Volatility of stock prices σ	0.1510
γ	5
T	50
Uncertainty $\sqrt{q_0}$	3%

Figure 2: Hedging Demand

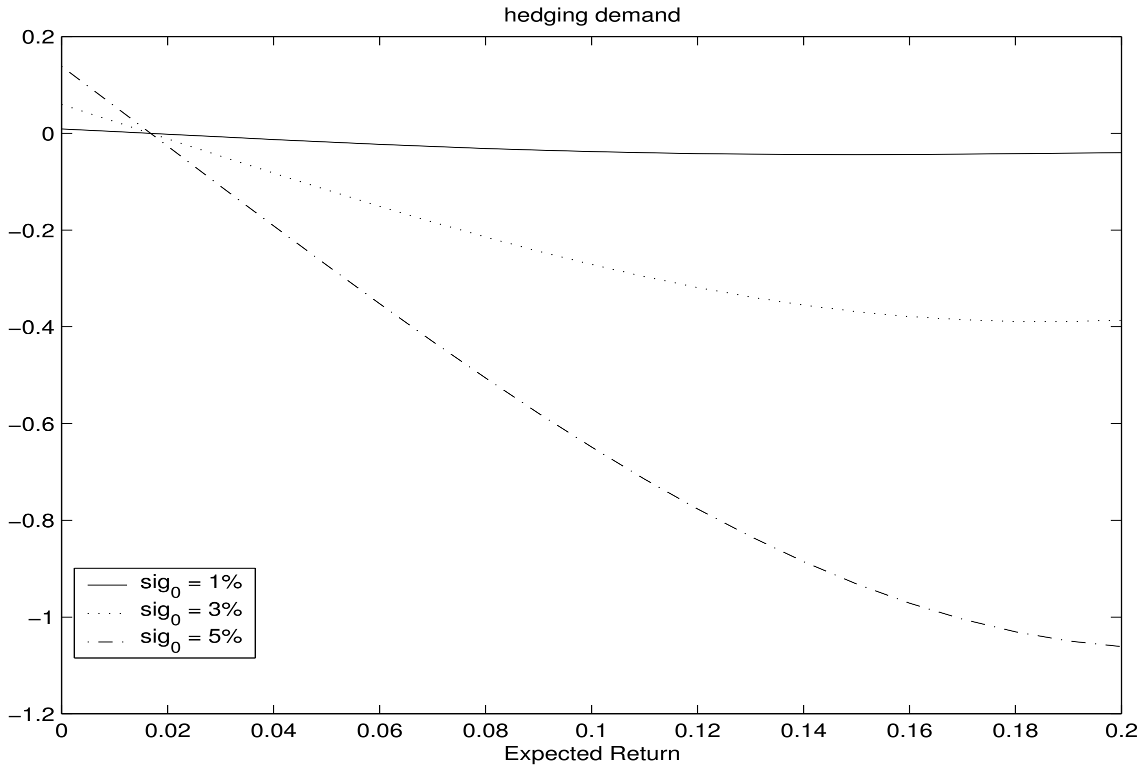
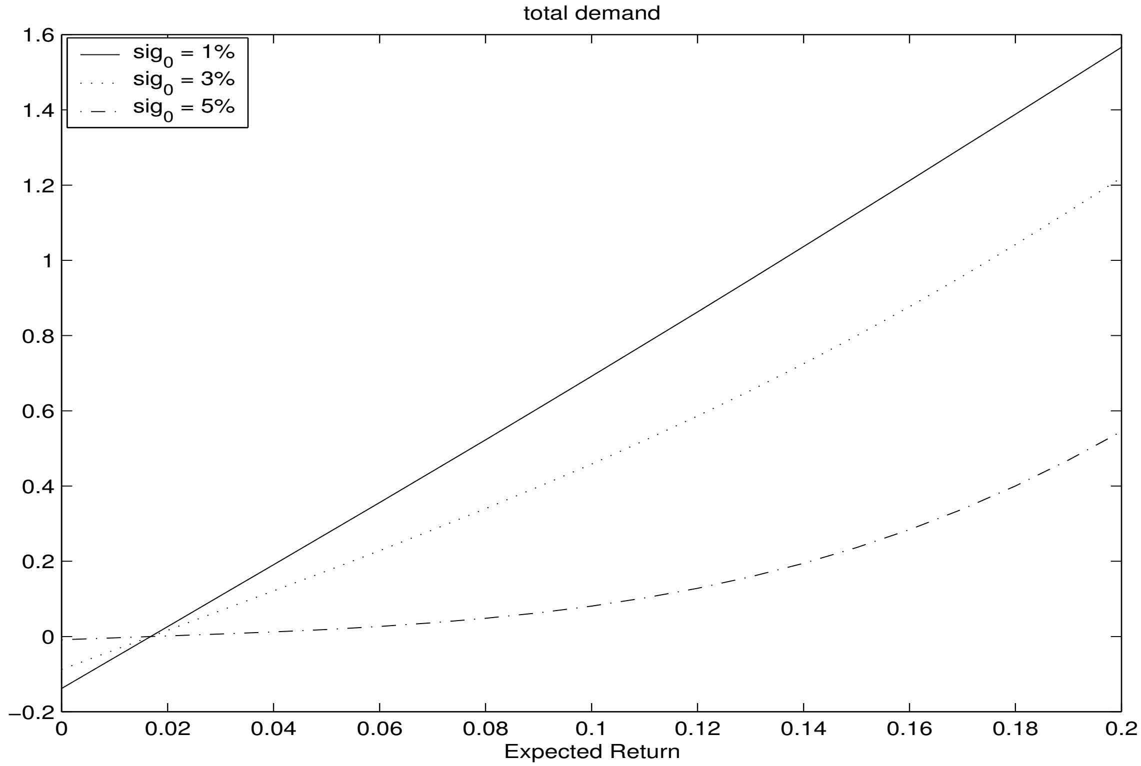


Figure 3: Total Demand



- As it can be guessed, the effect on hedging demands depends strongly on the level of risk aversion.

- Figure 4 shows the hedging demand and Figure 5 total demand for three different value of the risk aversion parameter γ

Figure 4: Hedging Demand

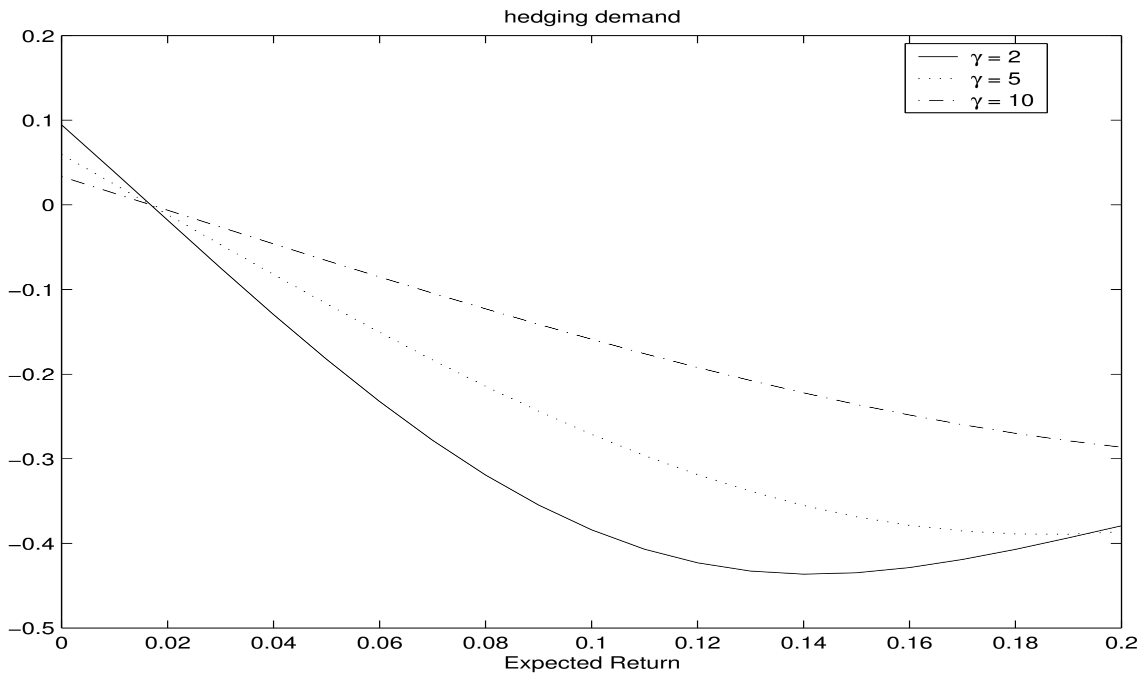


Figure 5: Total Demand