Modern Dynamic Asset Pricing Models

Lecture Notes 1 – Addendum.

Dynamic Portfolio Allocation Strategies

Pietro Veronesi

The University of Chicago
Booth School of Business
CEPR, NBER

Outline

1. Review of Merton / Samuelson Portfolio Allocation Problem
   - The Puzzles
2. Strategic Asset Allocation under Predictability of Stock Returns
   - The Problem and its solution
   - Implications for Dynamic Asset Allocation
3. Learning about Average Returns
   - Implications for Dynamic Asset Allocation
   - Comparison with the case of Predictability
4. Strategic Asset Allocation with Model Misspecification
   - The Problem and Its solution
   - The Example of Constant Investment Opportunity Set
5. Conclusion

Review of Merton/Samuelson Portfolio Allocation Problem

- There are n stocks. Stock i return

$$dR_t^i = \frac{dS_t^i + D_t^i dt}{S_t^i}$$

- $d\mathbf{R}_t = (dR_t^1, .., dR_t^n)'$
- Assume:

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

- $d\mathbf{B}_t = (dB_t^1, ..., dB_t^n)$ = vector of *independent* Brownian motions.
- **Investor problem**:

$$J(W_0, 0) = \max_{\{(C_t),(\boldsymbol{\theta}_t)\}} E_0 \left[\int_0^T u(C_t, t)\, dt \right]$$

- subject to

$$dW_t = \left\{ W_t \left(\boldsymbol{\theta}_t'(\boldsymbol{\mu} - r\mathbf{1}_n) + r \right) - C_t \right\} dt + W_t \boldsymbol{\theta}_t' \boldsymbol{\sigma} d\mathbf{B}_t$$

The Bellman Equation

- Bellman Equation:

$$0 = \sup_{(C_t,\boldsymbol{\theta})} u(C,t) + E\left[dJ(W,t)\right]/dt$$

- with boundary condition $J(W_T,T) = 0$
- Why this form?
  - The discrete time Bellman equation over a small Δ

$$J(W_t,t) = \max_{C,\theta} \left\{u(C,t)\Delta + E\left[J(W_{t+\Delta},t+\Delta)|W_t\right]\right\}$$

$$\Longrightarrow 0 = \max_{C,\theta} u(c,t)\Delta + E_t\left[J(W_{t+\Delta},t+\Delta) - J(W_t,t)\right]$$

- Note that by Ito's Lemma:

$$\begin{aligned} E[dJ(W,t)]/dt &= J_t + J_W E_t[dW]/dt + \frac{1}{2}J_{WW}E_t[dW^2]/dt \\ &= J_t + J_W\left\{W_t\left(\boldsymbol{\theta}_t'(\mu - r) + r\right) - C_t\right\} + \frac{1}{2}J_{WW}W_t^2\boldsymbol{\theta}_t'\boldsymbol{\sigma}\boldsymbol{\sigma}'\boldsymbol{\theta}_t \end{aligned}$$

The Optimal Consumption and Portfolio Allocation

- FOC with respect to C:

$$u_c\left(C_t, t\right) = J_W\left(W, t\right)$$

  – Example: Power utility

$$u\left(C_t, t\right) = e^{-\rho t}\frac{C_t^{1-\gamma}}{1-\gamma} \quad \Longrightarrow \quad C_t = e^{-\frac{\rho}{\gamma}t} J_W\left(W, t\right)^{-\frac{1}{\gamma}}$$

- FOC with respect to $\boldsymbol{\theta}_t$:

$$\boldsymbol{\theta}_t = \frac{1}{RRA(W)}\left(\boldsymbol{\sigma}\boldsymbol{\sigma}'\right)^{-1}\left(\boldsymbol{\mu} - r\mathbf{1}_n\right)$$

  – where

$$RRA(W) = -\frac{W J_{WW}\left(W, t\right)}{J_W\left(W, t\right)}$$

- We now solve for $J(W, t)$ in the power utility case.

The Explicit Solution via an Ordinary Differential Equation

1. Conjecture:

$$J(W,t) = e^{-\rho t}\frac{W^{1-\gamma}}{1-\gamma}F(t)^{\gamma}$$

2. Compute J_t, J_W and J_{WW};
3. Optimal consumption and portfolio holdings:

$$C_t = WF(t)^{-1}; \quad \text{and} \quad \boldsymbol{\theta}_t = \frac{1}{\gamma}(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}(\boldsymbol{\mu} - r\mathbf{1})$$

4. To find $F(t)$, substitute J_t, J_W and J_{WW} and optimal strategies in Bellman equation

$$0 = e^{-\rho t}\frac{C_t^{1-\gamma}}{1-\gamma} - \rho J + J\gamma\frac{F_t}{F} + W^{-1}(1-\gamma)J\left(W_t\left(\boldsymbol{\theta}_t'(\boldsymbol{\mu} - r\mathbf{1}_n) + r\right) - C_t\right)$$

$$-\frac{1}{2}\gamma(1-\gamma)W^{-2}JW_t^2\boldsymbol{\theta}_t'\boldsymbol{\sigma}\boldsymbol{\sigma}'\boldsymbol{\theta}_t$$

The Explicit Solution via a Ordinary Differential Equation

5. Simplify all that can be simplified, to find the ODE

$$0 = 1 - aF\left(t\right) + F_t$$

where $F(T) = 0$ and

$$a = \frac{1}{\gamma}\left\{\rho - \left(1-\gamma\right)r - \frac{1-\gamma}{2\gamma}\left(\boldsymbol{\mu} - r\mathbf{1}_n\right)'\left(\boldsymbol{\sigma}\boldsymbol{\sigma}'\right)^{-1}\left(\boldsymbol{\mu} - r\mathbf{1}_n\right)\right\}$$

6. The solution is

$$F\left(t\right) = \frac{1}{a}\left(1 - e^{-a(T-t)}\right)$$

   - As $t \rightarrow T$, consume a higher fraction of wealth.

7. The last point is to verify that "Conjecture" is indeed optimal.

The Puzzles

- For $n = 1$

$$\theta_t = (\mu - r)/(\gamma\sigma^2)$$

1. θ_t is independent of age t, and thus of remaining life $T - t$.
   - Against empirical evidence: an inverted U shaped θ_t
   - Against the typical recommendation of portfolio advisors.
2. Too large θ. Using $\mu - r = 7\%$ and $\sigma = 16\%$

Table: Portfolio Allocation

	Risk Aversion γ				
	2	4	6	8	10
θ	136%	68%	45%	34 %	27 %

   - Typical household holds between 6 % to 20 % in equity.
   - Conditional on participation, $\approx 40\%$ of financial assets.

Strategic Asset Allocation with Time Varying Expected Returns

- n stocks:

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

  - $\boldsymbol{\mu}_\mathbf{t} = E_t\left[d\mathbf{R}_t\right]$ is now time varying.

- For convenience (later), denote the expected *excess* return

$$\boldsymbol{\lambda}_\mathbf{t} = \boldsymbol{\mu}_t - r\mathbf{1}_n$$

- Assume a VAR process

$$d\boldsymbol{\lambda}_t = \left(\mathbf{A}_0 + \mathbf{A}_\mathbf{1}\boldsymbol{\lambda}_t\right) dt + \boldsymbol{\Sigma} d\mathbf{B}_t$$

- Note:
  - Assume $\boldsymbol{\sigma}$ is now $n \times m$.
  - E.g. $n = 1$ (1 stock), $m = 2$ (two shocks) with

$$\boldsymbol{\sigma} = (\sigma_1, 0) \quad \boldsymbol{\Sigma} = (\Sigma_1, \Sigma_2) \quad \Longrightarrow \text{Cov}\left(dR, d\lambda\right) = \boldsymbol{\sigma}\boldsymbol{\Sigma}' = \sigma_1\Sigma_1$$

The Bellman Equation with Time Varying Expected Returns

- **Investor problem**:

$$J\left(W_0, \lambda_0, 0\right) = \max_{\{(C_t),(\boldsymbol{\theta}_t)\}} E_0\left[\int_0^T u\left(C_t, t\right) dt\right]$$

- subject to

$$dW_t = \left\{W_t\left(\boldsymbol{\theta}_t'\boldsymbol{\lambda}_t + r\right) - C_t\right\} dt + W_t\boldsymbol{\theta}_t'\boldsymbol{\sigma} d\mathbf{B}_t$$

- The Bellman equation is

$$0 \;=\; \max_{C_t,\theta_t} e^{-\rho t}\frac{C_t^{1-\gamma}}{1-\gamma} + E_t\left[dJ_t\right]/dt$$

- with

$$E_t\left[dJ_t\right]/dt \;=\; J_t + J_W E_t\left[dW_t\right] + \frac{1}{2}J_{WW}E_t\left[dW_t^2\right]$$

$$+\mathbf{J}_\lambda' E_t\left[d\boldsymbol{\lambda}_t\right] + \mathbf{J}_{W\lambda}E_t\left[d\boldsymbol{\lambda}_t dW_t\right] + \frac{1}{2}tr\left(\mathbf{J}_{\lambda\lambda}E\left[d\boldsymbol{\lambda}_t d\boldsymbol{\lambda}_t'\right]\right)$$

Optimal Consumption and Portfolio Allocation

- Substitute expectations in Bellman equation:

$$0 = \max_{C_t,\boldsymbol{\theta}_t} e^{-\rho t}\frac{C_t^{1-\gamma}}{1-\gamma} + J_t + J_W\left(W_t\left(\boldsymbol{\theta}_t'\boldsymbol{\lambda}_t + r\right) - C_t\right) + \frac{1}{2}J_{WW}W_t^2\boldsymbol{\theta}_t'\boldsymbol{\sigma}\boldsymbol{\sigma}'\boldsymbol{\theta}_t$$
$$+\mathbf{J}_\lambda'\left(\mathbf{A}_0 + \mathbf{A}_1\boldsymbol{\lambda}_t\right) + \mathbf{J}_{W\lambda}W_t\boldsymbol{\Sigma}\boldsymbol{\sigma}'\boldsymbol{\theta}_t + \frac{1}{2}tr\left(\mathbf{J}_{\lambda\lambda}\boldsymbol{\Sigma}\boldsymbol{\Sigma}'\right)$$

- FOC with respect to C_t:

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

  - Same form as before.
  - But recall that J_W is not different.

- FOC with respect to $\boldsymbol{\theta}_t$:

$$\boldsymbol{\theta}_t = \frac{1}{RRA(W_t)}\left(\boldsymbol{\sigma}\boldsymbol{\sigma}'\right)^{-1}\boldsymbol{\lambda}_t - \left(\boldsymbol{\sigma}\boldsymbol{\sigma}'\right)^{-1}\boldsymbol{\sigma}\boldsymbol{\Sigma}'\frac{\mathbf{J}_{W\lambda}}{J_{WW}W}$$

  - There is one additional term.

Optimal Portfolio Allocation

- Optimal Portfolio Allocation:

$$\boldsymbol{\theta}_t = \boldsymbol{\theta}_t^M + \boldsymbol{\theta}_t^H$$

- Myopic Demand

$$\boldsymbol{\theta}_t^M = \frac{1}{RRA(W_t)} \left(\boldsymbol{\sigma}\boldsymbol{\sigma}'\right)^{-1} \boldsymbol{\lambda}_t$$

  – Same as before.

- Hedging Demand

$$\boldsymbol{\theta}_t^H = -\left(\boldsymbol{\sigma}\boldsymbol{\sigma}'\right)^{-1} \boldsymbol{\sigma}\boldsymbol{\Sigma}' \frac{\mathbf{J}_{W\lambda}}{J_{WW}W}$$

  – Recall that expected returns $\boldsymbol{\lambda}_t$ also (obviously) affect intertemporal utility.

  – $\Longrightarrow$ The asset allocation must "hedge" against the negative impact that the variation in expected returns has on the marginal utility.

- If $\boldsymbol{\theta}_t^H$ depends on age (t) and is negative, we may "resolve" the two puzzles.

Optimal Portfolio Allocation under Power Utility

- Solving this problem is substantially more complicated.
- Conjecture 1:

$$J\left(W_t, \boldsymbol{\lambda}_t, t\right) = e^{-\rho t}\frac{W_t^{1-\gamma}}{1-\gamma}F\left(\boldsymbol{\lambda}_t, t\right)^{\gamma}$$

- Compute J_t, J_W, J_{WW} , $\mathbf{J}_{W\boldsymbol{\lambda}}$, $\mathbf{J}_{\lambda}$ and $\mathbf{J}_{\lambda\lambda}$.
- This yields

$$C_t = W_t F^{-1} \quad \text{and} \quad \boldsymbol{\theta}_t = \frac{1}{\gamma}\left(\boldsymbol{\sigma}\boldsymbol{\sigma}'\right)^{-1}\boldsymbol{\lambda}_t + \left(\boldsymbol{\sigma}\boldsymbol{\sigma}'\right)^{-1}\boldsymbol{\sigma}\boldsymbol{\Sigma}'\frac{\mathbf{F}_{\lambda}}{F}$$

- To solve for $F\left(\lambda, t\right)$, substitute everything into the Bellman equation.

The Bellman Equation and its Solution

$$
\begin{aligned}
0 \;=\; & F^{-1} + ((1-\gamma)\, r - \rho)\,\frac{1}{\gamma} + \frac{F_t}{F} + \frac{1}{2} tr\left(\frac{\mathbf{F}_{\lambda\lambda}}{F}\boldsymbol{\Sigma}\boldsymbol{\Sigma}'\right) + \frac{(1-\gamma)}{2\gamma^2}\boldsymbol{\lambda}_t'\left(\boldsymbol{\sigma}\boldsymbol{\sigma}'\right)^{-1}\boldsymbol{\lambda}_t + \\
& + \frac{(1-\gamma)}{\gamma}\frac{\mathbf{F}_\lambda'}{F}\boldsymbol{\Sigma}\boldsymbol{\sigma}'\left(\boldsymbol{\sigma}\boldsymbol{\sigma}'\right)^{-1}\boldsymbol{\lambda}_t + \frac{\mathbf{F}_\lambda'}{F}\left(\mathbf{A}_0 + \mathbf{A}_1\boldsymbol{\lambda}_t\right) + \\
& + \frac{1}{2}(1-\gamma)\, tr\left(\left(\frac{\mathbf{F}_\lambda}{F}\frac{\mathbf{F}_\lambda'}{F}\right)\left(\boldsymbol{\Sigma}\boldsymbol{\sigma}'\left(\boldsymbol{\sigma}\boldsymbol{\sigma}'\right)^{-1}\boldsymbol{\sigma}\boldsymbol{\Sigma}' - \boldsymbol{\Sigma}\boldsymbol{\Sigma}'\right)\right)
\end{aligned}
$$

- This is horrible. There is:
  - A quadratic term in $\boldsymbol{\lambda}_t$;
  - A linear term in $\boldsymbol{\lambda}_t$;
  - A quadratic term in $\mathbf{F}_\lambda$.

- Yet, by applying recent techniques developed in Fixed Income, an analytical solution exists for the case

$$\boldsymbol{\Sigma}\boldsymbol{\sigma}'\left(\boldsymbol{\sigma}\boldsymbol{\sigma}'\right)^{-1}\boldsymbol{\sigma}\boldsymbol{\Sigma}' - \boldsymbol{\Sigma}\boldsymbol{\Sigma}' = \mathbf{0}$$

Towards an Analytical Solution

- Conjecture 2:

$$F(\boldsymbol{\lambda},t;T) = \int_t^T f(\boldsymbol{\lambda},t;\tau)\,d\tau$$

- with $f(\lambda,t,t) = 1$.
- After some algebra, we find the following PDE for $f(\boldsymbol{\lambda}_t,t;\tau)$:

$$\begin{aligned} 0 \;=\; & ((1-\gamma)\,r-\rho)\,\frac{1}{\gamma}f + f_t + \frac{1}{2}tr\left(\mathbf{f}_{\boldsymbol{\lambda}\boldsymbol{\lambda}}\boldsymbol{\Sigma}\boldsymbol{\Sigma}'\right) + \frac{(1-\gamma)}{2\gamma^2}\boldsymbol{\lambda}_t'\left(\boldsymbol{\sigma}\boldsymbol{\sigma}'\right)^{-1}\boldsymbol{\lambda}_t f + \\ & + \frac{(1-\gamma)}{\gamma}\mathbf{f}_{\boldsymbol{\lambda}}'\boldsymbol{\Sigma}\boldsymbol{\sigma}'\left(\boldsymbol{\sigma}\boldsymbol{\sigma}'\right)^{-1}\boldsymbol{\lambda}_t + \mathbf{f}_{\boldsymbol{\lambda}}'\left(\mathbf{A}_0 + \mathbf{A}_1\boldsymbol{\lambda}_t\right) \end{aligned}$$

- Perhaps this does not look any better to most, but it is a very standard PDE in Fixed Income Asset Pricing.
  - The solution is an exponential linear-quadratic function of $\boldsymbol{\lambda}_t$

An Analytical Solution

- Use method of undetermined coefficients.
- Conjecture 3:

$$f(\boldsymbol{\lambda}, t; \tau) = e^{\alpha_0(t;\tau) + \boldsymbol{\alpha}_1(t;\tau)'\boldsymbol{\lambda}_t + \frac{1}{2}\boldsymbol{\lambda}_t'\boldsymbol{\alpha}_2(t;\tau)\boldsymbol{\lambda}_t}$$

1. Take the derivatives f_t, $\mathbf{f}_\lambda$ and $\mathbf{f}_{\lambda\lambda}$
2. Substitute and pool terms together

- to obtain

$$\begin{aligned} 0 &= ((1-\gamma) r - \rho) \frac{1}{\gamma} + \frac{\partial \alpha_0(t;\tau)}{\partial t} + \boldsymbol{\alpha}_1(t,\tau)' \mathbf{A}_0 + \frac{1}{2} tr\left(\boldsymbol{\alpha}_2(t,\tau) \boldsymbol{\Sigma}\boldsymbol{\Sigma}'\right) + \frac{1}{2} tr\left(\boldsymbol{\alpha}_1(t,\tau)\boldsymbol{\alpha}_1(t,\tau)' \boldsymbol{\Sigma}\boldsymbol{\Sigma}'\right) \\ &+ \left(\frac{\partial \boldsymbol{\alpha}_1(t,\tau)'}{\partial t} + (1-\gamma)\boldsymbol{\alpha}_1(t,\tau)' \boldsymbol{\Sigma}\boldsymbol{\sigma}' \frac{1}{\gamma}(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1} + \boldsymbol{\alpha}_1(t,\tau)'\mathbf{A}_1 + \mathbf{A}_0'\boldsymbol{\alpha}_2(t,\tau) + \boldsymbol{\alpha}_1(t,\tau)\boldsymbol{\Sigma}\boldsymbol{\Sigma}'\boldsymbol{\alpha}_2(t,\tau)\right)\boldsymbol{\lambda}_t \\ &+ tr\left(\left(\frac{1}{2}\frac{\partial \boldsymbol{\alpha}_2(t,\tau)}{\partial t} + \frac{1}{2}(1-\gamma)\frac{1}{\gamma^2}(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1} + (1-\gamma)\boldsymbol{\alpha}_2(t,\tau)'\boldsymbol{\Sigma}\boldsymbol{\sigma}'\frac{1}{\gamma}(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1} + \boldsymbol{\alpha}_2(t,\tau)'\mathbf{A}_1 + \frac{1}{2}\boldsymbol{\alpha}_2(t,\tau)'\boldsymbol{\Sigma}\boldsymbol{\Sigma}'\boldsymbol{\alpha}_2(t,\tau)\right)\boldsymbol{\lambda}_t\boldsymbol{\lambda}_t'\right) \end{aligned}$$

- In order for the right hand side to be zero independently of $\boldsymbol{\lambda}_t$, the following must hold.

An Analytical Solution

- A system of ODE:

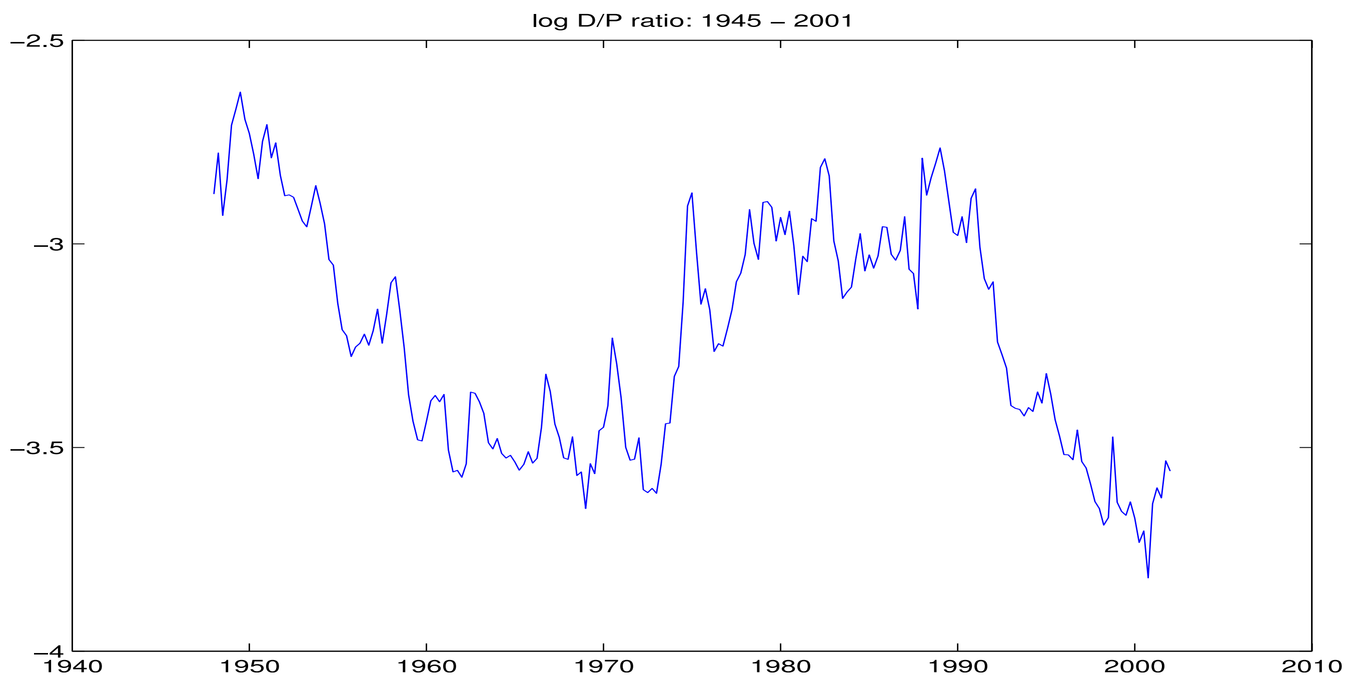
$$
\begin{aligned}
0 &= \frac{\partial \boldsymbol{\alpha}_2(t,\tau)}{\partial t} + (1-\gamma)\frac{1}{\gamma}\left(\frac{1}{\gamma} + 2\boldsymbol{\alpha}_2(t,\tau)'\boldsymbol{\Sigma}\boldsymbol{\sigma}'\right)(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1} + 2\boldsymbol{\alpha}_2(t,\tau)'\mathbf{A}_1 + \boldsymbol{\alpha}_2(t,\tau)'\boldsymbol{\Sigma}\boldsymbol{\Sigma}'\boldsymbol{\alpha}_2(t,\tau) \\
0 &= \frac{\partial \boldsymbol{\alpha}_1(t,\tau)'}{\partial t} + (1-\gamma)\boldsymbol{\alpha}_1(t,\tau)'\boldsymbol{\Sigma}\boldsymbol{\sigma}'\frac{1}{\gamma}(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1} + \boldsymbol{\alpha}_1(t,\tau)'\mathbf{A}_1 + \mathbf{A}_0'\boldsymbol{\alpha}_2(t,\tau) + \boldsymbol{\alpha}_1(t,\tau)\boldsymbol{\Sigma}\boldsymbol{\Sigma}'\boldsymbol{\alpha}_2(t,\tau) \\
0 &= \frac{\partial \alpha_0(t;\tau)}{\partial t} + ((1-\gamma)r-\rho)\frac{1}{\gamma} + \boldsymbol{\alpha}_1(t,\tau)'\mathbf{A}_0 + \frac{1}{2}tr(\boldsymbol{\alpha}_2(t,\tau)\boldsymbol{\Sigma}\boldsymbol{\Sigma}') + \frac{1}{2}tr\left(\boldsymbol{\alpha}_1(t,\tau)\boldsymbol{\alpha}_1(t,\tau)'\boldsymbol{\Sigma}\boldsymbol{\Sigma}'\right)
\end{aligned}
$$

- with final conditions $\boldsymbol{\alpha}_i(\tau,\tau) = 0$, $i = 0, 1, 2$.
- These ODEs can be easily solved numerically, independently of the dimension.
  - Just start with the final condition at τ and move backwards over time (it is three lines of code: one for each ODE).

Application 1: Portfolio Allocation under Predictability

- Let $n = 1$ and dR_t be the return on the aggregate stock market.
- Much of the literature uses the log dividend price ratio as a predictor.
- Let $x_t = \log(D_t/P_t)$ and let it follow the mean reverting process

$$dx_t = (\eta - \phi x_t)\, dt + \Sigma_{x1} dB_t^1$$



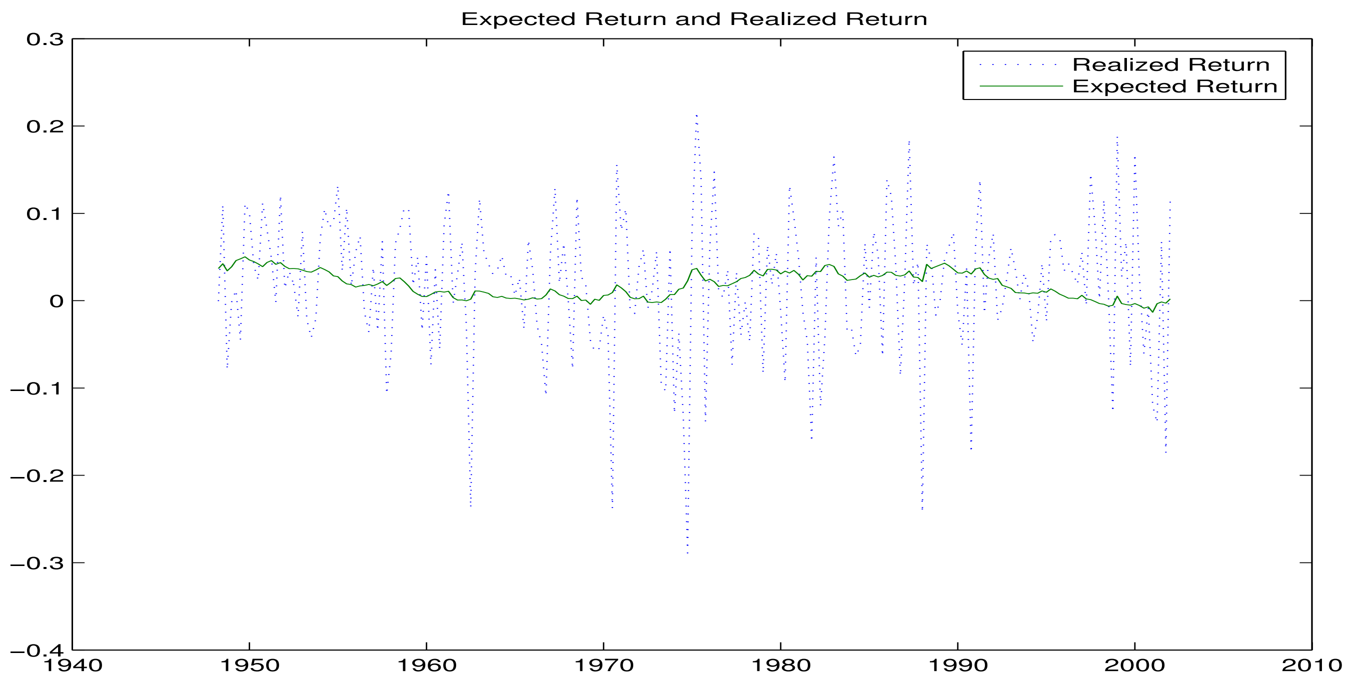
Application 1: Portfolio Allocation under Predictability

- Using x_t a predictor of excess stock returns, we can estimate

$$R_{t,t+dt} = \tilde{\beta}_0 + \tilde{\beta}_1 x_t + \epsilon_{t+dt}$$

Sample: 1947 - 2001. $dt = .25$

$\tilde{\beta}_0$	(t-stat)	$\tilde{\beta}_1$	(t-stat)	R^2
0.1898	(3.7042)	0.0531	(3.2424)	3.53%



Application 1: Portfolio Allocation under Predictability

- The annualized expected return $\lambda_t = E_t[R_{t,t+dt}/dt]$ is given by

$$\lambda_t \ = \ \beta_0 + \beta_1 x_t$$

- with $\beta_i = \tilde{\beta}_i/dt$
- Ito's Lemma implies

$$d\lambda_t = (A_0 + A_1\lambda_t)\, dt + \Sigma_1 dB_t^1$$

with

$$A_0 \ = \ \beta_1\eta + \phi\beta_0; A_1 = -\phi; \Sigma_1 = \beta_1\Sigma_{x1}$$

- The process for stock returns is

$$dR_t = (r + \lambda_t)\, dt + \sigma_1 dB_t^1 + \sigma_2 dB_t^2$$

Application 1: Portfolio Allocation under Predictability

Model:

$$\begin{aligned} d\lambda_t &= (A_0 + A_1\lambda_t)\, dt + \Sigma_1 dB_t^1 \\ dR_t &= (r + \lambda_t)\, dt + \sigma_1 dB_t^1 + \sigma_2 dB_t^2 \end{aligned}$$

Sample: 1947 - 2001. $dt = .25$

A_0	A_1	Σ_1	σ_1	σ_2
0.0077	-0.1405	-0.0317	0.1183	0.1057

- **Note 1**: Negative Σ_1 simply means $Cov(dR, d\lambda) = \Sigma_1\sigma_1 = -.0038 < 0$
  - Positive shocks to dividend yield increase expected returns but are *contemporaneously* negatively correlated with returns.
    - This is intuitive: dividend yield moves mainly because of prices.
    - If $P_t \downarrow \Longrightarrow dR_t < 0$ and $log(D/P) \uparrow$

A bad news ($dR < 0$) is not very bad, as it increases expected returns

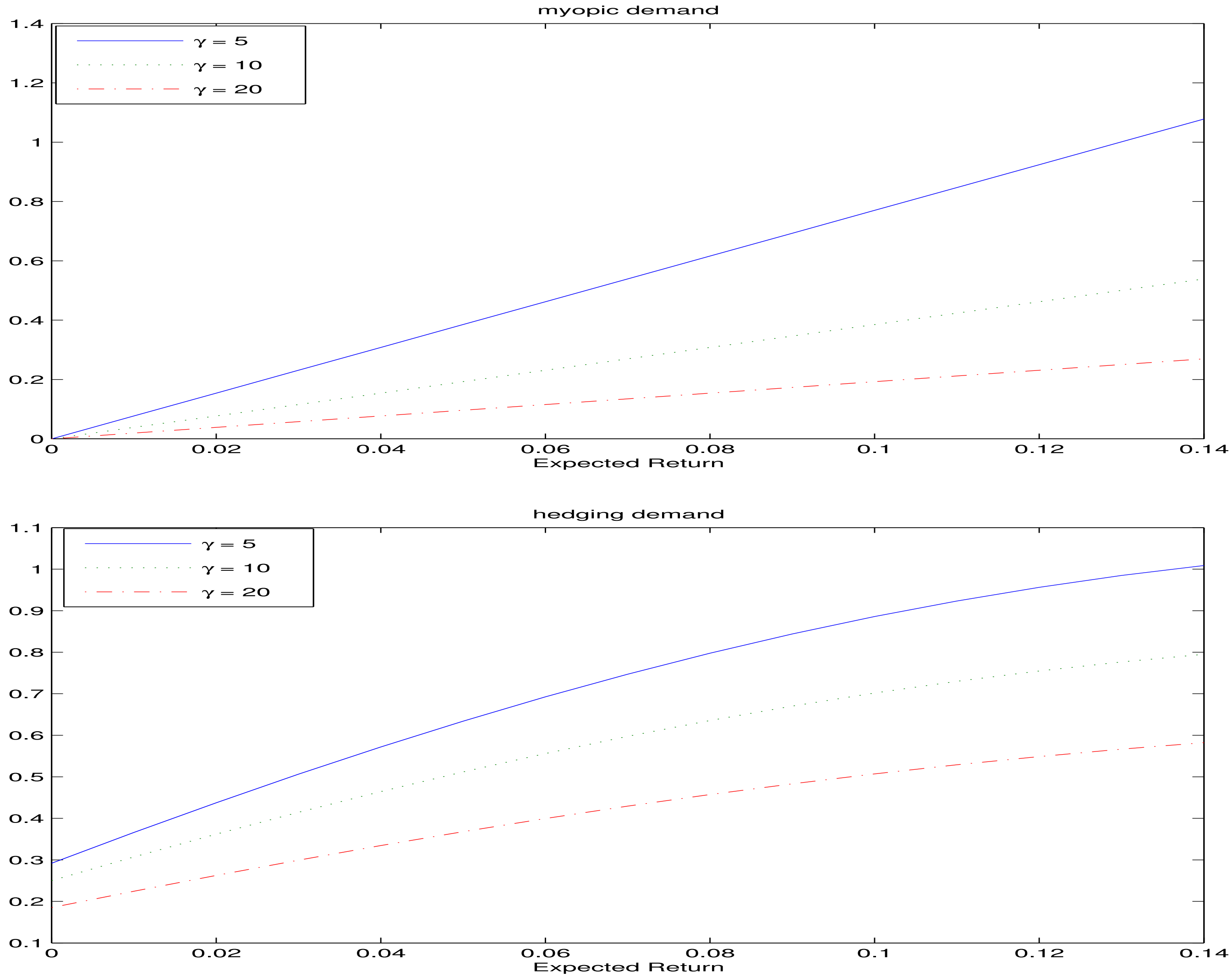
Application 1: Portfolio Allocation under Predictability

- **Note 2**: The condition for an exact analytical solution is violated:

$$\Sigma\sigma'\left(\sigma\sigma'\right)^{-1}\sigma\Sigma - \Sigma\Sigma' = 0 \Rightarrow \frac{\left(\sigma_1\Sigma_1\right)^2}{\sigma_1^2 + \sigma_2^2} = \Sigma_1^2 \Rightarrow \sigma_2^2 = 0$$

- $\Longrightarrow$ Exact formula really holds under the assumption of complete markets.
  - Stock returns span all of the uncertainty.
- Instead, we found $\sigma_2 > 0$.
  - Part of the problem is the use of quarterly data. At monthly frequency the (negative) correlation between returns and dividend yield is higher.
  - For the sake of argument, I will assume a perfect negative correlation between returns and dividend yield.
    - ∗ In what follows I then use $\sigma_1 = .1612$ and $\sigma_2 = 0$.

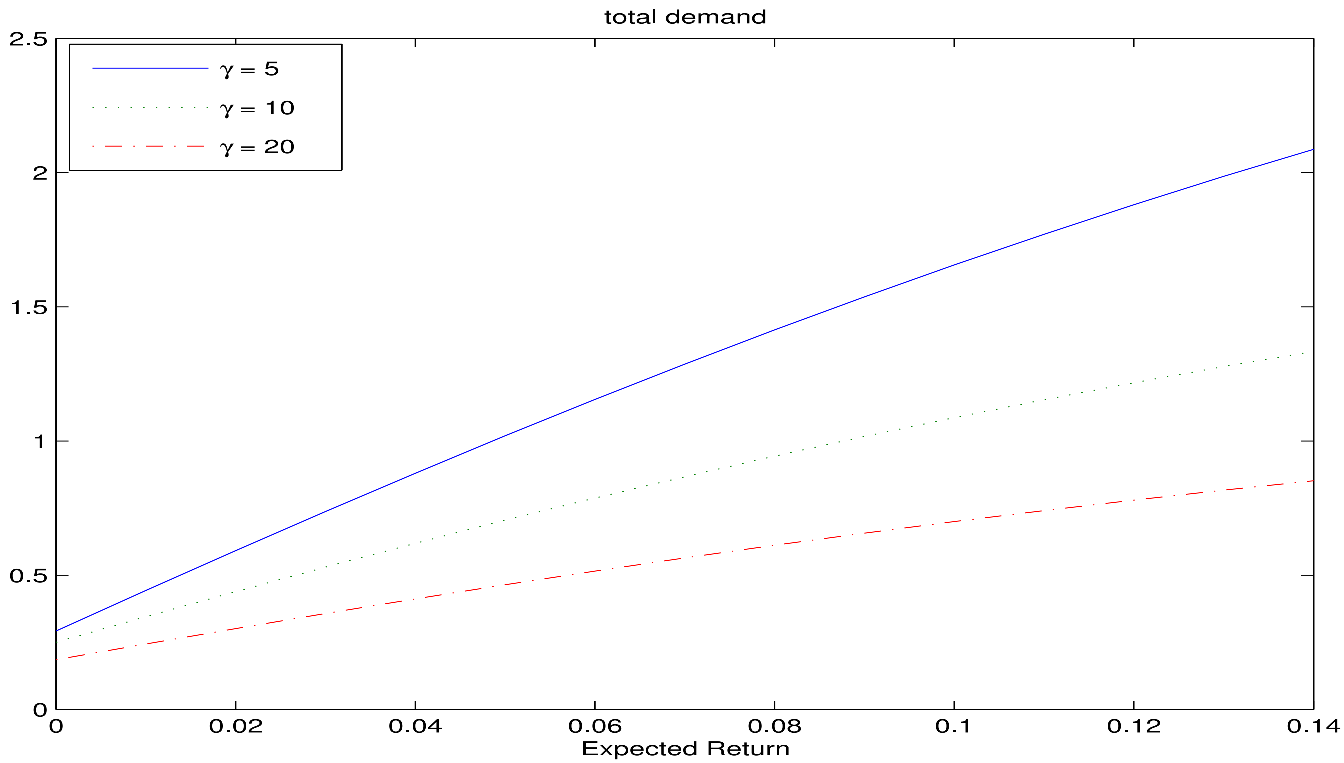
Myopic and Hedging Demand for Various Risk Aversion Parameter



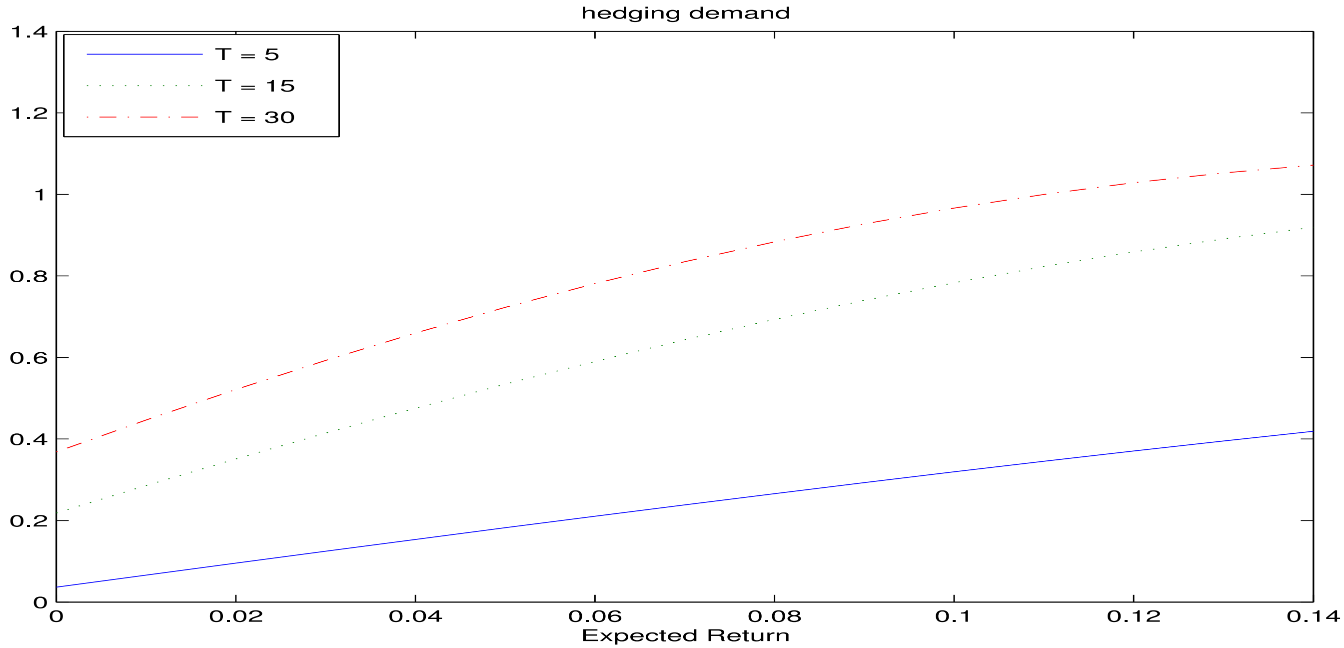
Hedging Demand with Predictable Returns

- **Finding 1:** The hedging demand is positive.
- The intuition is simple:
  - If we have a bad shock to returns, we have that μ_t increases (intuitively, the D/P increases, implying higher expected return).
  - But a higher λ_t implies that investor now want to buy more of the stock.
  - Anticipating this correlation, the investor buys more of the stock today, compared to the case where the hedging demand is zero.
- This finding is bad news for the portfolio holding puzzle:
  - We already showed that the agent would hold too much of the stock even with simple myopic demand (no time varying investment opportunity set).
    - ∗ The total demand now of the stock is even higher, deepening the puzzle.

Total Demand with Predictable Returns

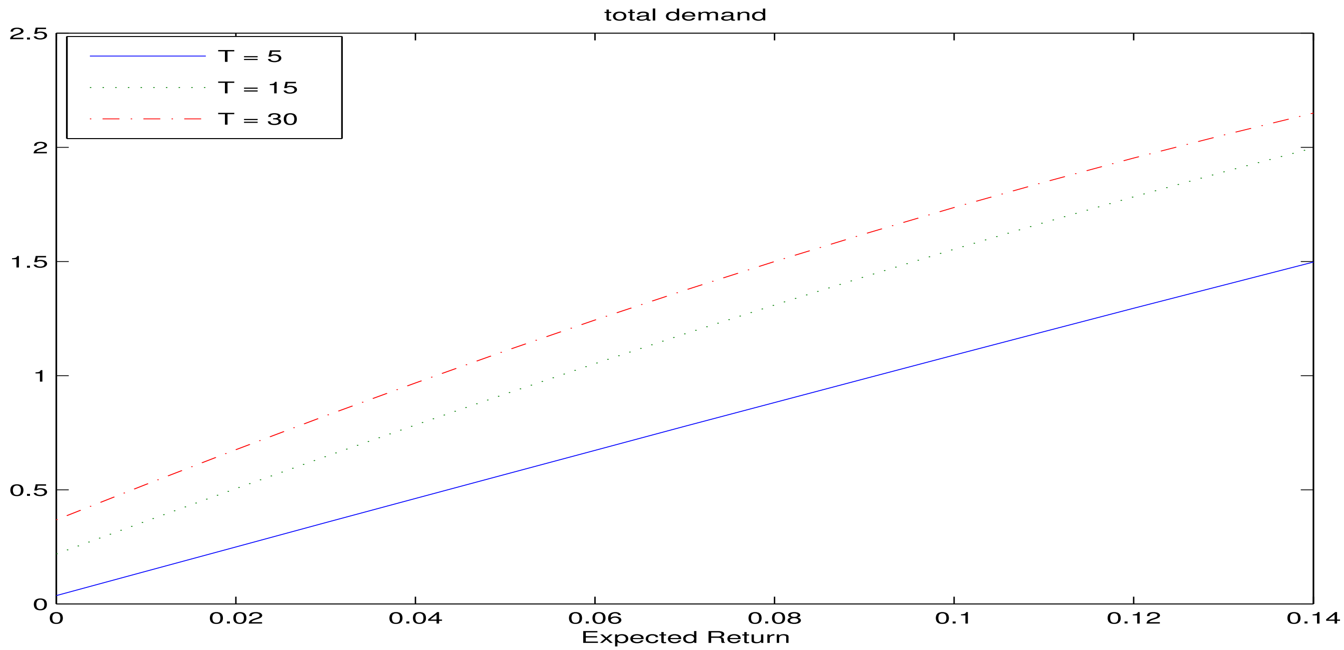


Hedging Demand for Various Life Expectancy



- **Finding 2:** Hedging demands help to address the life - cycle allocation puzzle.
  - As it can be see, the shorter the life expectancy T the lower the share in stocks, especially if current expected return is high.
  - In this case, mean reversion kicks in and the investor is wary about the negative consequences of a decrease in expected returns.

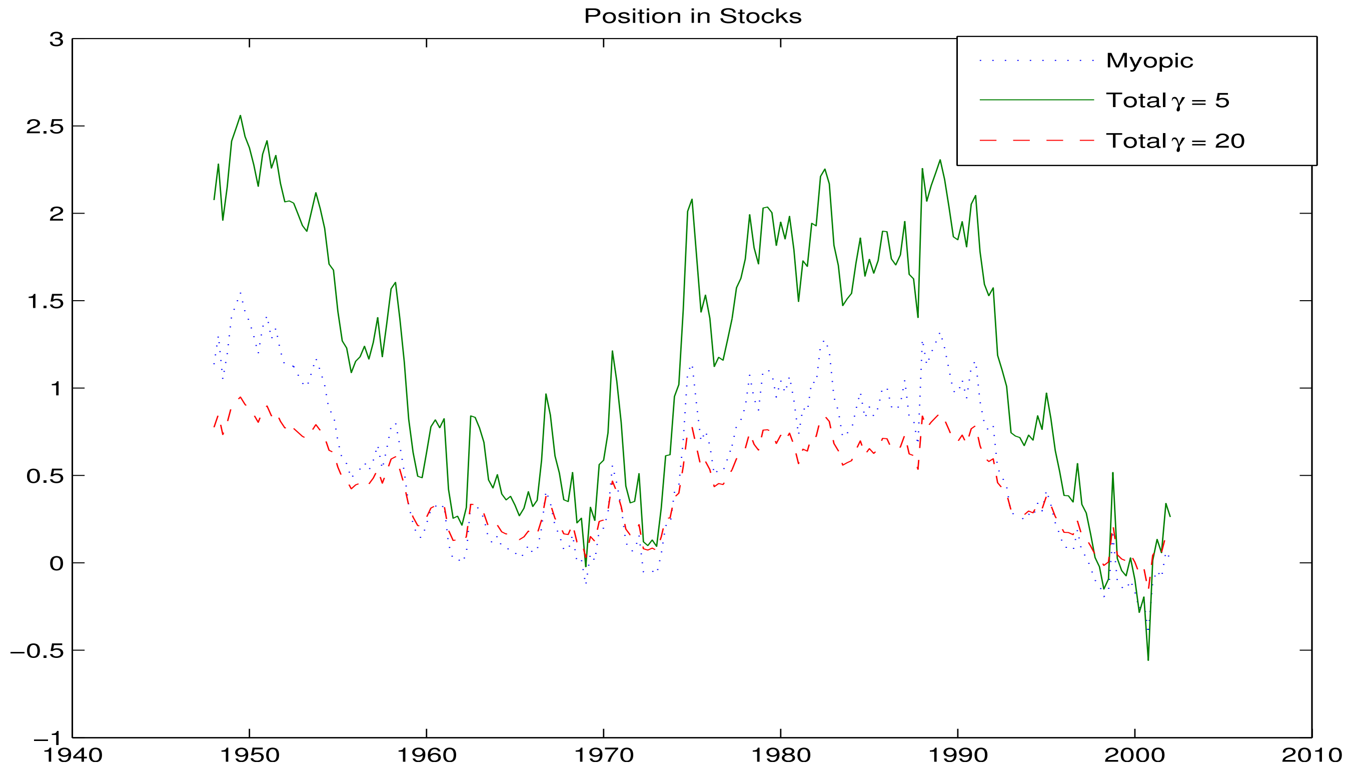
Total Demand over the Life Cycle



- Still, because of the hedging demand, an investor with 5 years to live would be still substantially exposed to stocks.

Strategic Asset Allocation over Time

- What is the variation over time of the optimal allocation to stock?
- Consider investor with $T = 15$ (constant) and $\gamma = 1, 5, 20$.



- The pattern for $\gamma = 20$ seems more reasonable than $\gamma = 1$ or 5.

Strategic Asset Allocation: Discussion

1. The predictability of stock returns is still source of heated debate.
   - Here we take the strong view that investors take empirical estimates as "true" parameters.
   - Much recent literature tried to relax this assumption, and use Bayesian methods in portfolio allocation
     - Kandel and Stambaugh (JF, 1997), Barberis (JF, 2000), Pastor (JF, 2000), Xia (JF, 2001).
     - These methodologies are very numerically intensive.

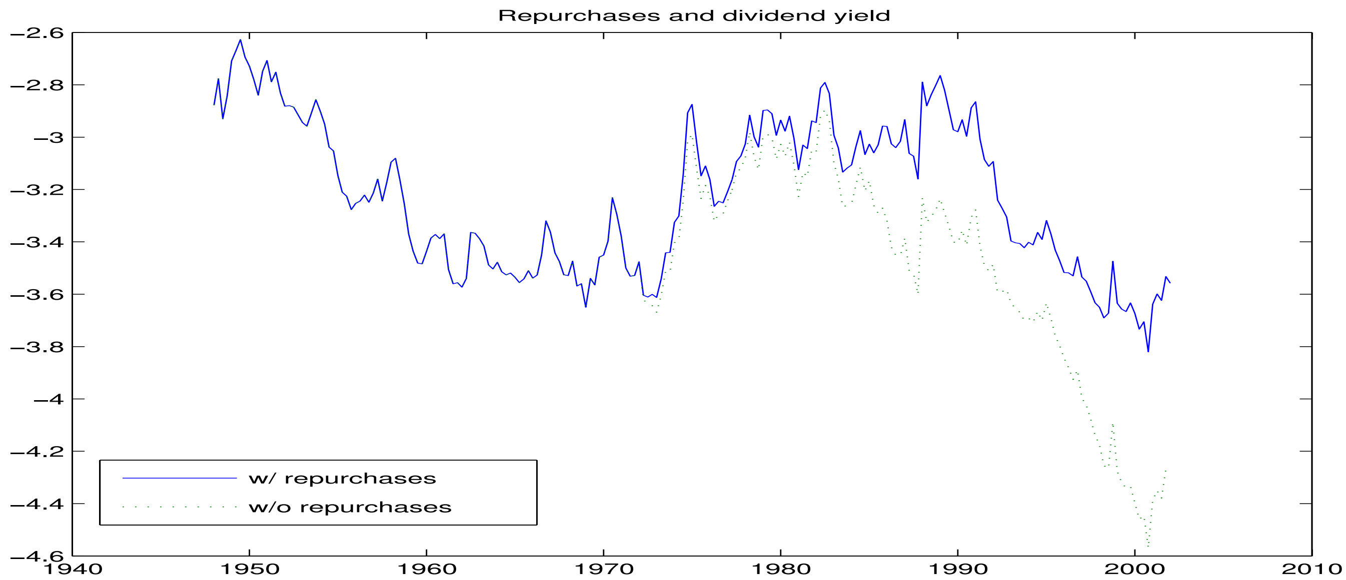
Strategic Asset Allocation: Discussion

2. As shown in Menzly, Santos and Veronesi (JPE, 2004), the dividend yield in which dividends are corrected for stock repurchases is a superior forecaster of future returns that the traditional dividend yield.

   – Without repurchases we have

Sample: 1947 - 2001. $dt = .25$

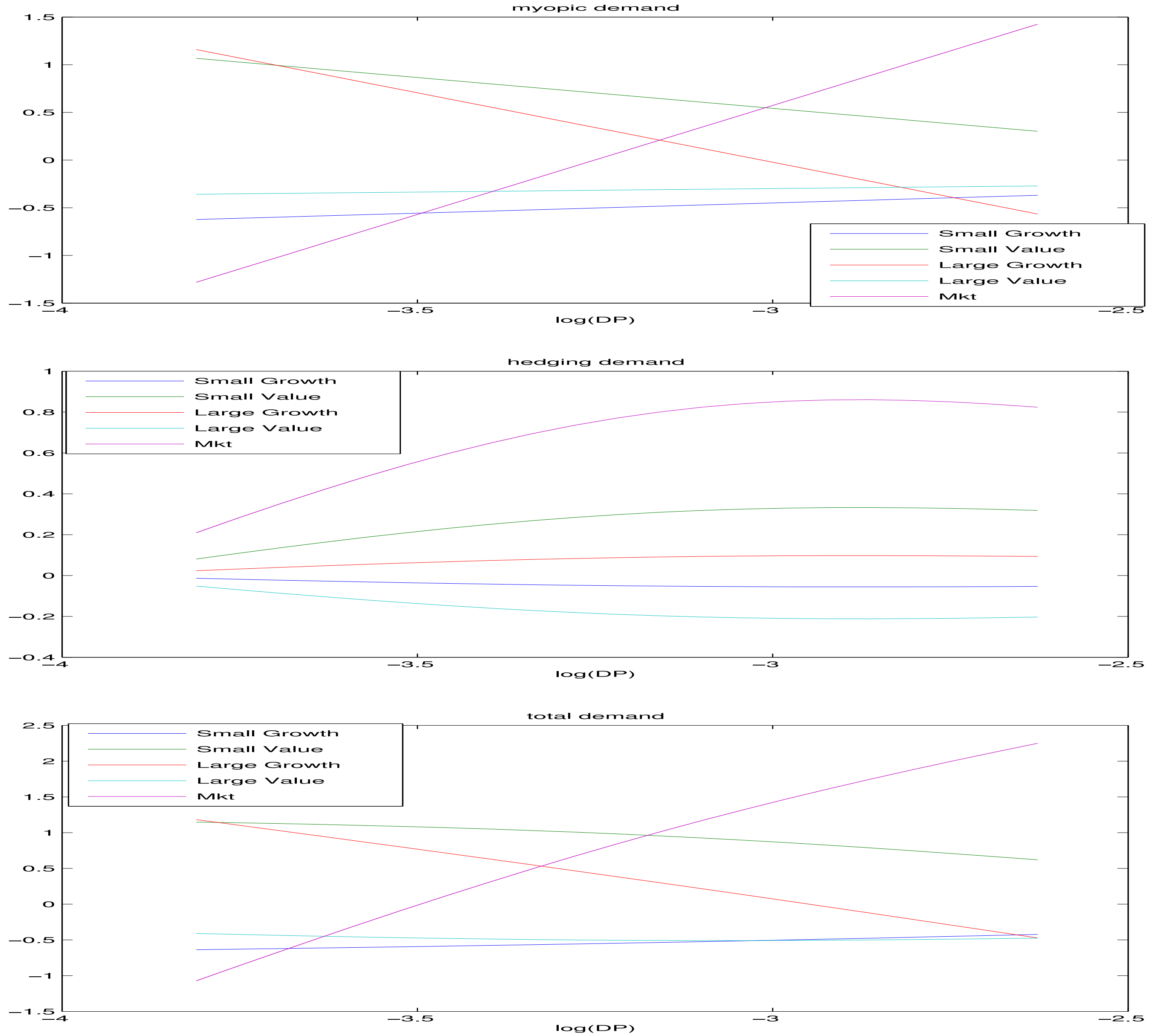
$\tilde{\beta}_0$	t-stat	$\tilde{\beta}_1$	t-stat	R^2
0.1233	2.5376	0.0310	2.0815	2.24%



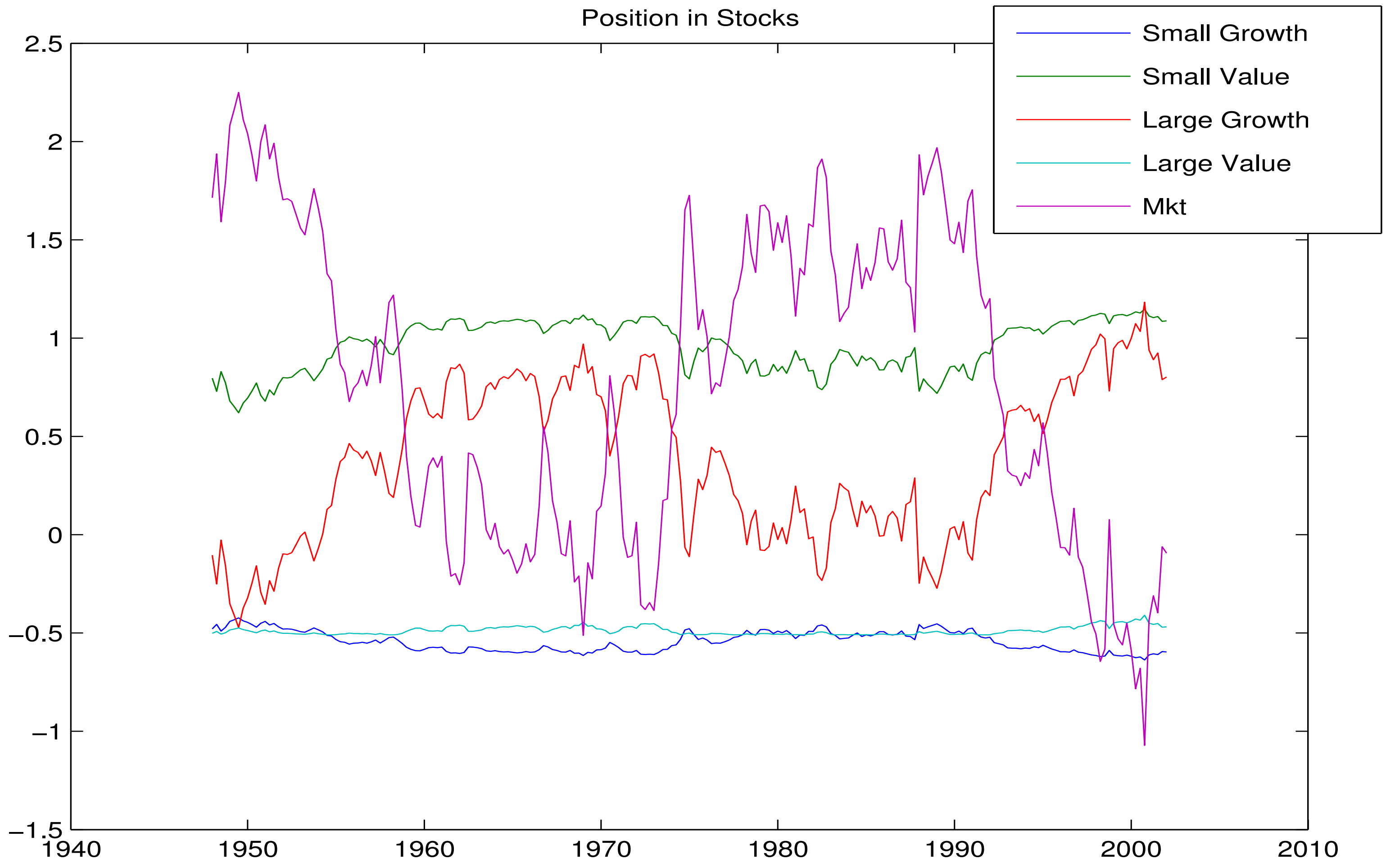
Strategic Asset Allocation: Discussion

3. The setting above can be easily extented to multiple assets and multiple predictors.
   - Analytical solutions are quite useful in this case.
   - Most models of strategic asset allocation do not go over the two or three assets.
   - As an illustration, next pictures show the strategic asset allocation for an investor who in addition to a market index, he has access to the returns from mutual funds specialized in 4 strategies:
     1. Value / Small Cap
     2. Value / Large Cap
     3. Growth / Small Cap
     4. Growth / Large Cap

Allocation to 6 Size - BM sorted portfolios and market



Allocation to 6 Size - BM sorted portfolios and market



Application 2: Learning about Average Returns

- Consider the same setting as in the original Merton problem

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

- Differently from Merton, assume that average returns $\boldsymbol{\mu}$ are not observable.
- Investors observe *realized* returns $d\mathbf{R}_t$ and infer the value of $\boldsymbol{\mu}$.
- Since the risk free rate r is observable, we can equivalently assume that agents infer the value of the average excess return $\boldsymbol{\lambda} = \boldsymbol{\mu} - r\mathbf{1}_n$.
- The following filtering result holds.

A Filtering Result

- Result: Let investors prior distribution at time 0 on $\boldsymbol{\lambda}$ be given by

$$\boldsymbol{\lambda}|_{t_0} \sim N\left(\widehat{\boldsymbol{\lambda}}_0, \widehat{\mathbf{q}}_0\right)$$

- Then, the posterior distribution at any time t is given by

$$\boldsymbol{\lambda}|_{t} \sim N\left(\widehat{\boldsymbol{\lambda}}_t, \widehat{\mathbf{q}}_t\right)$$

- where

$$\begin{aligned} d\widehat{\boldsymbol{\lambda}}_t &= \widehat{\boldsymbol{\Sigma}}_t d\widehat{\mathbf{B}}_t \\ \widehat{\boldsymbol{\Sigma}}_t &= \widehat{\mathbf{q}}_t \left(\boldsymbol{\sigma}'\right)^{-1} \\ \frac{d\widehat{\mathbf{q}}_t}{dt} &= -\widehat{\mathbf{q}}_t \left(\boldsymbol{\sigma}\boldsymbol{\sigma}'\right)^{-1} \widehat{\mathbf{q}}_t \end{aligned}$$

- The innovation process is

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

An Informational Equivalent Setting

- We can rewrite the system of returns then as follows

$$\begin{aligned} d\mathbf{R}_t &= \left(r + \widehat{\boldsymbol{\lambda}}\right) dt + \boldsymbol{\sigma} d\widehat{\mathbf{B}}_t \\ d\widehat{\boldsymbol{\lambda}}_t &= \widehat{\boldsymbol{\Sigma}}_t d\widehat{\mathbf{B}}_t \end{aligned}$$

- This is very similar to the previous case. Note the following:
  1. We are back to complete markets: Conditional on investors' information, the set of BMs that drive returns $d\mathbf{R}_t$ is the same that drive expected return $\widehat{\boldsymbol{\lambda}}_t$.
     - The reason is that the information filtration 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.
     - $\Longrightarrow$ The asset allocation solution is exact!

An Informational Equivalent Setting

2. The only difference from the problem discussed earlier is the fact that the volatility of $\widehat{\boldsymbol{\lambda}}_t$ depends on t.
   - However, this volatility declines deterministically.
   - Thus, the methodology developed earlier applies 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 zero as $t \to \infty$
   - This is because we assume $\boldsymbol{\lambda}$ is constant forever. Assuming some time variation in underlying average return will prevent the posterior variance from converging.
   - E.g. for the case $n = 1$,

$$q_t = \frac{1}{q_0^{-1} + \sigma^{-2} t}$$

An Informational Equivalent Setting

4. **Learning has a bite**: It has a prediction about the correlation between returns and expected returns.

$$Cov_t\left(d\mathbf{R}_t, d\boldsymbol{\lambda}_t\right) = \boldsymbol{\sigma}\widehat{\boldsymbol{\Sigma}}_t' = \boldsymbol{\sigma}\left(\boldsymbol{\sigma}\right)^{-1}\widehat{\mathbf{q}}_t = \widehat{\mathbf{q}}_t$$

– They are *p*ositively correlated: A negative innovation in returns decreases 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.

– This is opposite of what we found in our earlier exercise, 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 negatively correlated.

An Equivalent Portfolio Problem

- **Investor problem**:

$$J\left(W_0, \widehat{\lambda}_0, 0\right) = \max_{\{(C_t),(\boldsymbol{\theta}_t)\}} E_0\left[\int_0^T u\left(C_t, t\right) dt\right]$$

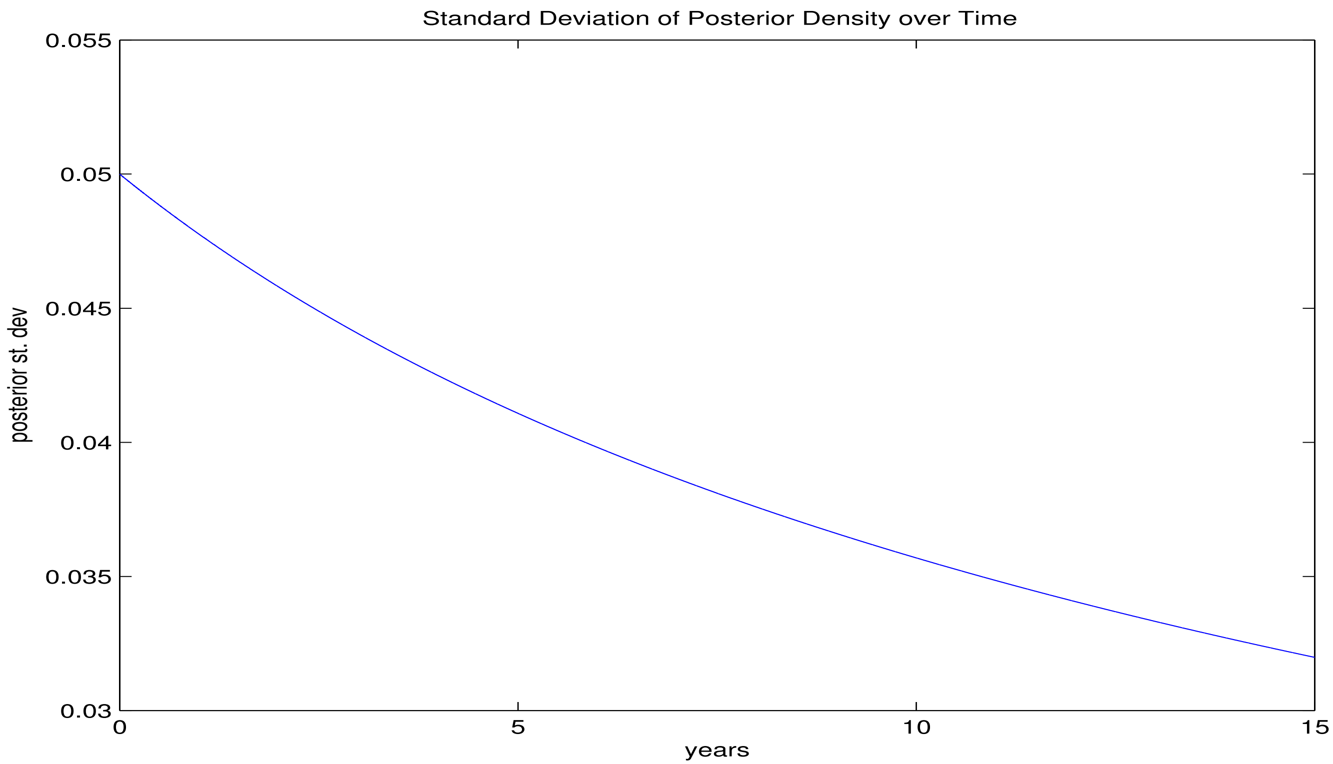
- subject to

$$dW_t = \left\{W_t\left(\boldsymbol{\theta}_t'\widehat{\boldsymbol{\lambda}}_t + r\right) - C_t\right\} dt + W_t\boldsymbol{\theta}_t'\boldsymbol{\sigma} d\widehat{\mathbf{B}}_t$$

- At this point, the solution is “almost” the same as before.
  - We need to set $\mathbf{A}_0 = \mathbf{A}_1 = 0$
  - Remember that $\widehat{\boldsymbol{\Sigma}}_t$ depends on time t.
    - ∗ The computation is in fact straightforward, as we can simply iterate forward the ODE that defines $\widehat{\mathbf{q}}_t$ (Riccati equation)

How Fast Would an Investor Learn?

- First, how fast does "uncertainty" declines?
  - From a prior uncertainty $\sqrt{q_0} = 5\%$, it declines rather slowly.

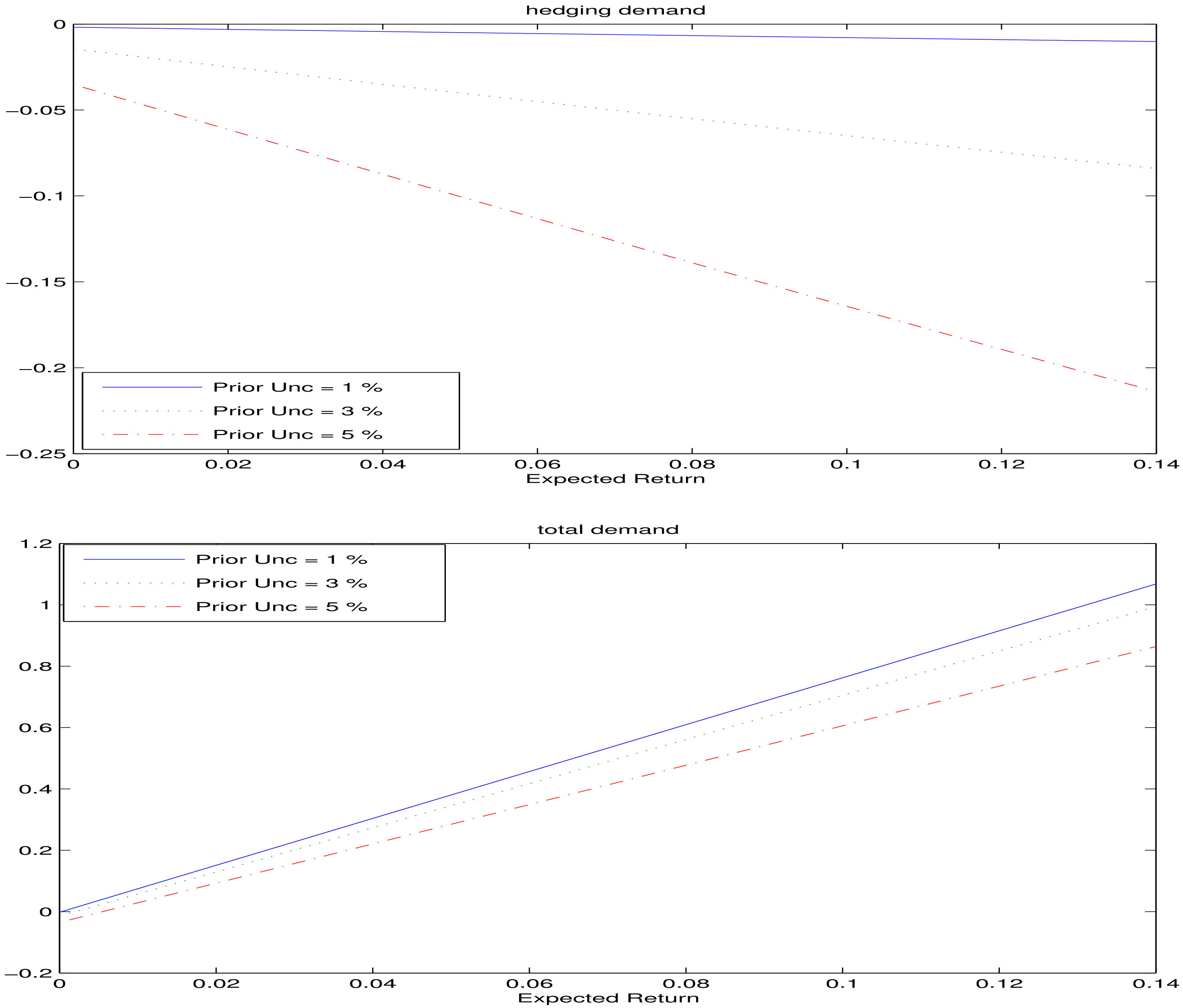


Strategic Asset Allocation with Learning: The Role of Prior Uncertainty

- 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.
- The following figures show the hedging demand and total demand for three different value of initial uncertainty

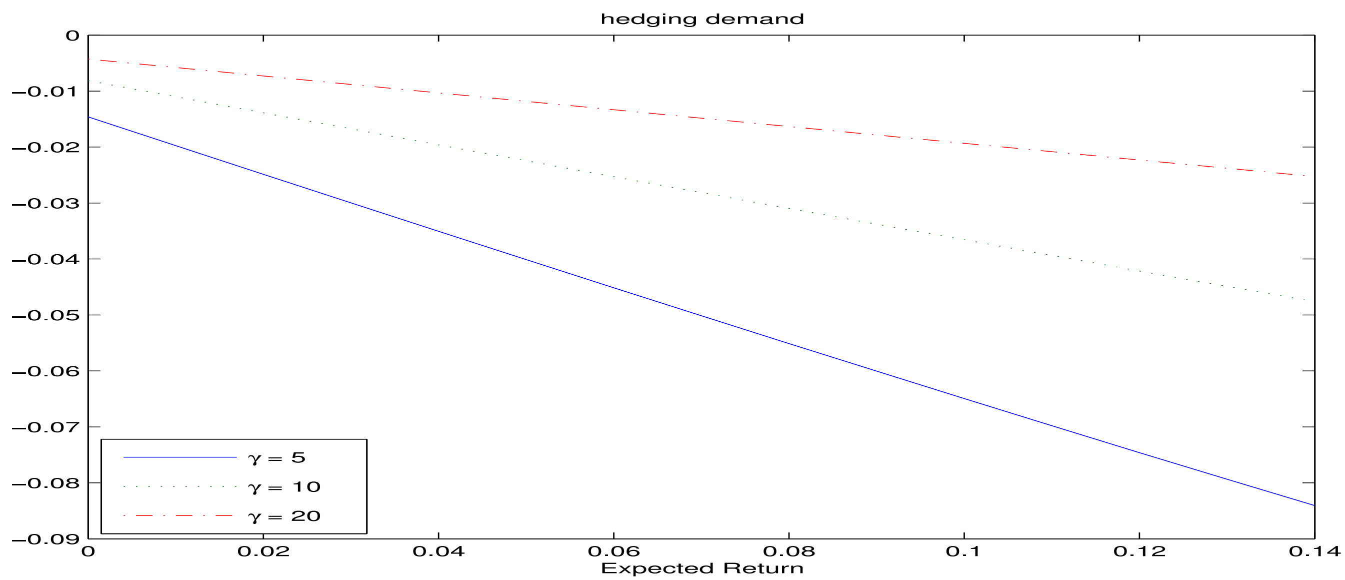
$$\sqrt{\widehat{q}_0} = 1\%, 3\%, 5\%$$

Strategic Asset Allocation with Learning: The Role of Prior Uncertainty



Strategic Asset Allocation with Learning: The Role of Risk Aversion

- What effect does risk aversion have on hedging demands?



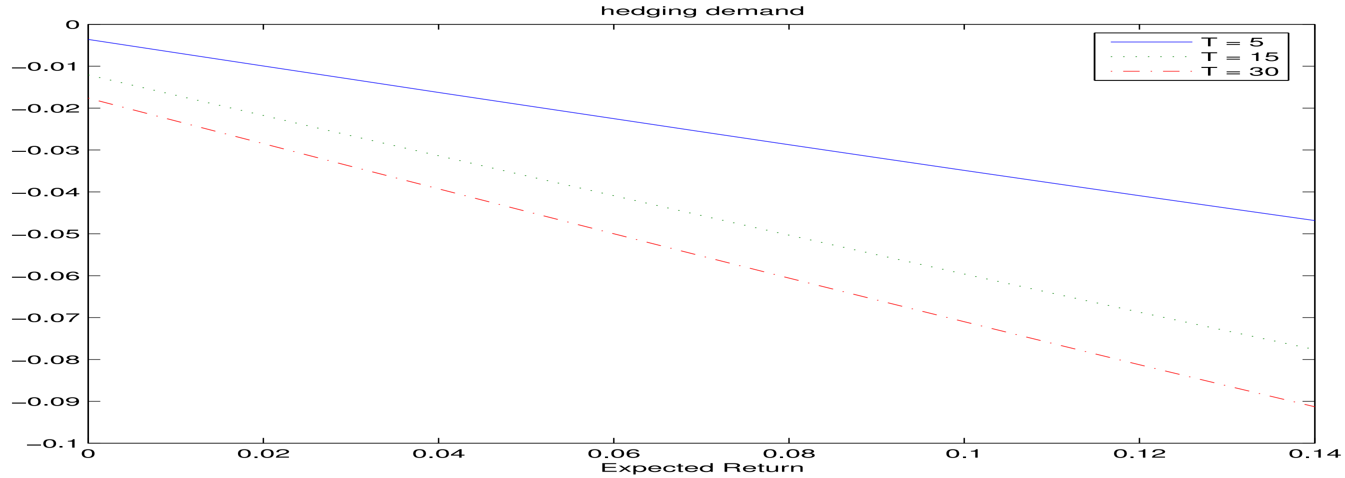
- Higher risk aversion decreases (in absolute value) the hedging demand.

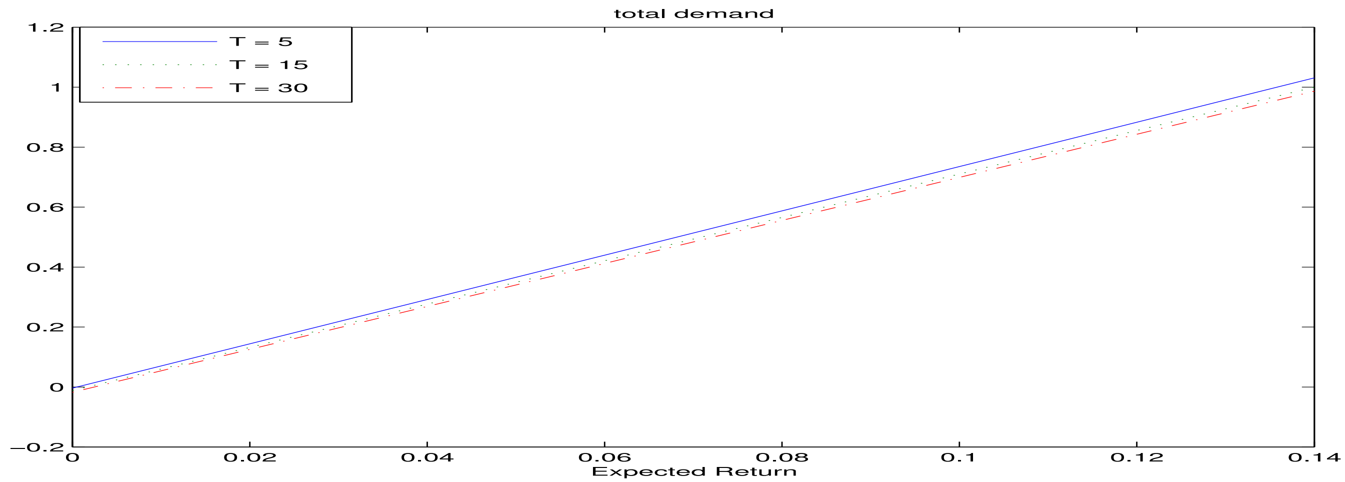
Strategic Asset Allocation with Learning: The Role of Risk Aversion

- Why does higher risk aversion decreases (in absolute value) the hedging demand?
  - This is due to the sensitivity of the consumption / wealth ratio C/W to changes in expected returns.
  - As we increase γ, the myopic demand for stocks decreases.
    * $\Longrightarrow$ The consumption to wealth ratio C/W becomes more and more insensitive to variation expected return.
    * $\Longrightarrow$ Eventually, changes in expected return have no impact on C/W, and thus no need of hedging demand.
    * $\Longrightarrow$ The relation between γ and hedging demand is non-linear, as hedging demand are close to zero both for γ close to 1 and for γ large.

Strategic Asset Allocation with Learning: The Life Cycle Implications

- How does learning affect the allocation of investors with different life expectancies?





Strategic Asset Allocation with Learning: The Life Cycle Implications

- Learning does not seem to have a large impact on the asset allocation as a function of time T.
- The little that is has goes in the opposite direction:
  - The reason, again, is the EIS.
  - The longer the horizon, the higher the impact of an increase in expected return on future consumption.
  - $\Longrightarrow$ larger decrease in θ_t due to consumption smoothing.

Strategic Asset Allocation with Learning over time

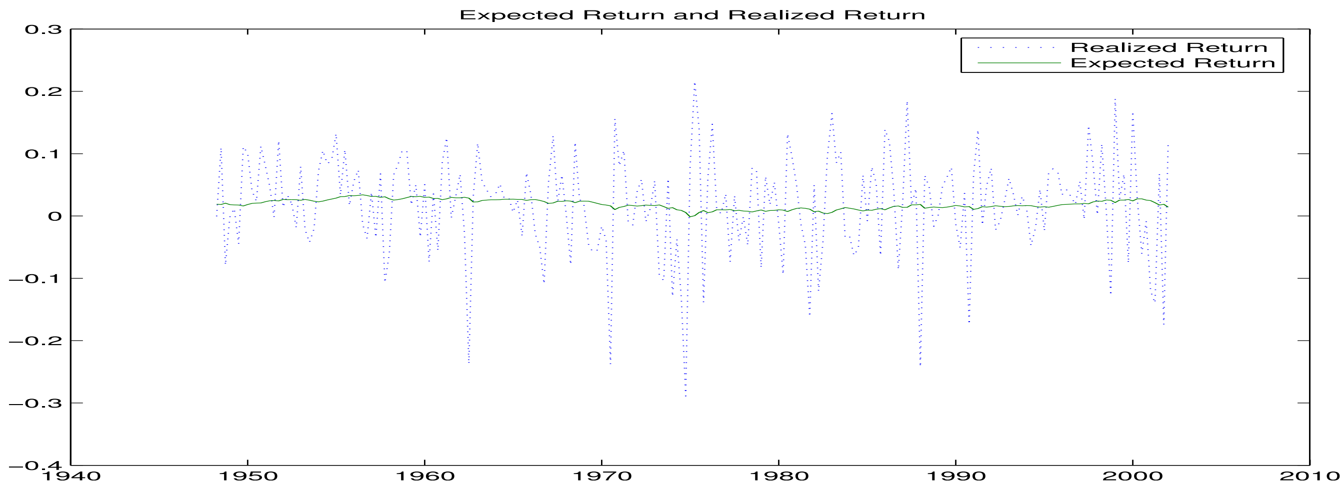
- Consider an investor in 1947 with prior uncertainty $\sqrt{q_0} = 5\%$.
  - How would his asset allocation change over time?

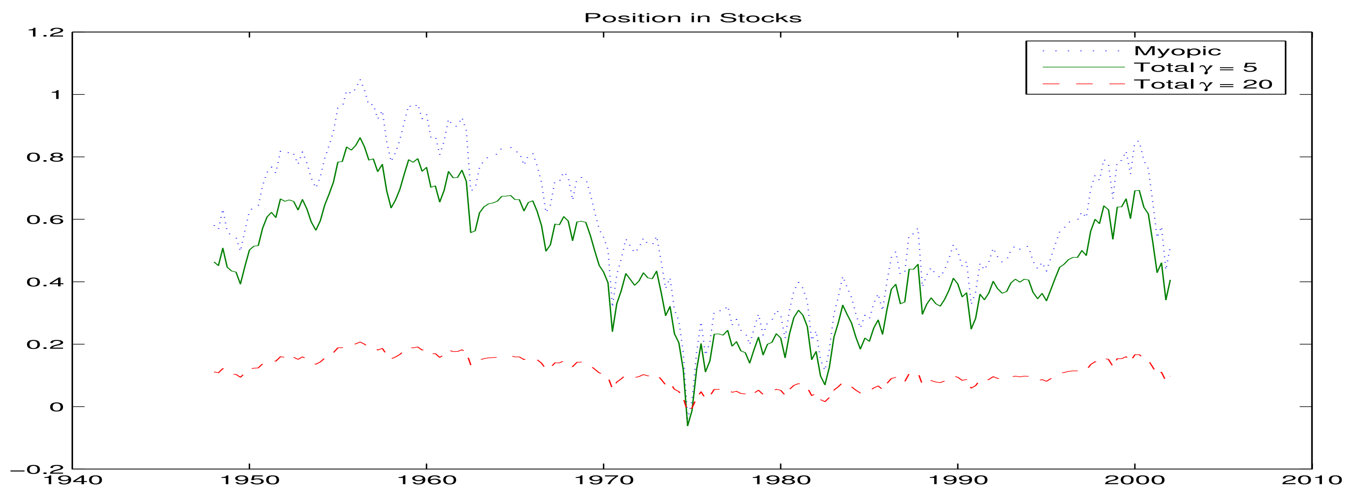
Strategic Asset Allocation with Learning over time

- Case 1: Assume a declining uncertainty over time

Strategic Asset Allocation with Learning over time

- Case 2: Assume a constant uncertainty (e.g. small probability of jumps)

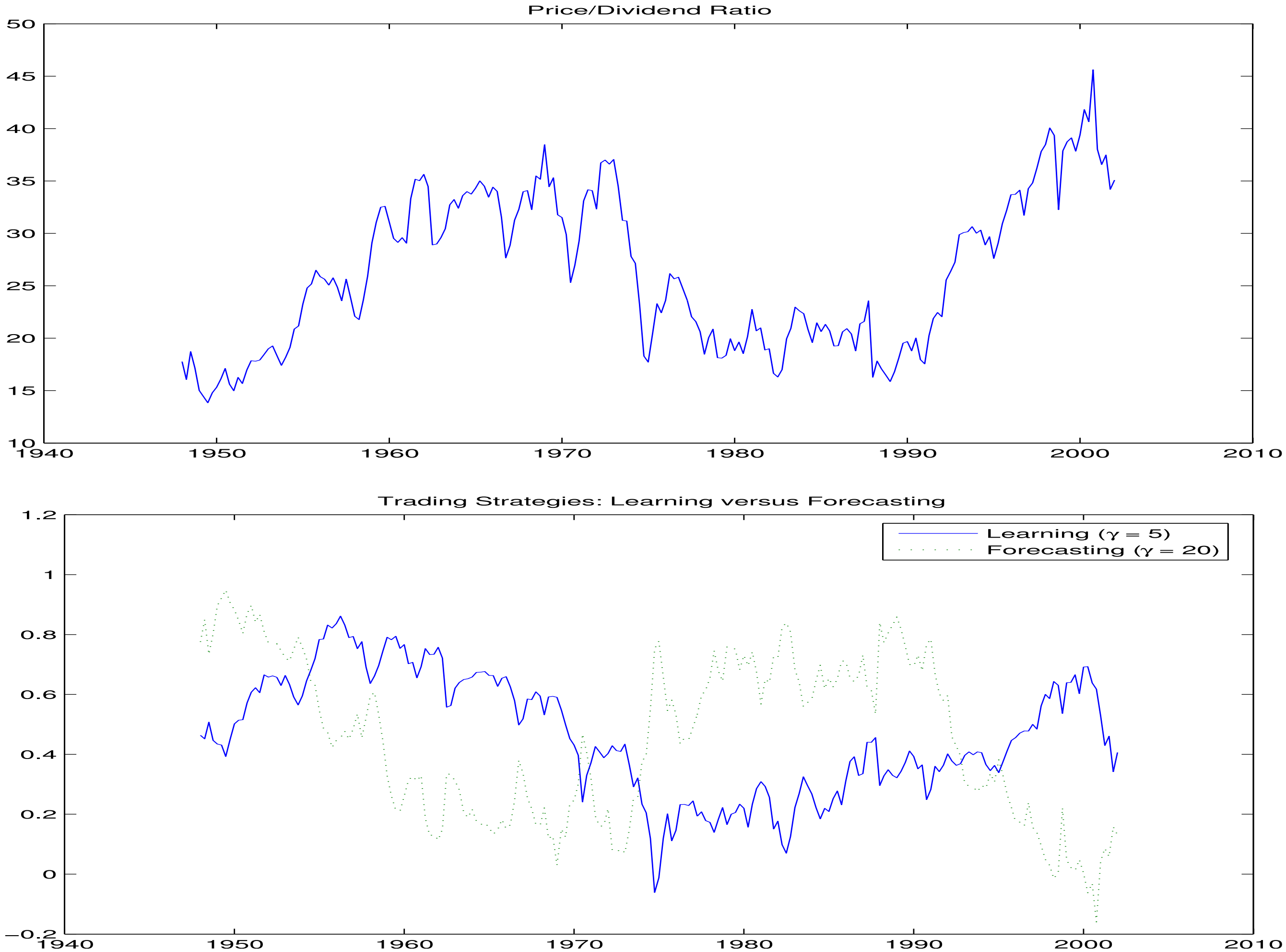




Strategic Asset Allocation and Expected Returns: Comparison

- Learning about average returns:
  - $\Longrightarrow$ Investor behave like "momentum" traders (or trend chasers)
    - They buy when prices increase.
- Forecasting returns using the dividend yield:
  - $\Longrightarrow$ Investors behave like reversal traders
    - They buy when prices drop

Strategic Asset Allocation and Expected Returns: Comparison



Strategic Asset Allocation with Learning *About* Predictability

- The predictive regression

$$r_{t+1} = \alpha + \beta \log\left(\frac{D_t}{P_t}\right) + \varepsilon_{t+1}$$

  is not reliable, as the β has large standard errors, and it changes value.

- How much would an investor invest in the stock market if needs to "estimate" β as well?
  - Kandel and Stambaugh (1997, JF) examine one period problem using Bayesian methodologies.
  - Barberis (JF, 2000) consider multiperiod model incorporating uncertainty about β. It does not allow for learning about β.
  - Xia (JF, 2001) uses a continuos time model to investigate the learning about β effect.
  - Brandt, Goyal, Santa Clara and Stroud (2005) use a simulation-based method to investigate a general model with learning.

Strategic Asset Allocation with Learning *About* Predictability

- Consider the single stock case, with

$$dR_t = (r + \lambda_t)\, dt + \sigma_R dB_t$$

  – where, denoting by x_t a predictor, we assume

$$\lambda_t = \overline{\lambda} + \beta_t(x_t - \overline{x})$$

- Let β_t and x_t follow the linear processes

$$\begin{aligned} d\beta_t &= \phi(\overline{\beta} - \beta_t)dt + \sigma_\beta dB_{\beta,t} \\ dx_t &= \kappa(\overline{x} - x_t)dt + \sigma_x dB_{x,t} \end{aligned}$$

- The investor observes returns dR and the predictor x. However, the investor does not observe β_t.
- In addition, the investor knows all of the parameters.

Another Filtering Result

- The observation equation is given by $d\mathbf{Y}_t = (dR_t, dx_t)'$

$$d\mathbf{Y}_t = (\mathbf{F}_0(\mathbf{Y}_t, t) + \mathbf{F}_1(\mathbf{Y}_t, t)\beta_t)dt + \boldsymbol{\omega}(\mathbf{Y}_t, t)d\mathbf{B}_t$$

- The state equation can be written as

$$d\beta_t = (a_0(\mathbf{Y}_t, t) + a_1(\mathbf{Y}_t, t)\beta_t)dt + \eta(\mathbf{Y}, t)d\mathbf{B}_t$$

- Denote by $\boldsymbol{\Gamma}(\mathbf{Y}_t, t) = \boldsymbol{\eta}\boldsymbol{\omega}'$, $\Sigma(\mathbf{Y}_t, t) = \boldsymbol{\eta}\boldsymbol{\eta}'$ and $\boldsymbol{\Phi}(\mathbf{Y}_t, t) = \boldsymbol{\omega}\boldsymbol{\omega}'$.
- Let the prior $\beta_0 \sim N(\widehat{\beta}_0, \widehat{v}_0)$. Then [Liptser and Shiryayev (1978, Thm 12.1)]:
- The posterior of $\beta_t \sim N(\widehat{\beta}_t, \widehat{v}_t)$ with

$$\begin{aligned}
d\widehat{\beta}_t &= (a_0(\mathbf{Y}_t, t) + a_1(\mathbf{Y}_t, t)\widehat{\beta}_t)dt + [v_t\mathbf{F}_1(\mathbf{Y}_t, t) + \boldsymbol{\Gamma}(\mathbf{Y}_t, t)]\,\Phi(\mathbf{Y}_t, t)^{-1}\boldsymbol{\omega}(\mathbf{Y}_t, t)d\widehat{\mathbf{B}}_t \\
\frac{d\widehat{v}_t}{dt} &= 2a_1(\mathbf{Y}_t, t)\widehat{v}_t + \Sigma(\mathbf{Y}_t, t) \\
&\qquad - [\widehat{v}_t\mathbf{F}_1(\mathbf{Y}_t, t) + \boldsymbol{\Gamma}(\mathbf{Y}_t, t)]\,\boldsymbol{\Phi}(\mathbf{Y}_t, t)^{-1}\,[\widehat{v}_t\mathbf{F}_1(\mathbf{Y}_t, t) + \boldsymbol{\Gamma}(\mathbf{Y}_t, t)]' \\
d\widehat{\mathbf{B}}_t &= \boldsymbol{\omega}^{-1}\,[d\mathbf{Y}_t - E\,(d\mathbf{Y}_t)]
\end{aligned}$$

Another Filtering Result - 2

- In our special case, we get, more explicitly:

$$\begin{aligned} d\widehat{\beta}_t &= \phi(\overline{\beta} - \widehat{\beta}_t)dt + v_1(x_t, v_t)d\widehat{B}_t + v_2(x_t, v_t)d\widehat{B}_{x,t} \\ \frac{d\widehat{v}_t}{dt} &= -2\phi\widehat{v}_t + \sigma_\beta^2 - \left[v_1(x_t, \widehat{v}_t)^2 + v_2(x, \widehat{v}_t)^2 + 2v_1(x_t, \widehat{v}_t)v_2(x_t, \widehat{v}_t)\rho_{R,s}\right] \end{aligned}$$

- with

$$\begin{aligned} v_1(x_t, v_t) &= \frac{v_t(x_t - \overline{x}) + \sigma_R\sigma_\beta(\rho_{\beta,R} - \rho_{\beta,x}\rho_{x,R})}{\sigma_R(1 - \rho_{x,R}^2)} \\ v_2(x_t, v_t) &= \frac{-v_t(x_t - \overline{x})\rho_{x,R} + \sigma_R\sigma_\beta(\rho_{\beta,x} - \rho_{\beta,R}\rho_{x,R})}{\sigma_R(1 - \rho_{x,R}^2)} \end{aligned}$$

- Note:
  - The dynamics of the posterior mean $\widehat{\beta}_t$ displays stochastic variation.
  - The dynamics of the poterior variance $\widehat{v}_t$ is locally deterministic, but stochastic.
  - Intuition: The investor tries to learn about regression coefficient β_t
  - $\Longrightarrow$ the amount of information depends on $(x_t - \overline{x})$.

Strategic Asset Allocation with Learning *About* Predictability

- The rest of the set up is the same. In this case, the value function will depend on two additional state variables, $\widehat{\beta}$ and $\widehat{v}$

$$J(W_t, \widehat{\beta}, \widehat{v}, t) = e^{-\rho t} \frac{W^{1-\gamma}}{1-\gamma} F(\widehat{\beta}, \widehat{v}, t)$$

- The optimal portfolio is given by

$$\begin{aligned} \theta_t = & \frac{\overline{\lambda} + \widehat{\beta}(x_t - \overline{x})}{\gamma \sigma_R^2} \\ & + \frac{F_x}{\gamma \sigma_R^2 F} \sigma_R \sigma_x \rho_{R,x} + \frac{F_\beta}{\gamma \sigma_R^2 F} \sigma_R \sigma_\beta \rho_{R,\beta} \\ & + \frac{F_\beta}{\gamma \sigma_R^2 F} \widehat{v}_t (x_t - \overline{x}) \end{aligned}$$

- A lot of hedging demands!
  - The last term is due to learning

Strategic Asset Allocation with Learning *About* Predictability

- Xia finds $F_\beta < 0$ (i.e. $J_\beta > 0$): Stronger predictability increases the J
- This implies that hedging demand for parameter uncertainty is negative when $x_t > \overline{x}$ and positive when $x_t < \overline{x}$.
  - $\Longrightarrow$ Invest heavily when D/P is low and get out when D/P is high.
- Intuitively, the conditional covariance between realized returns dR and state variable $d\widehat{\beta}$ changes over time.
  - When the $x_t > \overline{x}$, then an unusually high return implies that $\tilde{\beta}_t$ is too low
    - $\Longrightarrow \tilde{\beta}_t$ must go up in the next step.
  - When $x_t < \overline{x}$, then $\tilde{\beta}_t$ is multiplying a negative number. It follows that an unusually high return imply that $\tilde{\beta}_t$ is too high
    - $\Longrightarrow \tilde{\beta}_t$ must go down in the next step.
- As usual, if $cov(dR, d\widehat{\beta}) > 0 \Longrightarrow$ heding demand < 0 (and viceversa)
- The flip-flop in the conditional covariance implies changes in hedging demand, which is positive for low x_t and negative for high x_t.

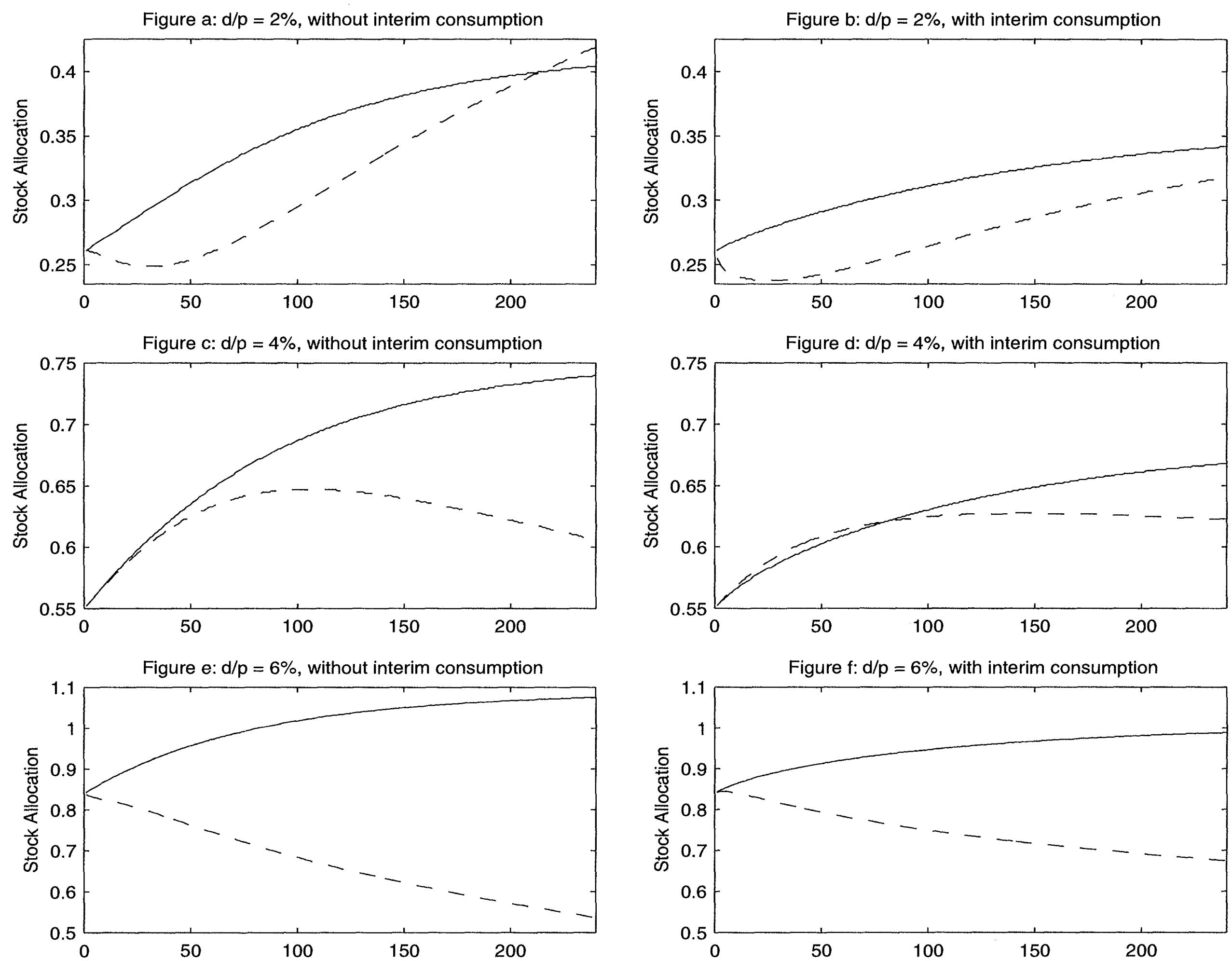


Figure 1. The term structure of optimal portfolio allocation: effect of learning ($\gamma = 5.0$ and $b_0 = 1.5$). This figure plots the optimal proportion of wealth allocated to stock for a horizon of 1 month to 240 months. The *risk-aversion* parameter, γ, is five and the investor's prior estimate, b_0, is 1.5. The optimal portfolio allocation when the investor faces a certain predictability is compared with the allocation when the investor faces uncertain predictability. When there is no parameter uncertainty, the optimal stock allocation increases with the horizon; when there is parameter uncertainty, the optimal stock allocation may increase or decrease with the horizon depending on the current value of the dividend yield d/p. $\nu_0 = 0.0$ (without parameter uncertainty): solid line; $\nu_0 = 4.0$ (with parameter uncertainty): dashed line.

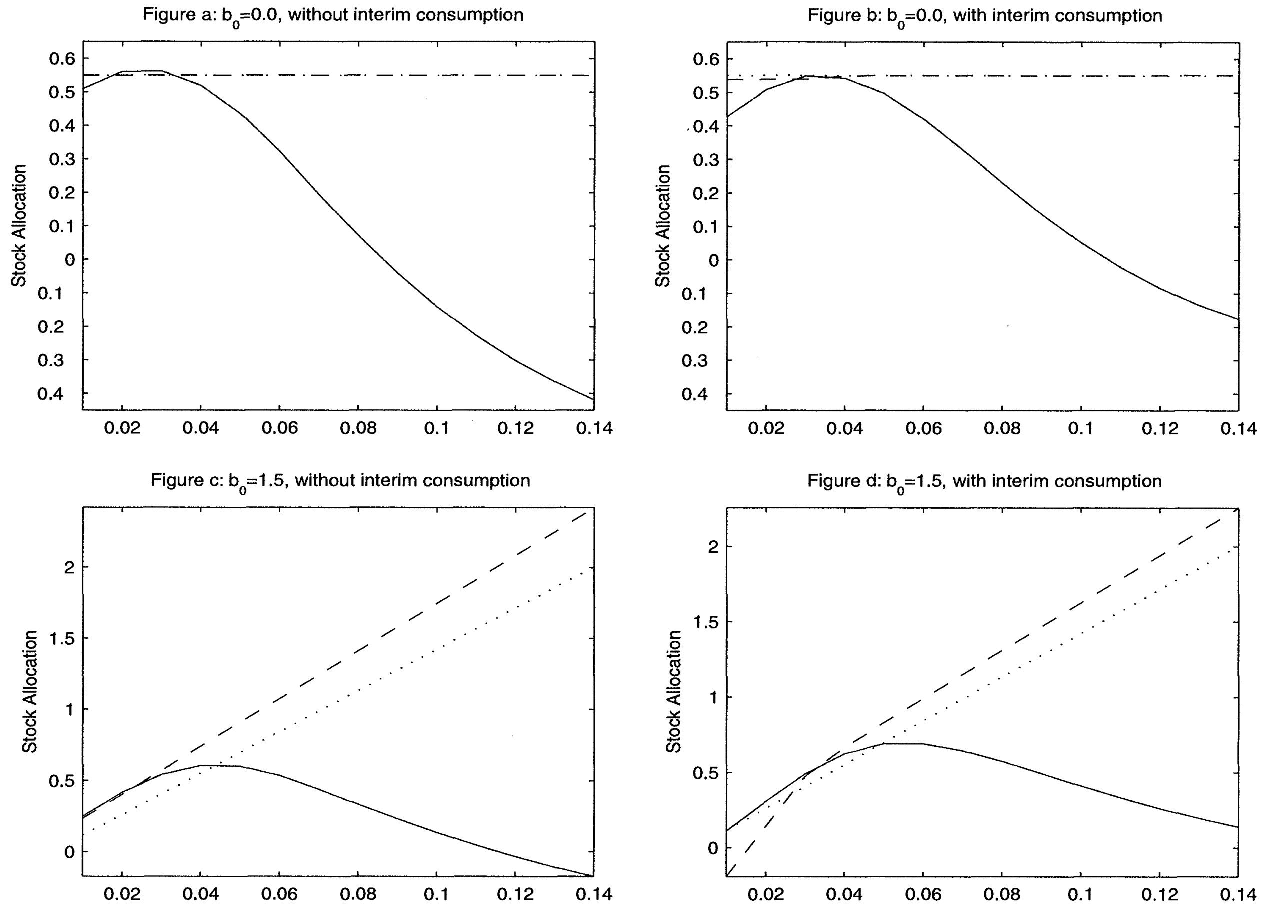


Figure 2. Optimal allocation conditional on dividend yield: effect of parameter uncertainty on market timing (γ = 5.0, ν_0 = 4.0, and T = 20 years). This figure plots the optimal proportion of wealth allocated to stock as a function of the current observation of dividend yield, d/p. The *risk-aversion* parameter, γ, is five, the investment horizon, T, is fixed at 20 years, and the investor's prior estimate, b_0, is zero in the first two panels and is 1.5 in the last two panels. The prior parameter uncertainty, ν_0, is four. The allocations under the myopic strategy, the dynamic strategy without learning, and the optimal strategy are compared in each panel. The stock allocation under the myopic and the dynamic strategy without learning is monotonically increasing in dividend yield, whereas the allocation under the optimal strategy first increases and then decreases with dividend yield. Optimal strategy: solid line; dynamic strategy without parameter uncertainty: dashed line; myopic strategy: dotted line.

Strategic Asset Allocation with Learning *About* Everything

- What if you do not know anything about the parameters?
  - In this case the portfolio selection is clearly rather hard. No analytical solutions.
  - Main issue: lot of state variables with non-lineas dynamics. For instance:

$$\mathbf{Y}_{t+1} = \mathbf{B}\mathbf{X}_t + \boldsymbol{\varepsilon}_{t+1}; \quad \text{where}$$
$$\mathbf{Y}_{t+1} = (r^e_{t+1}, log(D_t/P_t))'; \quad \mathbf{X}_t = (1, log(D_t/P_t))'; \quad \boldsymbol{\varepsilon}_{t+1} \sim N(0, \boldsymbol{\Sigma})$$

  - If non-informative priors about all of the parameters of $(\mathbf{B}, \boldsymbol{\Sigma})$ the posterior

$$p\left(\boldsymbol{\Sigma}^{-1} | \mathbf{Y}, \mathbf{X}\right) = Wishart(T-3, \mathbf{S}^{-1})$$
$$p\left(vec(\mathbf{B}) | \boldsymbol{\Sigma}, \mathbf{Y}, \mathbf{X}\right) = N\left(vec(\widehat{\mathbf{B}}), \boldsymbol{\Sigma} \otimes (\mathbf{X}'\mathbf{X})^{-1}\right)$$

  - where $\mathbf{S} = (\mathbf{Y} - \mathbf{X}\widehat{\mathbf{B}})'(\mathbf{Y} - \mathbf{X}\widehat{\mathbf{B}})$ and $\widehat{\mathbf{B}} = (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{Y})$
  - $\Longrightarrow$ posterior density fully characterized by matrices $(\mathbf{X}'\mathbf{X})$, $(\mathbf{X}'\mathbf{Y})$ and $\mathbf{S}$.
  - $\Longrightarrow$ state variables = elements of these matrices + $log(D/P) \Longrightarrow 11$
  - Learning implicit in the dynamics of posterior density.

A Simulation Approach to Portfolio Selection

- Brandt, Goyal, Santa-Clara, Stroud (RFS, 2005) use a simulations approach to portfolio selection to answer these questions.
- Idea is the following: Consider maximization problem

$$V_t(W_t, \mathbf{Z}_t) = \max_{\{\theta_s\}_{s=t}^T} E_t(u(W_T)) \quad \text{with} \quad W_{s+1} = W_s(\theta_s R^e_{s+1} + R^f)$$

- E.g. for CARA utility, one can rewrite the Bellman equation as follows

$$\begin{aligned}
V_t(W_t, \mathbf{Z}_t) &= \max_{\theta_t} E_t \left\{ \max_{\{\theta_s\}_{s=t+1}^T} E_{t+1} \left[\frac{W_T^{1-\gamma}}{1-\gamma} \right] \right\} \\
&= \max_{\theta_t} E_t \left\{ \max_{\{\theta_s\}_{s=t+1}^T} E_{t+1} \left[W_t \frac{\left(\prod_{s=t}^T (\theta_s R^e_{s+1} + R^f) \right)^{1-\gamma}}{1-\gamma} \right] \right\} \\
&= \max_{\theta_t} E_t \left\{ \frac{[W_t(\theta_t R^e_{t+1} + R^f)]^{1-\gamma}}{1-\gamma} \max_{\{\theta_s\}_{s=t+1}^T} E_{t+1} \left[\left(\prod_{s=t}^T (\theta_s R^e_{s+1} + R^f) \right)^{1-\gamma} \right] \right\} \\
&= \max_{\theta_t} E_t \left\{ u(W_{t+1}) \psi_{t+1}(\mathbf{Z}_{t+1}) \right\}
\end{aligned}$$

A Simulation Approach to Portfolio Selection

- Because of homotheticity, renormalize $W = 1$ to obtain

$$\frac{1}{1-\gamma}\psi_t(\mathbf{Z}_t) = \max_{\theta_t} E_t \left\{ u(\theta_t R^e_{t+1} + r)\psi_{t+1}(\mathbf{Z}_{t+1}) \right\}$$

  – $\Longrightarrow$ FOC: $\quad 0 = E_t \left\{ \partial_1 u(\theta_t R^e_{t+1} + r)\psi_{t+1}(\mathbf{Z}_{t+1}) R^e_{t+1} \right\}$

  The procedure is then as follows:

- First, Taylor expansion (2nd or 4th order) around $W_t R^f$:

$$\frac{\psi_t(Z_t)}{1-\gamma} \approx \max_{\theta} E_t \left[u\left(W_t R^f\right)\psi_{t+1}(Z_{t+1}) + \partial_1 u\left(W_t R^f\right)\psi_{t+1}(Z_{t+1})\left(W_t\theta_t R^e_{t+1}\right) + \frac{1}{2}\partial_1^2 u\left(W_t R^f\right)\psi_{t+1}(Z_{t+1})\left(W_t\theta_t R^e_{t+1}\right)^2 \right]$$

  – FOC imply

$$\theta_t \approx -\frac{\partial_1 u\left(W_t R^f\right)}{\partial_1^2 u\left(W_t R^f\right) W_t} \frac{E_t\left[\psi_{t+1}(\mathbf{Z}_{t+1})\left(R^e_{t+1}\right)\right]}{E_t\left[\psi_{t+1}(\mathbf{Z}_{t+1})\left(R^e_{t+1}\right)^2\right]} \approx \frac{1}{\gamma}\frac{E_t\left[\psi_{t+1}(\mathbf{Z}_{t+1})\left(R^e_{t+1}\right)\right]}{E_t\left[\psi_{t+1}(\mathbf{Z}_{t+1})\left(R^e_{t+1}\right)^2\right]}$$

  – $\Longrightarrow$ If we could compute these expectations, we are done.

A Simulation Approach to Portfolio Selection

- Second, simulate many paths of $\mathbf{Y}_t = (R_t^e, \mathbf{Z}_t)$
  - E.g. in the predictability with learning,
    * start from posterior $p_t(B, \Sigma)$ and sample the parameters B, Σ;
    * simulate realizations for R_t^e and $log(D/P)$;
    * update the postior given the realization (i.e. the 10 parameters of the posterior density);
    * obtain a series of $\mathbf{Y}_t = (R_t^e, log(D/P)_t, \mathbf{Z}_t)$ for $t = 1, .., T$;
  - Let a realized simulation be denoted by $\mathbf{Y}_t^m$, $m = 1, ..., 10000$
- Third, compute expectations through a regression.
  - One difficulty in evaluating the expecations is that ψ_{t+1} itself depends on an conditional expectation

$$\psi_{t+1}(\mathbf{Z}_{t+1}) = E_{t+1}\left[\left(\prod_{s=t+1}^{T-1} \theta_s R_{s+1}^e + R^f\right)^{1-\gamma}\right]$$

A Simulation Approach to Portfolio Selection

– Use law of iterated expectations to obtain (check computations!)

$$\theta_t \approx \frac{1}{\gamma} \frac{E_t\left[\left(\prod_{s=t+1}^{T-1} \theta_s R^e_{s+1} + R^f\right)^{1-\gamma} \left(R^e_{t+1}\right)\right]}{E_t\left[\left(\prod_{s=t+1}^{T-1} \theta_s R^e_{s+1} + R^f\right)^{1-\gamma} \left(R^e_{t+1}\right)^2\right]}$$

– So, the question is how to compute the expectations $E\left[h_{t+1}|\mathbf{Z}_t\right]$ for h_{t+1} any of those arguments above.

– Use a regression technique: parametrize

$$E\left[h_{t+1}|\mathbf{Z}_t\right] = \xi(\mathbf{Z}_t)\beta_t$$

where $\xi(\mathbf{Z}_{t+1})$ is a set of basis functions, and β_t has to be estimtaed. E.g.

$$\xi(\mathbf{Z}_t) = [1, Z_t, Z_t^2, ...]$$

– Where is β coming from? Rewrite

$$h_{t+1} = \xi(\mathbf{Z}_t)\beta_t + \epsilon_{t+1}$$

– If we had some observations h_{t+1} and some Z_t we could run the regression.

A Simulation Approach to Portfolio Selection

- Use simulations and a backward computation to obtain the observations h_t's.
- In particular, moving backward, for each t, compute realizations of h_{t+1}. E.g.

$$h_{t+1}^{m} = \left(\prod_{s=t+1}^{T-1} \widehat{\theta}_s^{m} R_{s+1}^{e\ m} + R^f \right)^{1-\gamma} \left(R_{t+1}^{e\ m} \right)$$

  – where the $\widehat{\theta}_s^{m}$ is the estimated optimal policy in future steps (already computed, simulation by simulation using the formula below).
- Use *cross-sectional* regression to estimate β_t across simulations

$$h_{t+1}^{m} = \xi(\mathbf{Z}_t^{m})\beta_t + \epsilon_{t+1}^{m}$$

- The *fitted* value correspond to the expectation

$$E[h_{t+1}|\mathbf{Z}_t] = \xi(\mathbf{Z}_t^{m})\widehat{\beta}_t$$

  – Substituting the expectations in the portfolio rule (for the two different h_{t+1}):

$$\widehat{\theta}_t \approx \frac{1}{\gamma} \frac{\xi(\mathbf{Z}_t^{m})\widehat{\beta}_{N,t}}{\xi(\mathbf{Z}_t^{m})\widehat{\beta}_{D,t}}$$

  – where $\widehat{\beta}_{N,t}$ and $\widehat{\beta}_{D,t}$ = regression coef. to compute the "Num." and "Denom."

A Simulation Approach to Portfolio Selection

- Brandt et al. (2005, RFS) use a more general model, including
  - Intermediate consumption;
  - Higher order expansion;
  - Various basis functions.
- See paper for tables and figures (cannot copy and paste them).

Strategic Asset Allocation with Model Misspecification

- What if investors are uncertain about the "model" and would like to take decisions that are "robust" to small misspecification?
  - We now discuss preferences for robustness and their implications for strategic portfolio allocation
  - The framework is the one of Anderson, Hansen, Sargent (ReStud 1999) as well as Maenhout (RFS, 2004)

- Consider (again!) the usual setting, with

$$d\mathbf{R}_t = (r + \boldsymbol{\lambda}_t)\, dt + \boldsymbol{\sigma} d\mathbf{B}_t$$

$$d\boldsymbol{\lambda}_t = (\mathbf{A}_0 + \mathbf{A}_1 \boldsymbol{\lambda}_t)\, dt + \boldsymbol{\Sigma} d\mathbf{B}_t$$

  - Let P denote the probability measure that is defined by these processes.
  - We call this the "reference model".

Modeling “Model Misspecification”

- The investor is worried about “small” model misspecification.
- Two questions:
  1. How can we model a model misspecification?
  2. How can we model investor “aversion” to such misspecification?

- We can model “model misspecification” by introducing a set of “plausible” probability measures Q that are “close” to the original one P.
- In continuous time, we can “perturb” the reference model and obtain new probability measures Q by replacing $d\mathbf{B}_t$ by

$$d\mathbf{B}_t = d\widehat{\mathbf{B}}_t + \mathbf{h}_t dt$$

- where $\mathbf{h}_t$ is another stochastic process.

Modeling "Model Misspecification"

- The class of misspecified models is then those defined by the

$$d\mathbf{R}_t = (r + \boldsymbol{\lambda}_t)\, dt + \boldsymbol{\sigma} \left(d\widehat{\mathbf{B}}_t + \mathbf{h}_t dt \right)$$

$$d\boldsymbol{\lambda}_t = (\mathbf{A}_0 + \mathbf{A}_1 \boldsymbol{\lambda}_t)\, dt + \boldsymbol{\Sigma} \left(d\widehat{\mathbf{B}}_t + \mathbf{h}_t dt \right)$$

- for "plausible" $\mathbf{h}_t$ processes.
- How can we introduce "preferences" for robustness?

Modeling "Model Misspecification"

- The *multiplier robust control problem* can be formulated as

$$\sup_{C,\boldsymbol{\theta}} \inf_{\mathbf{h}} \left\{ \widehat{E} \left[\int_0^T e^{-\rho t} \left(u\left(C_t\right) + \frac{\eta}{2} \mathbf{h}_t \mathbf{h}_t' \right) dt \right] \right\}$$

  – subject to the "perturbed" budged equations

$$dW_t = \left(W_t \left(\boldsymbol{\theta}_t' \boldsymbol{\lambda}_t + r \right) - C_t \right) dt + W_t \boldsymbol{\theta}_t' \boldsymbol{\sigma} \left(d\widehat{\mathbf{B}}_t + \mathbf{h}_t dt \right)$$

- Here η is a penalty imposed on the discrepancy between Q and P.
- For given η, the "robust" investor
  1. considers the probabilities Q (each defined by a process $\mathbf{h}_t$) that lead to low utility ($\inf_{\mathbf{h}}$ part)
  2. maximizes utility taking into account these worst case scenarios ($\max_{C,\boldsymbol{\theta}}$ part)

Modeling "Model Misspecification"

- A high η implies a choice of $\mathbf{h}_t$ that is close to 0, i.e. a probability Q that is close to P, because we are taking the "inf" with respect to $\mathbf{h}_t$.
  - If $\eta = 0$, we consider all the possible Q's.
  - If $\eta = \infty$, we consider only P.

Strategic Asset Allocation with Model Misspecification

- How can we solve this "max min" problem?
- It is convenient to stack all the state variables. Define $\mathbf{Y}_t = \left(W_t, \boldsymbol{\lambda}_t'\right)'$, so that we have

$$d\mathbf{Y}_t = \boldsymbol{\mu}_Y \left(\mathbf{Y}_t, \boldsymbol{\theta}_t, C_t\right) dt + \boldsymbol{\sigma}_Y \left(\mathbf{Y}_t, \boldsymbol{\theta}_t, C_t\right) \left(d\widehat{\mathbf{B}}_t + \mathbf{h}_t dt\right)$$

- The following Bellman Isaac condition is the necessary condition for the solution to the max min problem
- There exists a value function $J(Y)$ such that

$$\delta J \;=\; \max_{C,\boldsymbol{\theta}} \min_{\mathbf{h}} \left\{ u(C) + \frac{\eta}{2}\mathbf{h}\mathbf{h}' + \left(\mu_Y + \sigma_Y \mathbf{h}'\right)' \mathbf{J}_Y + \frac{1}{2} tr\left(\boldsymbol{\sigma}_Y' \mathbf{J}_{YY} \boldsymbol{\sigma}_Y'\right) \right\}$$

Towards a Solution to the Asset Allocation

- Solving for the minimum $\mathbf{h}$, one obtains

$$\mathbf{h}' = -\frac{1}{\eta}\boldsymbol{\sigma}_Y'\mathbf{J}_Y$$

- Notice that then

$$\begin{aligned}\frac{\eta}{2}\mathbf{h}\mathbf{h}' &= \frac{1}{2\eta}\mathbf{J}_Y'\boldsymbol{\sigma}_Y\boldsymbol{\sigma}_Y'\mathbf{J}_Y \\ \boldsymbol{\sigma}_Y\mathbf{h}' &= -\frac{1}{\eta}\boldsymbol{\sigma}_Y\boldsymbol{\sigma}_Y'\mathbf{J}_Y\end{aligned}$$

- Substitute into Bellman Isaac equation to find

$$\delta J = \max_{C,\boldsymbol{\theta}}\left\{u\left(C\right) - \frac{1}{2\eta}\mathbf{J}_Y'\sigma_Y\sigma_Y'\mathbf{J}_Y + \mu_Y'\mathbf{J}_Y + \frac{1}{2}tr\left(\sigma_Y'\mathbf{J}_{YY}\boldsymbol{\sigma}_Y\right)\right\}$$

- This is similar to earlier problem.

Optimal Consumption and Asset Allocation under Model Misspecification

- The FOC with respect to consumption lead to the usual condition

$$u_C = J_W$$

  - But J_W is different from before. It will depend on robustness preferences

- Instead, the FOC for optimal portfolio weights imply

$$\begin{aligned}\boldsymbol{\theta}_t \;=\; & \frac{-J_W}{W_t\left(J_{WW} - \frac{1}{\eta}J_W^2\right)}\left(\boldsymbol{\sigma}\boldsymbol{\sigma}'\right)^{-1}\left(\boldsymbol{\lambda}_t\right) \\ & + \frac{-1}{W_t\left(J_{WW} - \frac{1}{\eta}J_W^2\right)}\left(\boldsymbol{\sigma}\boldsymbol{\sigma}'\right)^{-1}\boldsymbol{\sigma}\boldsymbol{\Sigma}'\mathbf{J}_{W\boldsymbol{\lambda}} \\ & + \frac{\frac{1}{\eta}J_W}{W_t\left(J_{WW} - \frac{1}{\eta}J_W^2\right)}\left(\boldsymbol{\sigma}\boldsymbol{\sigma}'\right)^{-1}\boldsymbol{\sigma}\boldsymbol{\Sigma}'\mathbf{J}_{\boldsymbol{\lambda}}\end{aligned}$$

Strategic Asset Allocation under Model Misspecification

- The portfolio rule has then three components:

  1. Standard myopic demand.
     - Notice that the denominator is adjusted for robustness, implying a lower investment in the stocks (because $J_W^2 \frac{1}{\eta} > 0$).
  2. The standard Merton's hedging demand.
  3. An additional hedging demand arising from robustness preferences.
     - If $\eta \to \infty$, i.e. we consider the class of probability Q that are closer and closer to the reference P, we have back the usual results.
     - Note in particular that the last term drops out.

An Exact Solution for the Original Merton Problem

- Consider the original setting without time varying expected returns.
  - i.e. $\mathbf{A}_0 = 0$, $\mathbf{A_1} = 0$ and $\mathbf{\Sigma} = 0$
- In this case, the FOC with respect to $\mathbf{h}_t$ yield

$$\mathbf{h}_t = -\frac{1}{\eta}\boldsymbol{\sigma}'_W J_W$$

- and the Bellman Isaac equation is then given by

$$\delta J = \max_{C,\boldsymbol{\theta}} \left\{ u\left(C\right) - \frac{1}{2\eta} J_W^2 \boldsymbol{\sigma}_W \boldsymbol{\sigma}'_W + \mu_W J_W + \frac{1}{2} J_{WW} \boldsymbol{\sigma}_W \boldsymbol{\sigma}'_W \right\}$$

- Using $u_c = J_W$ we obtain

$$\boldsymbol{\theta}_t = \frac{-J_W}{W_t\left(J_{WW} - \frac{1}{\eta}J_W^2\right)} \left(\boldsymbol{\sigma}\boldsymbol{\sigma}'\right)^{-1} \left(\boldsymbol{\mu} - r\mathbf{1}_d\right)$$

An Exact Solution for the Original Merton Problem

- One complication with the previous problem is that, generically, it is not "scale invariant"
  - It is hard to solve as the solution depends on wealth.

- Maenheut (2004) proposes to scale the penalty parameter η by the value function J itself, in a way to make the model again scale independent.

$$\eta = \eta\left(J\right) = \eta^{*}\left(1-\gamma\right) J\left(W,t\right)$$

- The value function is then given by

$$J\left(W,t\right) = \left(\frac{1-e^{-a(T-t)}}{a}\right)^{\gamma} \frac{W^{1-\gamma}}{1-\gamma}$$

- where

$$a = \frac{1}{\gamma}\left[\rho - \left(1-\gamma\right) r - \frac{1-\gamma}{2\left(\gamma+\eta\right)}\left(\boldsymbol{\mu} - r\mathbf{1}_n\right)'\left(\boldsymbol{\sigma}\boldsymbol{\sigma}'\right)^{-1}\left(\boldsymbol{\mu} - r\mathbf{1}_n\right)\right]$$

An Exact Solution for the Original Merton Problem

- The optimal consumption and asset allocation are

$$C_t = \frac{a}{1 - e^{-a(T-t)}} W_t$$

$$\boldsymbol{\theta}_t = \frac{1}{\gamma + 1/\eta^*} \left(\boldsymbol{\sigma}\boldsymbol{\sigma}'\right)^{-1} \left(\boldsymbol{\mu} - r\mathbf{1}_d\right)$$

  – Preferences for robustness clearly go in the right direction to "solve" the asset allocation puzzle

  – A lower η^* translates into a higher "aversion" to model misspecification.

  – In this case, the allocation to stocks decreases.

  – Yet, the allocation is still independent of life expectancy $T - t$.

    * We need to introduce predictability for that.

How much pessimism is plausible?

- Clearly, by decreasing η^* we can match any empirically observed level of asset holdings.
- However, the question is then what is a "reasonable" level of η^*.
- Consider the case $n = 1$ (one stock) for simplicity.
  - For each level of η^*, there is a given worst case scenario, defined by the FOC

$$h_t = -\frac{1}{\eta}\sigma_W J_W = -\frac{1}{(1+\gamma\eta^*)\,\sigma}(\mu - r)$$

  - where I substitute for $\sigma_W = W\theta_t\sigma$, J_W and $\eta = \eta^*(1-\gamma)J$.
- A robust investor thinks that stock returns are given by

$$dR_t = (\mu + \sigma h_t)dt + \sigma d\widehat{B}_t$$

How much pessimism is plausible?

- Thus, the equity premium for a robust investor is

$$E_t^h[dR - r] = (\mu + \sigma h_t) - r = (\mu - r)\left(1 - \frac{1}{1 + \gamma\eta^*}\right)$$

- We can use the "implied" perceived equity premium of the robust investor as a reasonable metric to asset whether η^* is too small.

Optimal Portfolio Allocation under Robustness

	γ									
	2		4		6		8		10	
η	θ	$E^h[dR]$	θ	$E^h[dR]$	θ	$E^h[dR]$	θ	$E^h[dR]$	θ	$E^h[dR]$
0.1	22.79	1.17	19.53	2.00	17.09	2.63	15.19	3.11	13.67	3.50
0.2	39.06	2.00	30.38	3.11	24.86	3.82	21.03	4.31	18.23	4.67
0.5	68.36	3.50	45.57	4.67	34.18	5.25	27.34	5.60	22.79	5.83
1	91.15	4.67	54.69	5.60	39.06	6.00	30.38	6.22	24.86	6.36
2	109.38	5.60	60.76	6.22	42.07	6.46	32.17	6.59	26.04	6.67
10	130.21	6.67	66.69	6.83	44.83	6.89	33.76	6.91	27.07	6.93
100	136.04	6.97	68.19	6.98	45.50	6.99	34.14	6.99	27.32	6.99

Recent Applications of Robust Control

- The approach of robust control theory has found numerous applications in finance in recent times.

1. Liu, Pan and Wang (JF, 2005): uncertainty on rare events to explain options premia, along with the standard result on return equity premium.
2. Routledge and Zin (2004): rare events and market liquidity. $\Longrightarrow$ uncertainty aversion may lead agents not to trade after big market events.
3. Uppal and Wang (JF, 2003): extend the above model to the case of different aversions to uncertainty across assets.
   – For some assets there is less "ambiguity" about the probabilities.
   – Under-diversification: even a limited amount of aversion to uncertainty on some stocks $\Longrightarrow$ over-invest in those with less uncertainty aversion.

4. Boyle, Uppal, Wang (2005) use a similar setting to "explain" the over-investment in "own stock" puzzle.

Conclusion

- The last decade has seen a boom in research about optimal asset allocation.
- The groundwork set by Samuelson and Merton has found application only recently, as researchers were able to solve long-standing problems
  - The concept of hedging demands date back 30+ years
  - But only recently these hedging demands have been characterized in a quantitative fashion.
- Yet, we are still far from explaining all of the puzzles in a nice, convincing theory.
  - Predictability has the right implication for life cycle, but wrong for asset allocation magnitudes
  - Learning has the right implication for the magnitudes, but wrong for life cycle
  - Preferences for robustness imply unreasonable levels of pessimism.