ECON 2723, Asset Pricing, Section 3

Jennifer Walsh

October 13, 2023

Admin

- Problem set is due on Thursday, November 2.
- I will hold section **next week**, October 20, on Chapter 6.
- Please fill out a mid-semester feedback form to make section / office hours as useful as possible!
- Not covered today: 5.4.3, 5.5, 5.6 of the book.

Contents

Stochastic Discount Factor, Recap

- Recall from the lecture that M in complete markets is

$$P = \sum_s q(s)X(s) = \sum_s \pi(s)\frac{q(s)}{\pi(s)}X(s) = \sum_s \pi(s)M(s)X(s) = E[MX]$$

- Divide through by P to get fundamental equation in return form

$$\begin{aligned} 1 &= E[M(1+R)] \\ &= E[M]E[1+R] + Cov(M, 1+R) \end{aligned}$$

  Same equation for risk free asset gives

$$1 = E[M](1+R^f)$$

  subtract the two to get

$$0 = E[M]E[R-R^f] + Cov(M, 1+R)$$

$$\implies E[R-R^f] = -(1+R^f)Cov(M, 1+R)$$

- Utility maximization with initial wealth W_0:

$$\text{Max}\, u(C_0) + \sum_{s=1}^{S} \beta\pi(s)u(C(s)) \quad \text{subject to} C_0 + \sum_{s=1}^{S} q(s)C(s) = W_0$$

  implies

$$M(s) = \beta\frac{U'(C(s))}{U'(C_0)}$$

Existence of SDF

Law of one price (weaker than absence of arbitrage) + portfolio formation $\Rightarrow$ There exists an SDF that prices all payoffs within a set of tradable payoffs/assets

- Proof by construction: find a linear combination of payoffs/assets that prices all assets.

$$\mathbb{X} = (X_1 \dots X_N)' = \begin{pmatrix} X_1(1) & \cdots & X_N(1) \\ \vdots & \ddots & \vdots \\ X_1(S) & \cdots & X_N(S) \end{pmatrix}' \text{ and } X^* = \mathbb{X}'c^*$$

$$\begin{pmatrix} P(X_1) \\ \vdots \\ P(X_N) \end{pmatrix} = \begin{pmatrix} \mathbb{E}[X_1'X^*] \\ \vdots \\ \mathbb{E}[X_N'X^*] \end{pmatrix}$$

Write the condition in matrix form as

$$\mathbb{P} = \mathbb{E}[\mathbb{X}\mathbb{X}'c^*] \implies c^* = \left(\mathbb{E}[\mathbb{X}\mathbb{X}']\right)^{-1}\mathbb{P} \implies \mathbb{X}'c^* = X^* = \mathbb{X}\left(\mathbb{E}[\mathbb{X}\mathbb{X}']\right)^{-1}\mathbb{P}$$

- **Such SDF X^* is not guaranteed to be positive for each state of the world**

Positive SDF $\Leftrightarrow$ Absence of arbitrage (stronger than *law of one price*)

- Proof in the textbook
- **SDF is not guaranteed to lie in the set of tradable assets**

It is important to understand that only the SDF that is a linear combination of asset payoffs is unique. There may be many other SDFs of the form $M = X^* + \varepsilon$, where $\mathrm{E}[\mathbb{X}\varepsilon] = \mathrm{E}[\mathbb{X}(M - X^*)] = 0$. When a riskfree asset is traded these must all have higher variance than X^*, as discussed further below in the context of volatility bounds on the SDF. Equivalently, X^* is the projection of every SDF onto the space of tradable payoffs. Thus it can be thought of as the portfolio of available assets that best mimics the behavior of every SDF.

SDF Factor Structure [SKIP]

- We can rewrite the excess return equation to get

$$E[R-R^f] = \underbrace{\frac{Cov(M,1+R)}{Var(M)}}_{\text{Quantity of risk } \beta} \underbrace{\left[-\frac{Var(M)}{E[M]}\right]}_{\text{Price of risk } \lambda}$$

- Suppose that for factors f_k that have mean zero and orthogonal to each other

$$M = a - \sum b_k f_k$$

Risk free rate is $1+R^f = 1/E[M] = a$. Excess return is

$$E[R-R^f] = -(1+R^f)Cov(M,1+R) = -\frac{1}{a}Cov(\sum b_k f_k, R)$$
$$= -\frac{1}{a}\sum b_k Cov(f_k,R) = \sum \frac{Cov(f_k,R)}{Var(f_k)}\left(-\frac{b_k}{a}Var(f_k)\right) = \sum \beta_k \cdot \lambda_k$$

- SDF for CAPM takes the same form $M = a - b(R^W - E[R^W])$. To find a and b can price two assets: risk free bond and the wealth portfolio:

$$1 = E[a - b(R^W - E[R^W])](1+R^f) \implies (1+R^f) = \frac{1}{E[a]}$$
$$1 = E[a - b(R^W - E[R^W])(1+R^W)]$$

Factor structure

- What did we learn?
- Factor models (with mean zero factors f) are equivalent to linear models for the discount factor m, that is

$$\mathbb{E}(1+R^i) = \gamma + \lambda' \beta_i \iff m = a + b' f.$$

-

$$E[R-R^f] = \underbrace{\frac{Cov(M,1+R)}{Var(M)}}_{\text{Quantity of risk } \beta} \underbrace{\left[-\frac{Var(M)}{E[M]}\right]}_{\text{Price of risk } \lambda}$$

- Given a and b from the SDF, can determine γ and λ of the factor model. Conversely, if you know γ and λ from the factor model, you can learn a and b of the SDF.

HJ Bound

- Volatility bounds connect Sharpe ratios and the volatility of the SDF, and we can interpret them as the **mean-variance frontier of all discount factors that price a given set of assets**. Hansen & Jagannathan do this without a risk free asset.
- Rearranging the equation on the previous slide:

$$
\begin{aligned}
E\left[R_{i,t+1} - R_{f,t+1}\right] &= -\left(1 + R_{f,t+1}\right)\sigma_t\left(M_{t+1}\right)\sigma_t\left(R_{i,t+1}\right)\operatorname{corr}_t\left(M_{t+1}, R_{i,t+1} - R_{f,t+1}\right)\right) \\
&= -\frac{\sigma_t\left(M_{t+1}\right)\sigma_t\left(R_{i,t+1} - R_{f,t+1}\right)\operatorname{corr}_t\left(M_{t+1}, R_{i,t+1} - R_{f,t+1}\right)\right)}{E_t\left[M_{t+1}\right]}
\end{aligned}
$$

$$
|E(R^e)| \leq \frac{\sigma_t\left(M_{t+1}\right)\sigma_t\left(R_{i,t+1} - R_{f,t+1}\right)}{E_t\left[M_{t+1}\right]}
$$

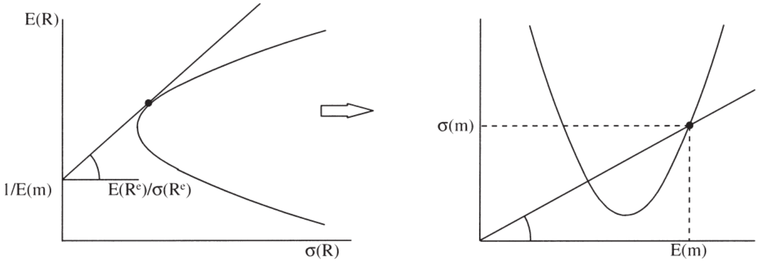
$$
\Rightarrow \frac{\sigma(M)}{\mathbb{E}(M)} \geq \frac{|\mathbb{E}(R^e)|}{\sigma(R^e)}
$$

- Hence, the boundary can be constructed by noting that

$$
\min_{\{\text{all } m \text{ that price } x \in \underline{X}\}} \frac{\sigma(m)}{\mathbb{E}(m)} = \max_{\{\text{all excess returns } R^e \in \underline{X}\}} \frac{\mathbb{E}(R^e)}{\sigma(R^e)}. \tag{1}
$$

  - In the Lecture, we fixed different $\mathbb{E}(m)$ and found the minimum variance SDF with each mean in order to build the frontier.
  - The HJ bound is a restriction on the set of *discount factors* that can price a given set of returns, as well as a restriction on the set of returns we will see given a specific discount factor.

HJ Bound (2)



1. Provided the maximum Sharpe ratio of a given set of test assets (tangency portfolio) on the left panel, determine the possible SDFs by tracing a line with such slope on the right panel.
2. Equivalently, given the ratio $\sigma(m)/\mathbb{E}(m)$ in the right panel, then the maximum Sharpe ratio on left panel can be determined.

HJ Bound and the Equity Premium Puzzle

- Consider the power utility function with $u'(c) = c^{-\gamma}$. Substituting into the HJ bound, we have:
$$\left| \frac{E(R^{mv}) - R^f}{\sigma(R^{mv})} \right| \leq \frac{\sigma\left[(c_{t+1}/c_t)^{-\gamma}\right]}{E\left[(c_{t+1}/c_t)^{-\gamma}\right]} \tag{2}$$
- The standard deviation is large if consumption is volatile or if γ is large.
- If consumption growth is lognormal, then
$$\left| \frac{E(R^{mv}) - R^f}{\sigma(R^{mv})} \right| \leq \sqrt{e^{\gamma^2 \sigma^2(\Delta \ln c_{t+1})} - 1} \approx \gamma \sigma(\Delta \ln c) \tag{3}$$
- If average real stock returns are 9% with standard deviation 16% and treasuries are 1%, then $SR = 0.5$. Aggregate consumption growth volatility has been about 1%, implying $\gamma = 50$.

Contents

Why do prices move?

- Dividend-Based Models: We defined returns as:

$$1+R_{t+1}=\frac{P_{t+1}+D_{t+1}}{P_t} \tag{4}$$

- In a world with constant expected returns (equiv. discount rates), this becomes

$$P_t=\mathrm{E}_t\left[\frac{P_{t+1}+D_{t+1}}{1+R}\right] \tag{5}$$

Solving forward:

$$P_t=\mathrm{E}_t\left[\sum_{k=1}^{K}\left(\frac{1}{1+R}\right)^k D_{t+k}\right]+\mathrm{E}_t\left[\left(\frac{1}{1+R}\right)^k P_{t+K}\right]. \tag{6}$$

Time Varying Discount Rates (1): Log-Linearization

- The evidence that prices are more volatile than realized dividends motivates time varying discount rates (i.e., expected returns).
- Working with time varying DR is hard as they feature both addition and multiplication

$$P_t = \mathbb{E}_t \sum_{j=1}^{\infty} \frac{D_{t+j}}{\prod_{k=1}^{j}(1+R_{t+k})}$$

- But we already encountered this problem in Chapter 2 and solved it by applying a log-linear approximation.

Time Varying Discount Rates (2): Log-Linearization

- Our approach here is similar to that in Chapter 2: log-linearize the return around a constant price-dividend ratio

$$
\begin{aligned}
r_{t+1} = \log(1+R_{t+1}) &= \log\left(\frac{P_{t+1}+D_{t+1}}{P_t}\right) = \log\left(\frac{P_{t+1}}{P_t}\left(1+\frac{D_{t+1}}{P_{t+1}}\right)\right) \\
&= p_{t+1} - p_t + \log(1+\exp(d_{t+1}-p_{t+1})) \\
&\approx p_{t+1} - p_t + \log(1+\exp(\overline{d-p})) + \underbrace{\frac{\exp(\overline{d-p})}{1+\exp(\overline{d-p})}}_{1-\rho}(d_{t+1}-p_{t+1}-\overline{d-p}) \\
&= p_{t+1} - p_t + k + (1-\rho)(d_{t+1}-p_{t+1}) \\
&= k + \rho p_{t+1} + (1-\rho) d_{t+1} - p_t
\end{aligned}
$$

- **Importantly:** in Chapter 2 the accuracy improved when time interval got smaller (since we approximated around zero risk premium). Here, accuracy depends on how much the price-dividend ratio moves in relation to its mean over time (so the approximation does **not** get better as the time interval shrinks).

Time Varying Discount Rates (3): Price-Dividend Ratio

- We can rearrange
$$r_{t+1} = k + \rho p_{t+1} + (1-\rho) d_{t+1} - p_t ,$$
and isolate p_t
$$p_t = k + \rho p_{t+1} + (1-\rho) d_{t+1} - r_{t+1} .$$

- Iterate this expression forward to get
$$p_t = \frac{k}{1-\rho} + \underbrace{\sum_{j=0}^{\infty} \rho^j (1-\rho) d_{t+1+j}}_{p_{CF,t}} - \underbrace{\sum_{j=0}^{\infty} \rho^j r_{t+1+j}}_{p_{DR,t}} + \underbrace{\lim_{j\to\infty} \rho^j p_{t+j}}_{=0} .$$
  - Higher price today ($p_t \uparrow$) with unchanged future dividends ($p_{CF,t} = const$) implies lower future returns as captured by $p_{DR,t} \downarrow$.
  - Higher returns in the future ($p_{DR,t} \uparrow$) with unchanged dividends ($p_{CF,t} = const$) implies lower price today ($p_t \downarrow$).

Time Varying Discount Rates (4): Price-Dividend Ratio

- When dividends follow a unit root process (i.e. they are very persistent) may want to work in differences

$$d_t - p_t = -\frac{k}{1-\rho} - \underbrace{\sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j}}_{dp_{CF,t}} + \underbrace{\sum_{j=0}^{\infty} \rho^j r_{t+1+j}}_{dp_{DR,t}}.$$

Return Decomposition (1)

- Take a **surprise operator** of the dividend-price ratio from the previous slide.

$$\underbrace{(\mathbb{E}_{t+1}-\mathbb{E}_t)(d_t-p_t)}_{=0} = \underbrace{-(\mathbb{E}_{t+1}-\mathbb{E}_t)\frac{k}{1-\rho}}_{=0}$$

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

$$\Longleftrightarrow$$

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

$$\Longleftrightarrow$$

$$\underbrace{(\mathbb{E}_{t+1}-\mathbb{E}_t)r_{t+1}}_{\tilde{r}_{t+1}} = \underbrace{(\mathbb{E}_{t+1}-\mathbb{E}_t)\sum_{j=0}^{\infty}\rho^j\Delta d_{t+1+j}}_{N_{CF,t}} - \underbrace{(\mathbb{E}_{t+1}-\mathbb{E}_t)\sum_{j=1}^{\infty}\rho^j r_{t+1+j}}_{N_{DR,t}}$$

Return Decomposition (2)

- We just found that

$$\tilde{r}_{t+1} = N_{CF,t} - N_{DR,t}.$$

- Some comments
  - Better information about future dividends *reduces* the volatility of returns.
  - This relation does not hinge on rational expectations (since it was derived from ex-post identities, it always holds for irrational investors that respect identities).
  - Predictable variation in returns, $V[\mathbb{E}_t(r_{t+1})]$, can be negligible at short horizons relative to the overall variation of returns, $V(r_{t+1})$, yet $(\mathbb{E}_t[r_{t+1}])$ can be an important determinant of the level of asset prices and the volatility of returns.[1]
  - The second component – news about future expected returns – is the main element of Intertemporal CAPM.
  - In most cases we work with covariances

$$Cov_t(r_{t+1}, m_{t+1}) = Cov_t(N_{CF,t}, \tilde{m}_{t+1}) - Cov_t(N_{DR,t}, \tilde{m}_{t+1}).$$

[1] More on this in the next slides.

Long vs. Short Run Predictability (1): Price-Dividend Ratio

- We want to write a simple statistical model where small but persistent changes in expected returns have large effects on prices. Such model will imply *weak form efficiency* but *no semi-strong* form efficiency.[2]

- Suppose that the statistical model for expected return is

$$\mathbb{E}_t[r_{t+1}] = \bar{r} + x_t, \text{ with } x_t = \phi x_{t-1} + \xi_t,$$

and return is a sum of expected and unexpected components,

$$r_{t+1} = \bar{r} + x_t + u_{t+1}.$$

[2] Remember, we need to specify the variables in information set at t. **Weak form efficiency**: includes past returns. **Semi-Strong form efficiency**: includes publicly available information (e.g. stock splits, dividends, earnings).

Long vs. Short Run Predictability (2): Price-Dividend Ratio

- Unexpected returns can be written as CF and DR news

$$r_{t+1} = \bar{r} + x_t + \underbrace{N_{CF,t+1} - N_{DR,t+1}}_{u_{t+1}}$$

- DR component of the price dividend ratio is

$$\begin{aligned} dp_{DR,t} &= \mathbb{E}_t \sum_{j=0}^{\infty} \rho^j r_{t+1+j} = \mathbb{E}_t \sum_{j=0}^{\infty} \rho^j [\bar{r} + x_{t+j} + u_{t+1+j}] \\ &= \frac{\bar{r}}{1-\rho} + \frac{x_t}{1-\phi\rho} \implies Var(dp_{DR,t}) = \frac{\sigma_x^2}{(1-\phi\rho)^2} \end{aligned}$$

Small volatility of expected returns (small σ_x^2) nevertheless leads to large variance in prices when persistence is high (ϕ close to 1).

Long vs. Short Run Predictability (3): Discount Rate News

- In this model discount rate news can be calculated as

$$
\begin{aligned}
N_{DR,t} &= (\mathbb{E}_{t+1} - \mathbb{E}_t) \sum_{j=1}^{\infty} \rho_j r_{t+1+j} \\
&= (\mathbb{E}_{t+1} - \mathbb{E}_t) \sum_{j=1}^{\infty} \rho_j \left[\bar{r} + x_{t+j} + u_{t+1+j}\right] \\
&= (\mathbb{E}_{t+1} - \mathbb{E}_t) \left[\rho x_{t+1} + \rho^2 x_{t+2} + \rho^3 x_{t+1} + \dots\right] \\
&= (\mathbb{E}_{t+1} - \mathbb{E}_t) \left[\rho(\phi x_t + \xi_{t+1}) + \rho^2(\phi^2 x_t + \phi \xi_{t+1} + \xi_{t+2}) + \dots\right] \\
&= (\mathbb{E}_{t+1} - \mathbb{E}_t) \left[\rho \xi_{t+1} + \rho^2 \phi \xi_{t+1} + \rho^3(\phi^2 \xi_{t+1}) + \dots\right] \\
&= \left[\rho \xi_{t+1} + \rho^2 \phi \xi_{t+1} + \rho^3(\phi^2 \xi_{t+1}) + \dots\right] \\
&= \frac{\rho \xi_{t+1}}{1 - \rho\phi} \approx \frac{\xi_{t+1}}{1 - \phi}
\end{aligned}
$$

  - We again find that, for a given variance of the expected return innovation ξ, the variance of the discount rate news is increasing in the persistence of expected return (determined by ϕ).
  - When persistence is high, returns today react a lot to small changes in expected returns since the effect accumulates over a long time.

Long vs. Short Run Predictability (4): Return Autocorrelation

$$
\begin{aligned}
&Cov(r_{t+1}, r_{t+1+j}) = \\
&= Cov(x_t + N_{CF,t+1} - N_{DR,t+1}, x_{t+j} + N_{CF,t+1+j} - N_{DR,t+1+j}) \\
&= Cov\left(x_t + N_{CF,t+1} - \frac{\rho\xi_{t+1}}{1-\rho\phi}, \phi^j x_t + \phi^{j-1}\xi_{t+1} + \phi^{j-2}\xi_{t+2} + \cdots + N_{CF,t+1+j} - \frac{\rho\xi_{t+1+j}}{1-\rho\phi}\right) \\
&= \phi^j\sigma_x^2 + Cov(N_{CF,t+1}, \phi^{j-1}\xi_{t+1}) - Cov\left(\frac{\rho\xi_{t+1}}{1-\rho\phi}, \phi^{j-1}\xi_{t+1}\right) \\
&= \phi^j \frac{\sigma_\xi^2}{1-\phi^2} + \phi^{j-1}Cov(N_{CF,t+1}, \xi_{t+1}) - \frac{\rho\phi^{j-1}}{1-\rho\phi}\sigma_\xi^2 \\
&= \phi^{j-1}\left[Cov(\underbrace{N_{CF,t+1}}_{\text{dividend news}}, \underbrace{\xi_{t+1}}_{\text{revision in exp. returns}}) + \sigma_\xi^2\left(\underbrace{\frac{\phi}{1-\phi^2}}_{\text{autocorr. returns}} - \underbrace{\frac{\rho}{1-\rho\phi}}_{\text{capital loss when realized returns rise}}\right)\right]
\end{aligned}
$$

- Some comments...
  - All autocovariances can be (close to) zero if the terms in squared brackets cancel out.
  - This implies that there is no predictability of future returns from past returns even though there exists a state variable x_t that predicts returns.
  - Prices can be weak form efficient but not semi-strong form efficient in the sense that not all relevant information regarding future returns is contained in past returns.

Running Predictive Regressions

- Often interested in estimating predictive regressions

$$r_t = \alpha + \beta x_{t-1} + u_t,$$

  for $t = 1, \ldots, T$, where:
  - r_t is the return on some asset during period t;
  - x_{t-1} is a scaled-price ratio (e.g., a dividend yield, book-to-price ratio, a yield spread) that is directly related to asset prices at $t-1$ or x_{t-1} is some corporate action that is indirectly related to asset prices;
  - u_t is the regression's disturbance;

- Two critical issues in applied time-series work in finance:
  1. Getting your standard errors right $\rightarrow$ Time-series data rarely generated by an iid process!
  2. Small sample Stambaugh bias because strict exogeneity fails $\rightarrow$ Although $E[u_t \mid x_{t-1}, x_{t-2}, \cdots] = 0$ so x_{t-1} is sequentially exogenous, often the case that $E[u_t \mid x_t, x_{t+1}, x_{t+2}, \cdots] \neq 0$ so the independent variable is not strictly exogenous. $\Rightarrow$ OLS estimates of α and β will typically be biased in finite samples.

Running Predictive Regressions: Stambaugh (1999) Bias

- **Kendall bias**: get a downward bias in estimating ϕ in

$$x_{t+1} = (1-\phi)\overline{x} + \phi x_t + \xi_{t+1}$$

The process appear to be less persistent that it actually is
When $\bar{x} = 0$,

$$E[\hat{\phi} - \phi] = -(\frac{1+3\phi}{T}) + o(\frac{1}{T^2})$$

- Stambaugh (1999) shows that this leads to a bias in a predictive return regression

$$r_{t+1} = \alpha + \beta x_t + u_{t+1}$$

when x_t is persistent.

- In particular, the bias is

$$E[\hat{\beta} - \beta] = \frac{cov(\xi, u)}{var(\xi)} E[\hat{\phi} - \phi]$$

- When x_t is Dividend-Price ratio then $cov(\xi, u) < 0$ so that

$$E[\hat{\beta} - \beta] = \underbrace{\frac{cov(\xi, u)}{var(\xi)}}_{<0} \underbrace{E[\hat{\phi} - \phi]}_{<0} > 0$$

which undermines predictive regressions.

Responses to Stambaugh Bias (1)

1. Lewellen (2004) builds a worst case scenario adjustment (given that $\phi < 1$) for β:

$$\mathbb{E}[\hat{\beta} - \beta | \hat{\phi}, \phi] = \frac{cov(\xi, u)}{var(\xi)}[\hat{\phi} - \phi] \overset{\phi<1}{\Longrightarrow} \hat{\beta}^{adj} = \hat{\beta} - \frac{cov(\xi, u)}{var(\xi)}[\hat{\phi} - 1]$$

   - Unpredictability of returns strongly rejected since $\hat{\phi} \approx 1$, and $\hat{V}(\hat{\beta}^{adj}) < \hat{V}(\hat{\beta})$ (US).

2. Cochrane (2008) takes the Campbell-Shiller approximation...

$$\begin{aligned} r_{t+1} &\approx k + \rho p_{t+1} + (1-\rho) d_{t+1} - p_t \\ &= k - \rho(d_{t+1} - p_{t+1}) + d_{t+1} - d_t + d_t - p_t \\ &= k - \rho(d_{t+1} - p_{t+1}) + \Delta d_{t+1} + (d_t - p_t). \end{aligned}$$

   - Then, regresses $r_{t+1}, d_{t+1} - p_{t+1}$, and Δd_{t+1} onto $d_t - p_t$ to get coefficients β, ϕ and β_d, which are linked by
   $$\beta = 1 + \beta_d - \rho\phi.$$
   - Suppose that $\rho = 0.96$ and $\phi \leq 1$. If $\beta = 0$, then $\beta_d = \beta - 1 + \rho\phi < 0$. However, the fact that β_d based on historical US data is close to zero implicitly means that $\beta > 0$.

3. Campbell and Yogo (2006) add $\xi_{t+1} = x_{t+1} - \phi x_t$ to the predictive regression.