- Last time: We log-linearized a simple accounting identity to obtain:

$$d_t - p_t = K - \sum_{i=0}^{\infty} \rho^i (1-\rho) \Delta d_{t+1+i} + \sum_{i=0}^{\infty} \rho^i r_{t+1+i}$$

- This is a very, very useful relation, because
  - The log dividend price ratio $(d_t - p)$ is stationary
  - Δd_{t+1+i} is stationary
  - r_{t+1+i} is stationary
  - The log dividend price ratio must forecast either future dividend growth or future returns
  - This is the relationship we had obtained a few slides ago, but now it is in a "regression" form.
  - We can use regressions to test this relation.
- The last observation is that we are dealing with future values. But this has never stopped us.

$$d_t - p_t = K - E_t \left\{ \sum_{i=0}^{\infty} \rho^i (1-\rho) \Delta d_{t+1+i} + \sum_{i=0}^{\infty} \rho^i r_{t+1+i} \right\}$$

Testing the above relationship:

- Regress: r_{t+1} on $d_t - p_t$
- Regress: Δd_{t+1} on $d_t - p_t$
- IMPORTANT: $d_t - p_t$ is the exogenous process.
- Which relationship is stronger?
- Note: This is a forecasting relationship: the dividend price ratio today helps us forecast tomorrow's (expected) returns
- It is different from the CAPM, APT, where we were trying to explain ex-post variations in returns.
- Here, we have a forecasting regression.

2 Long-Horizon Regressions:

People have noticed that the relationship

$$d_t - p_t = K - \sum_{i=0}^{\infty} \rho^i (1-\rho) \Delta d_{t+1+i} + \sum_{i=0}^{\infty} \rho^i r_{t+1+i}$$

might not hold in the short run, but it has got to hold in the long run.

- The above observation has prompted people to run regressions:
  - $r_{t+1} + ... + r_{t+k} = \alpha_1 + \beta_1 (d_t - p_t) + u_{t+k}$
  - $\Delta d_{t+1} + ... + \Delta d_{t+k} = \alpha_2 + \beta_2 (d_t - p_t) + e_{t+k}$
- One of the two regressions must be true, by definition.
- We have not used anything, but the definition of returns, no-bubbles condition, and a log-linearization.
- Q: What restrictions must be imposed on β_1 and β_2?
- The above restriction is never used in practice. How can it be implemented?
- The above regressions are never run at the dissagregated level.

- Good research topic:
  - Form portfolios on some characteristics
  - Test the above model as in Campbell and Shiller (1988, 1989, 1991, etc)
  - See if there is a pattern.

- Remark: As in the variance ratio tests, it is important to run the long horizon regressions using non-overlapping returns.
- If we use overlapping returns, we get spurious results.
- Correcting for the overlap is difficult and introduces another layer of complication.
- As a general rule, it is always good to stay away from overlapping returns regressions.
- It is also good to treat other people's overlapping returns results with a healthly dose of skepticism.
- Why? Because, most of the results are spurious, i.e. the product of the overlap, even though an underlying economic relation does not exist.

For concreteness, consider the short-horizon relation

$$\begin{aligned} Y_{t+1} &= \alpha + \beta X_t + \varepsilon_{1,t+1} \\ X_{t+1} &= \mu + \phi X_t + \varepsilon_{2,t+1} \end{aligned}$$

Suppose that under the null $\beta = 0$ and $\varepsilon_{1,t+1}$ and $\varepsilon_{2,t+1}$ are uncorrelated at all leads and lags.

- Suppose we form the long-horizon variables:

$$\begin{aligned} Z_t^k &= \sum_{i=0}^{k-1} Y_{t+i} \\ Q_t^k &= \sum_{i=0}^{k-1} X_{t+i} \end{aligned}$$

- Long horizon variables are run mostly in two fashions:

$$\begin{aligned} Z_{t+1}^k &= \gamma + \varphi X_t + u_{t+1} \\ Z_{t+1}^k &= \delta + \kappa Q_t^k + u_{t+1} \end{aligned}$$

- Note: The returns-dividend yield regressions discussed above are an example of the first kind (Campbell and Shiller (1988), etc.)

- Note: The second type of regressions is used in macroeconomics (tests of Fisher relation, etc. see Richardson (1993), Mishkin (1992), Fisher and Seater (1993))

- Why are those regressions problematic?

- Spurious Regression Result (Proven in Phillips (1986)):
  - Suppose that $y_t = y_{t-1} + \varepsilon_t = \sum_{j=0}^{\infty} \varepsilon_{t-j}$ and $x_t = x_{t-1} + u_t = \sum_{j=0}^{\infty} u_{t-j}$ and ε_t and u_t are mean zero and independent at all leads and lags.
  - Then, y_t and x_t are independent.
  - However, if we run the regression

    $$y_t = \gamma x_t + v_t$$

    the ols estimate $\hat{\gamma}$ wil not converge in probability to 0. Moreover its $t - statistic$ will diverge, producing spuriously significant results and the corresponding R^2 will also be very large.
  - All these results can be proven using a simple application of the Functional Central Limit Theorem (See Phillips (1986) and Hamilton (1996))

– For instance, using the FCLT from last lecture:

$$\begin{aligned}
\hat{\gamma} &= \frac{\sum_{t=1}^{T} y_t x_t}{\sum_{t=1}^{T} x_t^2} = \frac{\sum_{t=1}^{T}\left(\sum_{j=0}^{\infty}\varepsilon_{t-j}\right)\left(\sum_{j=0}^{\infty}u_{t-j}\right)}{\sum_{t=1}^{T}\left(\sum_{j=0}^{\infty}u_{t-j}\right)^2} \\
&= \frac{\frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{T^{1/2}}\sum_{j=0}^{\infty}\varepsilon_{t-j}\right)\left(\frac{1}{T^{1/2}}\sum_{j=0}^{\infty}u_{t-j}\right)}{\frac{1}{T^2}\sum_{t=1}^{T}\left(\sum_{j=0}^{\infty}u_{t-j}\right)^2} \\
&\Rightarrow \frac{\int_0^1 W_y(s)W_x(s)ds}{\int_0^1 W_x^2(s)ds}
\end{aligned}$$

– Hence, $\hat{\gamma}$ converges in distribution, but it does not converge in probability to 0 as we might have expected.

– Similar calculations might be applied to show that

$$t = \frac{\hat{\gamma} - 0}{se(\hat{\gamma})}$$

diverges at rate $T^{1/2}$. In other words, we have to rescale the t-statistic by $T^{1/2}$ to obtain a convergence in distribution.

– In small samples, these results imply that we will get large $\hat{\gamma}$ and $t - statistics$.

- Now, the long horizon regressions discussed above are very similar to the spurious regression setup. Valkanov (2003) shows that in small sample, these long-horizon regressions exhibit the same spuriousness symptoms, namely:
  - large γ, growing with the horizon
  - significant t, growing with the horizon
  - large R^2
- Richardson and Stock (1989) have shown that the same is true for the Variance Ratio (VR) test. In other words, if we use overlapping observations, the VR test will seem very large.