1 Advanced Empirical Asset Pricing

- Professor Rossen Valkanov (UCSD)
  Contact Info: rvalkanov@ucsd.edu
- Goals
  - Review important papers
  - Discuss important tools
  - Outline trends and topics for research
- Class Overview
  - Class 1: Predictive Regressions
    - Preliminaries
    - Short horizon predictive regressions
  - Class 2: Predictive Regressions
    - Time-varying expected returns
  - Class 3: Volatility Modeling
  - Class 4: New Methods

2 Predictive Regressions

A predictive (linear) regression is

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

where r_{t+1} is an asset return (mostly stock or bond) and x_t is a variable observable at time t (including lagged returns)

- Statistical issues
  - If predictors are persistent (borderline non-stationary)
  - If predictors are noisy proxy of expected returns
  - Spurious regressions
  - Horizon and overlap
  - Model selection and combinations
  - Stability of relationship
  - In-sample vs out-of-sample forecasts
- Economic issues
  - What are we testing: market efficiency or time-varying expected returns?
  - What predictive accuracy is economically reasonable?
  - Portfolio gains from predictability?
  - Economic constraints

2.1 Preliminaries–Conditional Means

Think of a variable y_{t+1} that is observable. We can always write it as

$$\begin{aligned} y_{t+1} &= E\left(y_{t+1}|I_t\right) + y_{t+1} - E\left(y_{t+1}|I_t\right) \\ &= E\left(y_{t+1}|I_t\right) + \varepsilon_{t+1} \end{aligned}$$

as long as $E\left(y_{t+1}|I_t\right)$ exists. This is a conditional expectation, where the conditioning is w.r.t. information available up to time t.

- $E\left(y_{t+1}|I_t\right)$ is what we will "model"
- ε_{t+1} is the "unexplained" or "unexpected" part.

  By construction, ε_{t+1} is orthogonal to I_t, i.e. $E\left(\varepsilon_{t+1}|I_t\right) = 0$.

  Interpretation: Nothing in I_t will help us forecast the conditional mean of ε_{t+1}

  Interpretation: Past information cannot be used to forecast the truly unexpected part of y_{t+1}.
- Why is this useful in finance?

Think of y_{t+1} as the log return of an asset, r_{t+1}. To be clear

- $P_t =$price of an asset at time t
- $D_{t+1} =$dividend the asset pays between t and $t+1$.
- The simple **gross** return is defined as

$$\begin{aligned} P_t \left(1 + R_{t+1}\right) &= P_{t+1} + D_{t+1} \\ 1 + R_{t+1} &= \frac{P_{t+1}}{P_t} + \frac{D_{t+1}}{P_{t+1}} \end{aligned}$$

- The simple (**net**) return is defined as

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

- We will work almost exclusively with a third quantity, the **log return**, defined as:

$$\ln\left(1 + R_{t+1}\right) = r_{t+1}$$

  Note that r_{t+1} is "close to" R_t, for values of R_t close to zero (Taylor expansion argument). This tends to be true if we consider short enough periods, t.

- Also, the **long-horizon** log return between t and $t+k$ is the sum of the one-period returns

$$\sum_{i=1}^{K} r_{t+i} = r_{t,t+k}$$

- We can write the log return r_{t+1} as

$$r_{t+1} = E\left(r_{t+1}|I_t\right) + \varepsilon_{t+1}$$

and the question is how to model $E\left(r_{t+1}|I_t\right)$. We will sometime use the shorthand notation

$$r_{t+1} = \mu_t + \varepsilon_{t+1}$$

- The simplest form for $E\left(r_{t+1}|I_t\right) = \alpha + \beta x_t$ where x_t is some "state variable" or "predictor".
  - state variable: captures the state of the economy
  - predictor: helps predict the variable of interest
  - Synonymous in most empirical contexts
  - Natural candidates: the dividend price ratio (see above slide), the short rate, the term spread, the default spread, etc.
- Hence, we are considering the linear predictive regression:

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

- Notice that the realized return, r_{t+1}, is different from the expected return, μ_t
- In asset pricing theory, we model expected returns. Indeed, that's what you do with the CAPM, APT (unconditional) and ICAPM and SDF (conditional)
- But in practice, we don't have expected returns.
- We have realized returns!
- If we write

$$r_{t+1} = \mu_t + \varepsilon_{t+1}$$

  there are two sources of time variation μ_t and ε_{t+1}
- Focus: μ_t is about time varying expected returns
- Focus: ε_{t+1} is about market efficiency

2.2 Preliminaries–Estimation

- Recall distinction between modeling and estimation
- The linear relation between r_{t+1} and x_t can be estimated in a number of ways:

  Least Squares

  Maximum Likelihood

  GMM

  In the linear case, we have (assuming $\alpha = 0$, to simplify notation)

$$\begin{aligned} E\left(\varepsilon_{t+1} x_t\right) &= 0 \\ E\left(\left(r_{t+1} - \beta x_t\right) x_t\right) &= 0 \\ E\left(r_{t+1} x_t\right) &= \beta E\left(x_t^2\right) \\ \beta &= \frac{E\left(r_{t+1} x_t\right)}{E\left(x_t^2\right)} \end{aligned}$$

  To estimate β,we use sample analogues to the moments above

$$\begin{aligned} \hat{\beta} &= \frac{\frac{1}{T}\sum_{t=1}^{T} r_{t+1} x_t}{\frac{1}{T}\sum_{t=1}^{T} x_t^2} = \frac{\frac{1}{T}\sum_{t=1}^{T}\left(\beta x_t + \varepsilon_{t+1}\right) x_t}{\frac{1}{T}\sum_{t=1}^{T} x_t^2} \\ &= \beta + \frac{\frac{1}{T}\sum_{t=1}^{T} \varepsilon_{t+1} x_t}{\frac{1}{T}\sum_{t=1}^{T} x_t^2} \end{aligned}$$

- To estimate $\hat{\beta}$ consistently, i.e. $\hat{\beta} \rightarrow^p \beta$, we need that
  $\frac{1}{T}\sum_{t=1}^{T} \varepsilon_{t+1} x_t \rightarrow^p E(\varepsilon_{t+1} x_t) = 0$
  $\frac{1}{T}\sum_{t=1}^{T} x_t^2 \rightarrow^p Q < \infty$
- Recall the difference between consistency, biasedness, and efficiency
- Under a lot of restrictive assumptions, $\hat{\beta}$ is BLUE (best linear unbiased), but those conditions are not met in finance!
- If we don't have a consistent estimator, not good!
- To have a consistent estimator, we only need that
$$E(\varepsilon_{t+1} x_t) = 0$$
- Pretty simple!

- One omitted detail: For the sums $\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{t+1}x_t$ and $\frac{1}{T}\sum_{t=1}^{T}x_t^2$ to converge, we need the processes to be ergotic (stationary with some added conditions). You can think of this as LLN for time series data (potentially dependent)
- If stationarity is not satisfied, bad things can happen to us!
- Good rule: When in doubt, make sure you are working with stationary data.

- Back to our predictive regression

$$\begin{aligned} r_{t+1} &= \mu_t + \varepsilon_{t+1} \\ &= \alpha + \beta x_t + \varepsilon_{t+1} \end{aligned}$$

- Now, we know that realized returns are close to i.i.d (hopefully not totally i.i.d, otherwise the predictive regression will be pointless)
  - Aside: How can we reconcile this expression with market efficiency (Fama (1970, JF))?
  - "Predictive regressions" means different things to different people–lots of confusion
- How about the stochastic properties of the predictor x_t?
- In practice x_t is very persistent, borderline non-stationary
  - Dividend price ratio: close to non-stationary
  - Short rate (3 month Tbill): close to non-stationary
  - Term spread: close to non-stationary
  - ...

- What do I mean by "close to non-stationary"?
- If we assume a simple AR(1) process for x_t

$$x_{t+1} = \phi x_t + u_{t+1}$$

a sufficient condition for this process to be stationary (covariance stationary) is that $|\phi| < 1$

- Since $E(x_{t-1}u_t) = 0$ by construction, we expect $\hat{\phi} \to^p \phi$
- Following the above calculations

$$\sqrt{T}\left(\hat{\phi} - \phi\right) \to^d N\left(0, \frac{\sigma^2}{\sigma^2/\left(1-\phi^2\right)}\right) = N\left(0, \left(1-\phi^2\right)\right)$$

As $T \to \infty$, $E\left(\hat{\phi}\right)$ closer to ϕ

As $T \to \infty$, $Var\left(\hat{\phi}\right)$ closer to 0.

- Notice what happens when $\phi \to 1$.
- What if $\phi = 0.997$?

Two Approaches:

- First Method: Kendall (1954) and Marriott and Pope (1954) [used by Stambaugh (1986)] have shown that under normality, the bias is approximately:

$$E\left(\hat{\phi}\right) = \phi - \frac{1+3\phi}{T} + O\left(1/T^2\right)$$

- Hence, If we get an estimate of, say, 0.9, with T=100, it would be downward biased.

- We can obtain a (sort of) first-order unbiased estimate using

$$\begin{aligned}\hat{\phi}^u &= 0. + \frac{1+3*0.9}{100} \\ &= 0.9 + 0.037 \\ &= 0.937\end{aligned}$$

- Comment: Can't fully correct for the bias, because it is a function of ϕ

- Comment: This is a small-sample issue. If we have enough observations, the bias will disappear

- Comment: Can we assume $\phi = 1$ in the correction? [Lewellen (2004)]

- However, we usually have a fixed number of $T's.$

- Even if we had more $T's,$we can also find a ϕ closer to 1 that would create problems!

- Second Method: Median Unbiased Estimator (MUE)
- Idea: If we simulate

$$Y_t = c + \phi Y_{t-1} + \varepsilon_t$$

for a given T, say 5000 times and for a grid of $\phi = [0, 1]$.

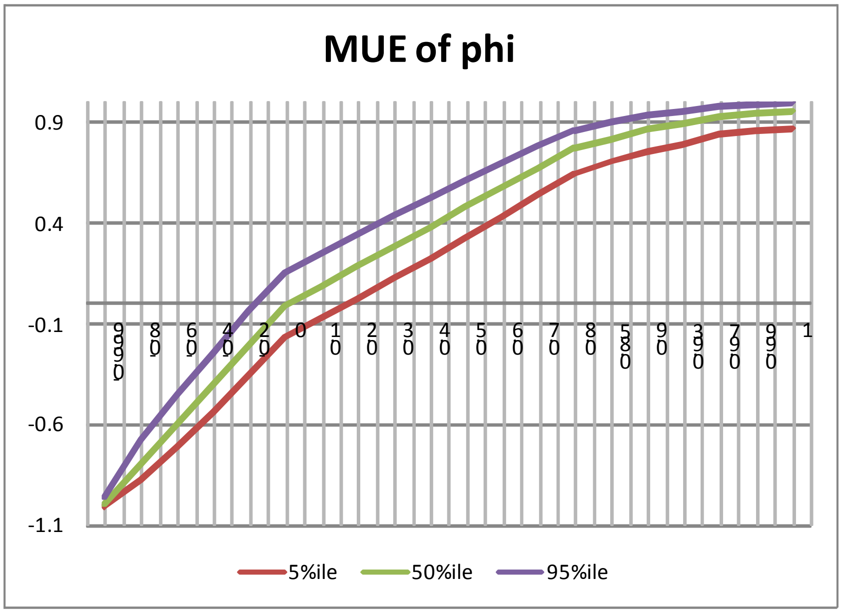
- Estimate $\hat{\phi}$ for all 5000 simulations.
- Then we plot the histograms of $\hat{\phi}$ as a function of ϕ
- Suppose $m_T(\phi)$ is the median of $\hat{\phi}$, when the true parameter is ϕ.
- We can invert the median function:

$$\phi^u = m^{-1}\left(\hat{\phi}\right)$$

- Ex: Andrews (Econometrica, 1993)
- Table II

	$T+1=90$			$T+1=100$			$T+1=125$			$T+1=150$			$T+1=200$		
α/Quantile	.05	.5	.95	.05	.5	.95	.05	.5	.95	.05	.5	.95	.05	.5	.95
−.999	−1.007	−.997	−.958	−1.006	−.997	−.961	−1.005	−.997	−.966	−1.004	−.997	−.970	−1.003	−.998	−.976
−.80	−.879	−.794	−.663	−.876	−.794	−.672	−.869	−.795	−.688	−.865	−.796	−.700	−.858	−.797	−.715
−.60	−.720	−.598	−.441	−.715	−.598	−.450	−.704	−.599	−.468	−.696	−.599	−.480	−.684	−.599	−.498
−.40	−.549	−.402	−.233	−.542	−.402	−.242	−.528	−.402	−.259	−.517	−.401	−.272	−.502	−.401	−.289
−.20	−.370	−.207	−.032	−.361	−.206	−.041	−.344	−.205	−.057	−.332	−.204	−.069	−.314	−.203	−.087
.00	−.184	−.011	.161	−.174	−.010	.154	−.155	−.008	.139	−.141	−.007	.127	−.121	−.005	.111
.10	−.088	.086	.256	−.078	.088	.249	−.058	.090	.234	−.043	.092	.224	−.023	.094	.208
.20	.008	.184	.349	.019	.186	.342	.040	.189	.329	.055	.191	.319	.076	.193	.304
.30	.107	.282	.440	.118	.284	.434	.140	.287	.422	.155	.289	.413	.177	.292	.399
.40	.207	.380	.530	.219	.382	.524	.241	.385	.513	.257	.388	.505	.278	.391	.493
.50	.309	.477	.617	.321	.480	.612	.343	.484	.603	.359	.486	.596	.381	.490	.585
.60	.413	.575	.702	.425	.577	.699	.448	.582	.691	.464	.585	.685	.486	.589	.676
.70	.519	.672	.785	.532	.675	.782	.555	.680	.776	.571	.684	.771	.593	.688	.764
.80	.628	.769	.863	.641	.773	.862	.664	.778	.858	.681	.782	.855	.702	.787	.850
.85	.684	.818	.901	.697	.821	.900	.721	.827	.897	.737	.831	.895	.758	.836	.891
.90	.741	.865	.936	.754	.869	.936	.778	.876	.934	.794	.880	.933	.815	.885	.931
.93	.775	.893	.956	.788	.897	.956	.813	.904	.955	.829	.909	.955	.850	.915	.953
.97	.820	.929	.981	.834	.933	.981	.859	.941	.981	.876	.947	.981	.898	.953	.981
.99	.840	.945	.994	.854	.950	.994	.881	.958	.994	.898	.964	.994	.921	.971	.994
1.0	.849	.952	.999	.863	.957	.999	.890	.965	.999	.908	.971	.999	.931	.978	.999

• For T=100, the picture is:



• For instance, if the OLS estimate of ϕ is $\hat{\phi} = 0.9$,then $\phi^{MUE} = 0.93$

• Similar to Kendall's correction.

• A couple of remarks on MUE estimators

As T increases, we don't get better estimates

We need to assume the distribution of ε_t (in the simulation)

Stock (JME, 1991) has an asymptotic version of the MUE estimator

- Simulation 1: Downward bias in $\hat{\phi}$
- The financial econometrics literature has found another way to deal with this problem: local-to-unity (LTU) process

$$\begin{aligned} x_{t+1} &= \phi x_t + u_{t+1} \\ \phi &= 1 + \frac{c}{T} \end{aligned}$$

- In the LTU process, as $T \to \infty$, $\phi \to 1$, so our test will never be consistent against $c \neq 0$ alternatives.
- What's the big deal with non-stationary data?

2.3 Spurious Regressions and Non-Stationary Time Series

- The easiest way is to work with stationary time series: r_t and (later) σ_t.
- Why? LLN (Ergoticity) and CLT are valid under stationarity
- But suppose you want to work with non-stationary time-series, i.e. prices, volume, number of investors in a particular fund, number of funds, etc. Those processes are inherently non-stationary.
- Let p_t be the log-price. We know that

$$p_t = p_{t-1} + \varepsilon_t$$

  or p_t is an AR(1) process with an unit-root.
- This process is non-stationary. We cannot apply the CLT.
- But we are still interested in testing the null $\phi = 1$ versus $\phi < 1$.
- Problem: Under the null, the process is non-stationary.
- Under the alternative, the process is stationary.

- It turns out that (Funtional CLT)

$$\frac{1}{\sqrt{T}}p_t = \frac{1}{\sqrt{T}}\sum_{s=1}^{t}\varepsilon_s$$
$$= \frac{1}{\sqrt{T}}\sum_{s=1}^{[rT]}\varepsilon_s \Rightarrow W(r)$$

  where $W(r)$ is a Brownian motion, $t = [rT]^{-}$, on $[0,1]$. I.e., $W(r) \sim N(0,r)$, $0 < r < 1.$

- Note that this result is asymptotic. We don't have to assume that the $\varepsilon' s$ are normal (hence the Functional CLT, or FCLT).

- Q: Can't we standardize the non-stationary processes by a power of T in order for them to converge.

- A: Yes.

- This is known as "fill-in" asymptotics
  As $T \to \infty$, the interval gets closer

- Let's get a "□avor" of how things work:

- Recall that

$$\hat{\phi} = \frac{\sum p_t p_{t-1}}{\sum p_{t-1}^2} = \frac{\sum p_{t-1}\left(\phi p_{t-1} + \varepsilon_t\right)}{\sum p_{t-1}^2}$$
$$= \phi + \frac{\sum p_{t-1}\varepsilon_t}{\sum p_{t-1}^2}$$

- If $\phi < 1$, we had

$$\hat{\phi} = \phi + \frac{\frac{1}{T}\sum \varepsilon_t p_{t-1}}{\frac{1}{T}\sum p_{t-1}^2} \to^p \phi$$
$$\sqrt{T}\left(\hat{\phi} - \phi\right) \sim N\left(0, \sigma_{\hat{\phi}}^2\right)$$

- But if $\phi = 1$, the results do not hold. But

$$\hat{\phi} = \phi + \frac{\sum \varepsilon_t p_{t-1}}{\sum p_{t-1}^2}$$
$$T\left(\hat{\phi} - \phi\right) \Rightarrow O_p(1)$$

- In other words, $\hat{\phi}$ is super-consistent.
- Q: But since we don't know the distribution of $\hat{\phi}$, can we use this result for testing?
- A: Yes, if we simulate the distribution.

- Dickey-Fuller (DF) Test:
  $H_o : \phi = 1$
  $H_a : \phi < 1$
- The test is: $t = \frac{\hat{\phi}-1}{se\left(\hat{\phi}\right)}$
- Suppose that ε_t follows an AR(p) process. The distribution of the DF test is in□uenced by these (nuisance0 parameters. Not good!
- To get rid of these parameters, we run the following regression:

$$p_t = \phi p_{t-1} + \zeta_1 \Delta p_{t-1} + \zeta_2 \Delta p_{t-2} + ... + \zeta_k \Delta p_{t-k} + \upsilon_t$$

- Then, $t = \frac{\hat{\phi}-1}{se\left(\hat{\phi}\right)}$.
- This is called the Augmented DF,or ADF test.

- Summary of ADF test–Testing for a unit root:

  Regress p_t on $p_{t-1}, \Delta p_{t-1}, \Delta p_{t-2}, ..., \Delta p_{t=k}$

  Form: $t = \frac{\hat{\phi}-1}{se\left(\hat{\phi}\right)}$

  Get the critical value from simulations.

- So, is working with non-stationary variables that easy?
- NO:
- Suppose p_t^1 and p_t^2 are the prices of the same asset traded on two markets. Then, it must be the case that

$$p_t^1 = p_t^2$$

- Empirically, this is almost true. We find

$$p_t^1 - p_t^2 = \varepsilon_t$$

  where ε_t is almost i.i.d., and $E(\varepsilon_t) = 0$.

- How do we take advantage of this?
- Regress:

$$p_t^1 = \gamma p_t^2 + \varepsilon_t$$

- If

$$\begin{aligned} \varepsilon_t &> 0 \\ p_t^1 &> \gamma p_t^2 \end{aligned}$$

- The asset in market 1 is "too expensive"
- For best way to exploit convergence trades (or mispricings) using cointegration: Liu and Timmermann (RFS, 2013)
- Easy, right?

- Wrong!
- Suppose we have two assets, with prices p_t and q_t. One might be tempted to look for arbitrage strategies as in

$$p_t = \gamma q_t + \varepsilon_t$$

- If $\gamma > 0$, there is a relationship, and we can trade.
- No.
- Suppose

$$\begin{aligned} p_t &= p_{t-1} + u_t \\ q_t &= q_{t-1} + v_t \\ cov\left(u_t v_t\right) &= 0 \end{aligned}$$

- Note: The two log-prices represent two *independent* discrete Brownian motions.

- But:

$$\hat{\gamma} = \frac{\sum_{t=1}^{T} p_t q_t}{\sum_{t=1}^{T} q_t^2} = \frac{\sum_{t=1}^{T} \left(\sum_{s=1}^{t} v_s\right) \left(\sum_{s=1}^{t} u_s\right)}{\sum_{t=1}^{T} \left(\sum_{s=1}^{t} v_s\right)^2}$$

$$= \frac{\frac{1}{T^2} \sum_{t=1}^{T} \left(\sum_{s=1}^{t} v_s\right) \left(\sum_{s=1}^{t} u_s\right)}{\frac{1}{T^2} \sum_{t=1}^{T} \left(\sum_{s=1}^{t} v_s\right)^2} = O_p(1)$$

- Similarly:

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

  is not consistent.

- The R^2 does not converge to $0, as\ T \to \infty$.

- Illustration: **spurious.m**

Co-integration:

- Suppose that p_t and q_t are unit root (integrated) processes, but there is a linear combination of p_t and q_t that is stationary. That is, there exists a vector $\gamma = \left[1 \ -\gamma_1 \right]$ such that

$$p_t - \gamma q_t = \varepsilon_t$$

  and ε_t is a stationary process. Then, p_t and q_t are said to be cointegrated.

- There are formal tests for cointegration, but they have low power against the alternative. Why? (Think spurious correlation).

- Cointegration is only between contemporaneous variables. I.e. $p_t - \gamma q_t$ is a cointegrating vector. But $\triangle p_t$ is not. However, the latter is also a way of stationarizing the process.

- Cointegration occurs “naturally” in economics. It is dictated by theory.

- Examples:

  Prices p_t^1 and p_t^2 of the same asset traded on two markets.

  Dividend price ratio is cointegrated, or:

  $$d_t - p_t$$

  must be stationary

  The long and the short rate must be cointegrated

  Consumption and GDP must be cointegrated, etc.

  Lettau & Ludvigson (2001), the “cay” ratio:

  $$c_t - \gamma_1 a_t - \gamma_2 y_t$$

2.4 Variance Ratio Tests and Market Efficiency

The starting point is

$$r_{t+1} = \mu_t + \varepsilon_{t+1}$$

- Let's assume $\mu_t = 0$ (i.e., daily returns)
- The variation in $r_{t_{+1}}$ comes from ε_{t+1}
- Suppose ε_{t+1} is iid (N?), then...
- Suppose ε_{t+1} is indepedent (relaxing identically distributed)
- Suppose ε_{t+1} is serially uncorrelated (weak-form)
  Various tests of weak-form efficiency...Can you enumerate a couple?
- Suppose ε_{t+1} is uncorrelated with other observable (to market participants) variables at time t (semi-strong form)
- Horizon is an issue
- At short horizons, we don't run predictive regressions. We work only with returns. Why?

- We want to estimate see if r_{t+1} is uncorrelated with past information, or

$$E\left(r_{t+1}x_t\right) = 0$$

- Suppose $r_{t+1} = \mu + \varepsilon_{t+1}$, then $E\left(r_{t+1}x_t\right) = E\left(\varepsilon_{t+1}x_t\right) = 0$ [null]
- Alternatives can be:

  $r_{t+1} = \phi r_t + \varepsilon_{t+1}$ (AR(1))

  $r_{t+1} = \phi_1 r_t + \phi_2 r_{t-1} + \varepsilon_{t+1}$ (AR(2))

  $r_{t+1} = \varepsilon_{t+1} + \theta\varepsilon_t$

  $r_{t+1} = \varepsilon_{t+1} + \theta_1\varepsilon_t + \theta_2\varepsilon_{t-1}$ (MA(2))

  $r_{t+1} = \phi_1 r_t + \varepsilon_{t+1} + \theta_1\varepsilon_t$ (ARMA(1,1))

  We have to estimate all those models to eliminate them against the null!

- The variance ratio test is a neat and "natural" way to test market efficiency.
- Work with log returns, r_t, which **under the null** are unforecastable, uncorrelated (or independent, or iid)
- Here is a useful observation about the variance of a two period (log) return $r_{t,t+2}$ [under the null]

$$\begin{aligned} Var\left(r_{t,t+2}\right) &= Var\left(r_{t+1} + r_{t+2}\right) \\ &= Var\left(r_{t+1}\right) + Var\left(r_{t+2}\right) \\ &= 2\sigma_r^2 \end{aligned}$$

- In general, for the q-th period (long-horizon) return $r_{t,t+q}$:

$$\begin{aligned} Var\left(r_{t,t+q}\right) &= Var\left(r_{t+1} + ... + r_{t+q}\right) \\ &= q\sigma_r^2 \end{aligned}$$

- Those implications were derived under the hypothesis (the null) that returns are unforecastable.

- So, we have a "natural" test for the Random Walk model (or equivalently, for unforecastable returns):
  $$Var\left(r_{t,t+q}\right) \overset{Under\ the\ Null\ of\ a\ Random\ Walk}{=} qVar\left(r_{t+1}\right)$$
  or
  $$VR(q) = \frac{Var\left(r_{t,t+q}\right)}{qVar\left(r_{t+1}\right)} = 1$$
- This is a simple test.
- Compute the variance of the q-period returns.
- Compute the variance of the 1-period returns.
- It must be the case that the ratio of $\frac{Var(r_{t,t+q})}{qVar(r_{t+1})}$ must be close to 1.
- Q: How close? We should not forget that the test $VR(q)$ is a random variable with a corresponding density, etc.
- Q: For what values of $VR(q)$ is the model a RW and for what values it is not?
- ASIDE: We can relate the $VR(q)$ test to autocorrelations (see CLM, 1997).
- ASIDE: What q to choose?

- So, under the null hypothesis that r_t are i.i.d., we can use the CLT to show that

$$\sqrt{Tq}\left(\widehat{VR(q)}-1\right)\sim^{a} N\left(0,2\left(q-1\right)\right)$$

(recall last lecture and the results about $\overline{X}$ and $\hat{\beta}$)

or

$$\left(\widehat{VR(q)}-1\right)\sim^{a} N\left(0,\frac{2\left(q-1\right)}{Tq}\right)$$

- Therefore, we can conduct testing in the usual way: Form a statistic

$$Z=\frac{\widehat{VR(q)}-1}{\sqrt{\frac{2(q-1)}{Tq}}}\sim^{a} N\left(0,1\right)$$

- So, if Z is greater than 1.96, we reject the null of IID returns at the 5% level.

- Note: We must be careful when computing the long-horizon returns $r_{t,t+q}$ from overlapping observations.
- Q: Why?
- A: Because, under the null hypothesis, $r_{t,t+q}$ are i.i.d. If we use overlapping observations, they will NOT be i.i.d. by construction.
- Some people advocate the use of overlapping observations and correcting for correlation from the overlap as:

$$\sqrt{Tq}\left(\widehat{VR^{overlap}}(q)-1\right) \sim^{a} N\left(0, \frac{2(2q-1)(q-1)}{3q}\right)$$

- The correction of the variance is supposed to correct for the overlap. At the same time we have more observations, so the test might be "better", or more powerful.
- People are deluding themselves!
- If we correct exactly for the overlap, whether we use overlap or not would make no difference. But the uncertainty introduced around how to deal with the overlap makes the second test less desirable.
- At the end of the day, we have the same amount of information (returns), not more.
- We will make that more formal later!

3 Predictive Regressions and Market Efficiency: Short Horizon Predictability

3.1 Short Horizon Predictability: Recent Papers

- Chordia, Roll, and Subrahmanyam, 2005. "Evidence on the speed of convergence to market efficiency," JFE.

  Chordia, Roll, and Subrahmanyam, 2008. "Liquidity and market efficiency," JFE.

- Paper: Evidence on the speed of convergence to market efficiency

  Q: Market is efficient at daily horizon. How about at shorter horizons?

  Premise: "Investors need time to absorb and act on new information". Speed of information impounding into prices

  Link between market efficiency and liquidity: "...more liquid should exhibit less pronounced return predictability and vice versa"

Interval choice: r_t is 5-minute return (somewhat arbitrary, but...)

Firm level returns (150 large firms)

Sample: 1993-2002 (TAQ database)

- ∗ Determine Buyer vs Seller Initiated Trades: Matching trades to quotes...Lee and Ready(1991) algorithm.
- ∗ OIB: Number of buyer minus seller initiated trades over a period of time for a given stock, as a fraction of all trades.
- ∗ Daily Stats: (Table 1)

	Trade $Return_t$	Midpoint $Return_t$	$OIB\#_t$	$OIBSh_t$	$OIB\$_t$	Trade $Return_t$	Midpoint $Return_t$	$OIB\#_t$	$OIBSh_t$	$OIB\$_t$	Trade $Return_t$	Midpoint $Return_t$	$OIB\#_t$	$OIBSh_t$	$OIB\$_t$
			1996					1999					2002		
$Return_{t-1}$[a]	0.002	0.007				−0.000	0.010				−0.030	−0.027			
	(0.38)	(1.15)				(−0.03)	(1.30)				(−4.98)	(−4.56)			
$OIB\#_t$	0.315	0.314				0.215	0.214				0.232	0.230			
	(22.32)	(22.22)				(12.18)	(11.98)				(13.75)	(13.52)			
$OIB\#_{t-1}$	0.055	0.056	0.336			0.019	0.022	0.371			0.022	0.022	0.249		
	(9.64)	(9.79)	(19.77)			(3.07)	(3.68)	(27.65)			(3.53)	(3.43)	(18.05)		
$OIBSh_t$	0.454	0.460	0.305			0.462	0.468	0.279			0.327	0.326	0.503		
	(40.80)	(41.12)	(21.91)			(37.21)	(37.28)	(16.32)			(25.82)	(25.56)	(42.82)		
$OIBSh_{t-1}$	0.002	−0.002	−0.055	0.125		−0.016	−0.014	−0.018	0.204		0.006	0.006	0.088	0.179	
	(0.31)	(−0.38)	(−5.20)	(14.09)		(−2.53)	(−2.16)	(−1.53)	(18.46)		(0.93)	(1.00)	(8.54)	(16.84)	
$OIB\$_t$	0.452	0.458	0.303	0.991		0.457	0.462	0.280	0.981		0.321	0.321	0.488	0.965	
	(41.34)	(41.57)	(22.04)	(609.3)		(37.13)	(37.14)	(16.64)	(437.8)		(27.67)	(27.43)	(44.18)	(201.4)	
$OIB\$_{t-1}$	0.000	−0.003	−0.057	0.124	0.128	−0.020	−0.018	−0.013	0.197	0.209	−0.007	−0.006	0.075	0.161	0.180
	(0.08)	(−0.63)	(−5.49)	(14.09)	(14.22)	(−3.18)	(−2.85)	(−1.09)	(18.37)	(18.39)	(−1.20)	(−1.11)	(8.27)	(16.63)	(17.22)

Predictive Regressions: Cross-Sectional Univerate(Panel) Regressions

Table 2: Main Results

Explanatory variable	Return interval (minutes)				
	Five	Ten	Fifteen	Thirty	Sixty
			1996		
Midpoint	−0.003	−0.025	−0.027	0.010	0.022
Return_{t-1}	(−0.50)	(−2.52)	(−2.22)	(0.40)	(0.60)
	(−0.12)	(−0.98)	(−1.03)	(0.42)	(0.61)
	0.45	0.39	0.41	0.23	0.33
$\text{OIB\#}_{t-1}$	7.73	3.72	2.50	1.81	0.94
	(18.04)	(6.02)	(3.36)	(2.27)	(1.17)
	(3.91)	(2.58)	(2.17)	(2.23)	(0.83)
	2.61	0.79	0.44	0.28	0.22
$\text{OIB\$}_{t-1}$	21.40	12.03	9.04	8.90	4.07
	(11.91)	(4.64)	(2.88)	(1.90)	(1.20)
	(3.54)	(2.40)	(2.36)	(0.84)	(0.29)
	1.04	0.38	0.21	0.22	0.18
			1999		
Midpoint	0.016	0.010	0.008	0.009	0.021
Return_{t-1}	(2.15)	(1.01)	(0.66)	(0.49)	(0.76)
	(0.99)	(0.70)	(0.58)	(0.56)	(1.07)
	0.19	0.14	0.13	0.14	0.19
$\text{OIB\#}_{t-1}$	2.83	1.42	0.94	0.50	0.53
	(8.56)	(3.26)	(2.04)	(0.94)	(0.81)
	(2.78)	(2.13)	(1.88)	(0.98)	(0.97)
	0.56	0.20	0.12	0.10	0.15
$\text{OIB\$}_{t-1}$	7.51	4.15	3.18	2.43	2.65
	(4.91)	(2.01)	(1.45)	(0.92)	(0.79)
	(2.54)	(1.99)	(2.02)	(1.25)	(1.02)
	0.18	0.07	0.05	0.07	0.12

Multivariate Regressions: Main Results (Table 5)

Explanatory variable	Return interval (minutes)									
	Five		Ten		Fifteen		Thirty		Sixty	
Dependent variable is the midpoint return$_t$, 1996										
Midpoint	−0.069	−0.035	−0.070	−0.051	−0.063	−0.051	−0.014	−0.008	0.007	0.005
Return$_{t-1}$	(−7.93)	(−4.37)	(−5.66)	(−4.51)	(−4.13)	(−3.62)	(−0.68)	(−0.45)	(0.10)	(0.04)
	(−3.31)	(−1.34)	(−3.05)	(−1.86)	(−2.43)	(−1.79)	(−0.56)	(−0.30)	(0.50)	(0.14)
OIB#$_{t-1}$	8.92		4.88		3.49		2.03		0.84	
	(19.12)		(7.53)		(4.59)		(2.24)		(0.96)	
	(3.75)		(3.01)		(2.70)		(2.00)		(1.60)	
OIB\$$_{t-1}$		23.48		15.26		12.51		8.95		5.01
		(12.61)		(5.97)		(4.06)		(1.84)		(0.99)
		(3.64)		(2.71)		(2.77)		(1.09)		(0.49)
R^2	3.32	1.59	1.42	0.94	1.02	0.78	0.53	0.43	0.52	0.43
Dependent variable is the midpoint return$_t$, 1999										
Midpoint	−0.011	0.005	−0.005	0.004	−0.004	0.002	0.001	0.003	0.013	0.014
Return$_{t-1}$	(−1.44)	(0.68)	(−0.46)	(0.35)	(−0.26)	(0.13)	(0.08)	(0.12)	(0.43)	(0.45)
	(−0.76)	(0.31)	(−0.34)	(0.26)	(−0.25)	(0.13)	(0.09)	(0.17)	(0.63)	(0.66)
OIB#$_{t-1}$	3.03		1.52		1.07		0.56		0.46	
	(8.29)		(3.08)		(1.99)		(0.86)		(0.76)	
	(2.72)		(2.24)		(2.12)		(1.22)		(0.86)	
OIB\$$_{t-1}$		7.20		3.88		3.11		2.38		1.94
		(4.53)		(1.75)		(1.27)		(0.75)		(0.50)
		(2.48)		(1.91)		(1.87)		(1.24)		(0.69)
R^2	0.70	0.34	0.30	0.20	0.23	0.17	0.21	0.19	0.29	0.26
Dependent variable is the midpoint return$_t$, 2002										
Midpoint	−0.037	−0.021	−0.035	−0.029	−0.025	−0.021	−0.008	−0.007	−0.004	−0.002
Return$_{t-1}$	(−4.57)	(−2.75)	(−3.08)	(−2.67)	(−1.77)	(−1.58)	(−0.38)	(−0.37)	(−0.13)	(−0.06)
	(−1.71)	(−0.97)	(−1.43)	(−1.69)	(−1.03)	(−0.90)	(−0.34)	(−0.35)	(−0.14)	(−0.08)
OIB#$_{t-1}$	1.65		0.75		0.52		0.21		0.23	
	(7.28)		(2.47)		(1.51)		(0.50)		(0.43)	
	(2.76)		(1.73)		(1.34)		(0.49)		(0.52)	
OIB\$$_{t-1}$		8.24		5.15		4.35		2.27		1.18
		(4.26)		(1.86)		(1.32)		(0.54)		(0.28)
		(2.64)		(1.20)		(1.54)		(0.58)		(0.26)
R^2	0.54	0.32	0.35	0.30	0.28	0.93	0.18	0.16	0.17	0.15

Bottom Line:

- OIB# predict returns at short horizons up to 30 minutes
- OIB# are positively correlated
- "In no more than thirty minutes, order imbalances lose their predictive ability.. " (p.291)
- Markets are semi-strong form efficient if you consider horizons of 30 mins or more

Paper: Chordia, Roll, and Subrahmanyam, 2008. "Liquidity and market efficiency," JFE.

- Additional Argument: Liquidity
- From Abstract

  "We find that such predictability is diminished when bid-ask spreads are narrower, and has declined over time with the minimum tick size."

  "Variance ratio tests suggest that prices were closer to random walk benchmarks in the more liquid decimal regime than in

  other ones."

  Economic punchline: "These findings indicate that liquidity stimulates arbitrage activity, which, in turn, enhances market efficiency."
- Variables

  Returns (5-minute), r_t

  OIB_t

  $QSPR_t = P^a - P^b$

  $ESPR_t = 2\left|P_t^* - P_t^{Midpoint}\right|$

● Effective Spread (Figure 2):

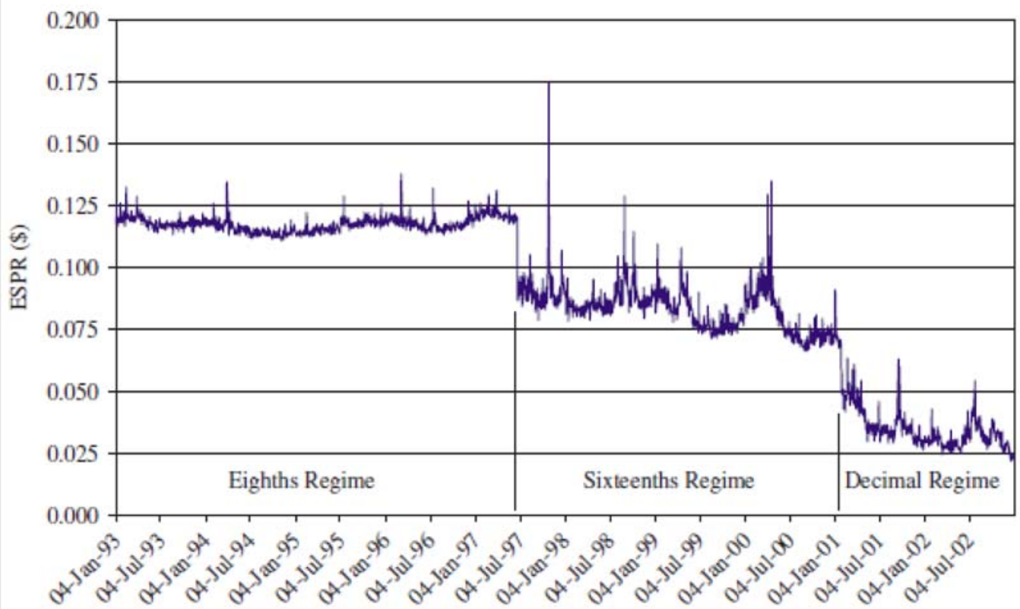


Fig. 2. Value-weighted daily average effective spread, NYSE, 1993–2002.

- Main Result: Table 5

		Coefficient		t-Statistic
Eighths Regime ($n = 86{,}692$)	$OIB\$_{t-1}$	0.0298		45.53
	$OIB\$_{t-1} * ILD$	0.0241		13.64
	Intercept	−0.0010		−5.62
	Adjusted R^2		0.0348	
Sixteenths Regime ($n = 69{,}538$)	$OIB\$_{t-1}$	0.0148		7.92
	$OIB\$_{t-1} * ILD$	0.0517		9.53
	Intercept	−0.0012		−2.99
	Adjusted R^2		0.0032	
Decimal Regime ($n = 36698$)	$OIB\$_{t-1}$	0.0059		1.55
	$OIB\$_{t-1} * ILD$	0.0278		3.23
	Intercept	−0.0003		−0.35
	Adjusted R^2		0.0005	

- Robustness Results: Table 6

		Large Firms		Mid-Cap Firms		Small Firms	
		Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statisti
Eighths Regime	$OIB\$_{t-1}$	0.0293	41.47	0.0192	37.79	0.0115	32.29
	$OIB\$_{t-1} * ILD$	0.0222	11.15	0.0097	7.58	0.0103	11.39
	Intercept	−0.0010	−4.68	−0.0006	−4.12	0.0000	−0.21
	Adjusted R^2	0.0277		0.0227		0.0194	
Sixteenths Regime	$OIB\$_{t-1}$	0.0094	4.63	0.0228	17.67	0.0207	21.39
	$OIB\$_{t-1} * ILD$	0.0514	8.62	0.0299	8.68	0.0318	12.37
	Intercept	−0.0008	−1.82	−0.0019	−6.15	−0.0015	−5.79
	Adjusted R^2	0.0019		0.0079		0.0129	
Decimal Regime	$OIB\$_{t-1}$	0.0029	0.74	0.0111	3.44	0.0245	8.84
	$OIB\$_{t-1} * ILD$	0.0287	3.11	0.0187	2.74	0.0259	4.45
	Intercept	−0.0002	−0.20	−0.0005	−0.82	−0.0022	−3.63
	Adjusted R^2	0.0003		0.0007		0.0039	

Variance Ratio Tests (Table 10)

	Eighths	Sixteenths	Decimals
Panel A: Five minutes/daily variance ratios			
Large firms	1.21	1.21	1.10
Mid-cap firms	1.81	1.82	1.36*
Small firms	2.28	2.38	1.79†
Panel B: Per hour open/close variance ratios			
Large firms	6.17	7.14	6.69
Mid-cap firms	11.85	7.88	18.01**
Small firms	8.95	15.33	22.16*
Panel C: First order autocorrelations of daily returns			
Large firms	0.0572	−0.0228	0.0039
	(0.055)	(0.493)	(0.932)
Mid-cap firms	0.0924	0.0697	0.0533
	(0.002)	(0.036)	(0.244)
Small firms	0.2121	0.0814*	0.0339*
	(0.000)	(0.014)	(0.458)

Summary

- These two papers argued that equity markets were mostly efficient.
- They were all about the properties of "unexpected" returns, or ε_{t+1}
- Similar tests would argue that markets deviate slightly from market efficiency
- Story not very different from Chordia et al. (2005): Investors are boundedly rational, because they face
  - Cognitive limitations/Limited attention (Sims (1995), Hong and Stein (1999))
  - Neglected stocks (Merton (1987))
  - Information will be impounded into prices with a lag
- Same statistical tests!

- Literature:
  - Recent Papers: Huberman and Regev (2001), Barber and Odean (2006), DellaVigna and Pollet (2006), Hou (2006), Menzly and Ozbas (2006), Hong, Torous, and Valkanov (2007), Cohen and Frazzini (2008)
  - Recent Paper: "Complicated Firms," JFE 2012
  - Premise: Some firms are more "complicated" to understand by investors than others
  - Compare returns of complicated firms vs portfolio of simple firms that replicates the complicated firms
  - The portfolio of simple firms should predict the complicated firms

Table 1: Summary Statistics

	Min	Median	Max	Mean	Std Dev
Panel A: Time series (33 annual observations, 1977 – 2009)					
Number of conglm firms in the sample per year	542	840	1288	898	198
Number of standalones in the sample per year	919	1948	3563	2069	632
Full sample % coverage of CRSP universe (EW)	51.39	79.04	86.02	77.53	6.78
Full sample % coverage of CRSP universe (VW)	61.24	89.82	93.26	85.69	7.41
Conglm firms % of CRSP universe (EW)	15.44	21.93	37.54	24.04	6.39
Conglm firms % of CRSP universe (VW)	32.07	43.90	57.20	44.56	6.19
Standalones % of CRSP universe (EW)	26.03	53.10	70.37	53.49	8.96
Standalones % of CRSP universe (VW)	29.17	40.31	54.98	41.13	6.51
Panel B: Firms (Pooled firm–year observations)					
Conglm firm market-cap percentile (NYSE)	0.00	0.51	1.00	0.50	0.32
Standalone market-cap percentile (NYSE)	0.00	0.30	1.00	0.36	0.28
Conglm firm book-to-market percentile (NYSE)	0.00	0.65	1.00	0.60	0.27
Standalone book-to-market percentile (NYSE)	0.00	0.55	1.00	0.51	0.29
# of industries per conglomerate	2	2	10	2.64	0.94
Percent of sales per industry segment	0	0.27	1	0.36	0.29

Main Results: Table 4

Dep Variable	RET_t		$RET_t - INDRET_t$		$RET_t - PCRET_t$	
*100	(1)	(2)	(3)	(4)	(5)	(6)
$PCRET_{t-1}$	**7.408**	**6.896**	**3.047**	**4.652**	**3.260**	**4.098**
	(5.84)	(6.67)	(2.72)	(5.35)	(2.56)	(3.21)
RET_{t-1}		**-4.422**		**-4.183**		**-4.583**
		(-6.88)		(-6.72)		(-7.18)
$INDRET_{t-1}$		**4.783**		-1.341		-0.296
		(3.85)		(-1.27)		(-0.25)
SIZE	-0.052	-0.048	-0.029	-0.023	-0.034	-0.031
	(-1.24)	(-1.12)	(-1.49)	(-1.05)	(-1.56)	(-1.32)
B/M	**0.212**	**0.229**	**0.209**	**0.225**	**0.217**	**0.234**
	(2.35)	(2.50)	(2.93)	(3.02)	(2.91)	(3.02)
MOM	**0.285**	**0.283**	**0.296**	**0.311**	**0.265**	**0.270**
	(2.51)	(2.46)	(2.89)	(3.02)	(2.45)	(2.54)
TURNOVER	**-0.027**	**-0.029**	**-0.025**	**-0.027**	**-0.029**	**-0.031**
	(-3.36)	(-3.51)	(-3.67)	(-3.88)	(-3.92)	(-4.09)
Adj R^2	0.06	0.07	0.03	0.04	0.03	0.04

- Robustness Check: Table 5

Dep Variable	Conglomerate Return(t)					
*100	(1)	(2)	(3)	(4)	(5)	(6)
$PCRET_{t-1}$	**8.504**	**5.995**	**8.456**	**7.871**	**7.033**	**6.720**
	(5.77)	(4.60)	(5.09)	(5.38)	(5.24)	(6.23)
$PCRET_{t-1}$* *Herfindahl > median*	**-3.458** (-3.33)					
$PCRET_{t-1}$* *Idio Vol > median*		**3.159** (2.43)				
$PCRET_{t-1}$* *MktCap > NYSE median*			**-3.181** (-2.23)			
$PCRET_{t-1}$* *Res Inst Own > median*				-1.698 (-1.20)		
$PCRET_{t-1}$* *Turnover > median*					0.361 (0.24)	
$PCRET_{t-1}$* *#Analyst > median*						-0.500 (-0.37)
CONTROLS	YES	YES	YES	YES	YES	YES
Adj R^2	0.09	0.09	0.09	0.08	0.08	0.08

- Summary: We have looked at predictive regressions in the context of market inefficiency
  - Short horizons (from 5-minutes to 2-3 days)
  - Market Efficiency: At what horizon are returns uncorrelated?
  - Market Inefficiency: Emphasize the inefficiency and design an empirical test that convincingly tests for it (controlling for previous effects)
- Next Class: Predictability and Time Varying Expected Returns