

1 Introduction

Goal

The goal of this project is to test for autocorrelations in stock returns, which persist for economically meaningful short-term timespans but are not, on average, significantly different from zero in the long term. A rejection of the null hypothesis under such tests would establish evidence against the random walk hypothesis; that is, that past returns are not independent of future returns in some short-run contexts. Such a finding would have clear implications for trading strategy.

Motivation and hypotheses

The motivation for this project is to expand upon previous tests of the random-walk hypothesis, which have found autocorrelations of stock returns to be on average zero. I hypothesize that this is the case, but that there are periods in the time series of returns ("bubbles") in which returns do in fact exhibit significant autocorrelation (either positive, when a bubble is increasing, or negative, when a bubble "bursts").

Idea

Test the random walk hypothesis (i.e. $E(r_{t+1}|r_t) = E(r_{t+1})$) using a time-varying coefficient model

$$r_{i,t} = X_t \beta_{i,t} + R_{i,t}^{(p)'} \Phi_{i,t}^{(p)} + \epsilon_{i,t}$$

where:

- $r_{i,t}$ is the return for asset i in time t
- $X_t = [X_t^{(0)} \quad \dots \quad X_t^{(k)}]$ is a vector of common factors at time t (i.e. Fama-French portfolio factors) and a constant term
- $\beta_{i,t} = [\beta_t^{(0)} \quad \dots \quad \beta_t^{(k)}]'$ is a vector of time-varying coefficients at t
- $R_{i,t} = [r_{i,t-1} \quad \dots \quad r_{i,t-p}]'$ is a vector of lagged returns for asset i at time t
- $\Phi_{i,t}^{(p)} = [\phi_{i,t}^{(1)} \quad \dots \quad \phi_{i,t}^{(p)}]'$ is a vector of time-varying AR coefficients at t
- $\epsilon_{i,t}$ is a stochastic error term for which we assume mean zero and constant conditional variance.¹

If the random walk hypothesis holds, then $\Phi_{i,t}^{(p)} = 0$ for all assets i over all times t . However, if there are time periods in which this hypothesis does not hold, we should see that at least one component of $\Phi_{i,t}^{(p)}$ is significantly different from zero. The project proposes to use a local linear estimation method such as that of Ang and Kristensen (2011); and bootstrap methods presented by Chen, Gao, Li, and Silvapulle (2018); McKinnon (2007); and Lee and Ullah (2001) to derive confidence intervals and bands, and test statistics, for these coefficients. This will allow for statistical inference on whether or not $\Phi_{i,t}^{(p)}$ is significantly different from zero at certain values of t . It will be of particular interest whether this relationship holds significantly during periods classified ex post as asset pricing bubbles.

2 Methodology

Estimation method

This project proposes to use the local linear regression method of Ang and Kristensen (2011) to estimate time-varying $\beta_{i,t}$ and $\Phi_{i,t}^{(p)}$ using daily returns data from the 30 industry portfolios from Ken French as

¹These assumptions are not crucial to inference. If the constant variance assumption is loosened, the residual bootstrap procedures will be adjusted to sample from the set of standardized residuals, rather than the set of centered residuals.

individual assets over the 36-year period from 1/1980 - 12/2016. For each asset k at each time t in this span, we estimate $\beta_{i,t}$ and $\Phi_{i,t}^{(p)}$ by minimizing the weighted sum of squares

$$\sum_{i \in 1 \dots T} K_{hT}(t_i - t) (R_{k,i} - (X_t \beta_{i,t} + R_{i,t}^{(p)'} \Phi_{i,t}^{(p)}))^2$$

where $K_{hT}(z) = K(\frac{z}{hT})$, $K(\cdot)$ is a kernel and h is a bandwidth. The optimal estimators are given by

$$[\beta_{i,t} \quad \Phi_{i,t}^{(p)}]' = (V_k' W_{hT} V_k)^{-1} (V_k' W_{hT} R_k)$$

where V_k is the matrix holding values for X_t and $R_{k,t}$ in all t , and W_{hT} is the n -dimensional diagonal matrix of kernel weights.

Bootstrap inference procedures

2.0.1 Pointwise confidence bands

This project proposes to use the bootstrap procedure outlined in Chen, Gao, Li, and Silvapulle (2018) to create approximate pointwise confidence intervals for each/any of the estimated $\hat{\Phi}_{i,t}^{(p)}$. The procedure yields pointwise intervals in the form $\hat{\phi}_{i,t}^{(j)} \pm \hat{c}_{\frac{\alpha}{2}}(t) \hat{sd}(\hat{\phi}_{i,t}^{(j)})$ as follows (note that the below assumes a single asset, so we drop the subscript k):

- I. Estimate $\hat{\phi}_{i,t}^{(j)}$ for all t using the local linear method described above. Obtain the estimated residual $\hat{\epsilon}_t$ for each t using these coefficient estimates, and create the set of centered residuals $\tilde{\epsilon}_t = \hat{\epsilon}_t - \bar{\epsilon}$.
- II. Resample $\{\epsilon_t^*\}$ with replacement from the set of centered residuals $\tilde{\epsilon}_t$. Using these residuals, generate $r_t^* = X_t \hat{\beta}_t + R_t^{(p)'} \hat{\Phi}_t^{(p)} + \tilde{\epsilon}_t$ recursively. That is, start with independent variables at some $t = 0$, and then use each subsequent r_t^* as the lagged return(s) in r_{t+1}^* .
- III. Use $\{(r_t^*, V_t)\}$ to estimate $\hat{\phi}_{i,t}^{(j)}$ for all $t = 1 \dots n$.
- IV. Repeat steps II and III B times to obtain B bootstrap estimates $\hat{\phi}_{i,t}^{(j)*}$ at each $t = 1 \dots n$.
- V. Estimate $\hat{sd}(\hat{\phi}_{i,t}^{(j)})$ as the sample standard deviation of the B bootstrap estimates at each t .²
- VI. Compute $Q_{j,t}^* = \frac{(\hat{\phi}_{i,t}^{(j)*} - \hat{\phi}_{i,t}^{(j)})}{\hat{sd}(\hat{\phi}_{i,t}^{(j)})}$ for each of B bootstrap replications. Obtain $\hat{c}_{\frac{\alpha}{2}}(t)$ as the upper $\frac{\alpha}{2}$ percentile of the empirical distribution.

Plots of these confidence bands will be produced for each of the 30 assets. It will be of interest whether we see confidence intervals for the autoregressive coefficients which don't include zero during periods classified ex post as asset pricing bubbles.

2.0.2 Simultaneous confidence bands

This project proposes to extend the confidence bounds in the above subsection by applying a size adjustment as described in MacKinnon (2007). (For example, a Bonferroni adjustment.) Using the same procedure as described above, we would define the upper and lower bounds for each $\hat{\phi}_{i,t}^{(j)}$ on the time interval t_1 to t_2 by using the $\hat{c}_{\frac{\alpha'}{2}}(t)$ percentile, where α' is adjusted for the size of the time interval, i.e. $t_2 - t_1$.

²Note that this can be obtained as the square root of the j^{th} diagonal entry in the covariance matrix of $[\beta_{i,t} \quad \Phi_{i,t}^{(p)}]'$. This matrix can be estimated using the formula on page 8 of MacKinnon (2007).

2.0.3 Hypothesis tests

This project proposes to apply the bootstrap hypothesis testing procedure proposed in Lee and Ullah (2001) to test a set of null and alternative model specifications. These different specifications will be primarily concerned with the time-varying and non-zero nature of the autoregressive coefficients. The procedure for this test is as follows:

- I. Estimate time varying coefficients $[\beta_{i,t} \quad \Phi_{i,t}^{(p)}]'$ for all t using the local linear regression described above.
- II. Estimate the factor (and, if specified, autoregressive) coefficients in some restricted model. (For example, all coefficients are time-invariant; or all autoregressive coefficients are equal to zero).
- III. Calculate test statistic $\hat{\tau} = \frac{(RSS_{restricted} - RSS_{unrestricted})}{RSS_{unrestricted}}$, which compares the RSS from the model in I (unrestricted) to the RSS from the model in II (restricted).
- IV. Generate B bootstrap samples under the restricted model, using the error bootstrap procedure described in the pointwise confidence bands section. (That is, generate the estimated residuals from the restricted model in II, and resample from these with replacement B times to recursively build B bootstrap datasets.) Use each of these B datasets, generated under the null, to estimate $\tau_i^* = \frac{(RSS_{restricted} - RSS_{bootstrap_i})}{RSS_{bootstrap_i}}$.
- V. Step IV gives us a distribution of the test statistic under the null hypothesis. The p-value for the test is then given by $p = \frac{1}{B} \sum I(\tau_i^* > \hat{\tau})$.

This procedure can be modified to test a hypothesis that some coefficients are equal to zero by using this parameter-restricted model as the null.

This project proposes to apply a second test procedure from Lee and Ullah (2001) which compares the fitted values for the coefficients under the null and alternative model specifications. This test follows the procedure described above, but calculates the test statistic $\hat{Q} = \frac{1}{n} \Sigma (\tilde{\Phi}_i^{(p)} - \hat{\Phi}_i^{(p)})$ where $\tilde{\Phi}_i^{(p)}$ is the vector of autoregressive coefficients under the null and $\hat{\Phi}_i^{(p)}$ is the vector of autoregressive coefficients under the unrestricted alternative (i.e. the time-varying case). We calculate a p-value by comparing this test statistic to the empirical distribution of $Q^* = \frac{1}{n} \Sigma (\tilde{\Phi}_i^{(p)} - \Phi_{i^*}^{(p)})$. Here the $\Phi_{i^*}^{(p)}$ are estimated for B bootstrap replications, with the dataset for each replication generated as described above.

References

- Ang, Andrew, and Kristensen, Dennis, Testing Conditional Factor Models, National Bureau of Economic Research, Working Paper 17561, 2011, <http://www.nber.org/papers/w17561.pdf>
- Xiangjin B. Chen, Jiti Gao, Degui Li and Param Silvapulle (2018) Nonparametric Estimation and Forecasting for Time-Varying Coefficient Realized Volatility Models, Journal of Business and Economic Statistics, 36:1, 88-100, DOI: 10.1080/07350015.2016.1138118
- MacKinnon, James G; Bootstrap Hypothesis Testing, Department of Economics, Queen's University, 2007, http://qed.econ.queensu.ca/working_papers/papersqed/wp_1127.pdf
- Lee, Tae-Hwy, and Ullah, Aman, Nonparametric Bootstrap Specification Testing in Economic Models, 2001, <http://www.faculty.ucr.edu/taleepaper/giles.pdf>