HW: Predictive Regression

February 8, 2023

In this HW we construct a 2-stage regression for predicting hedged CDS spread "returns" using boxcar and discounted least squares for the predictive phase, and compare the two.

1 Data

The class website contains 5 year CDS rates for debt from several companies over a multi-year range in `Liq5YCDS.delim`. Read this data, and load the corresponding adjusted close prices for the corresponding equity[1].

CDS spreads are not directly investable in the way equities are, but we can still learn a lot from treating them in ways similar to those we use for treating asset prices. In particular, just like equity prices, spreads are bounded below by zero and have no functional upper bound.

CDS spreads do not necessarily get meaningful updates every day, causing Epps effect-type problems in daily data. Thus, we will compute weekly Wednesday to Wednesday returns r^{Equity} on the adjusted equity close prices, and similar "returns" r^{CDS} on the CDS spreads. Compute weekly Wednesday to Wednesday returns r^{Equity} on the adjusted equity close prices, and similar "returns" r^{CDS} on the CDS spreads. In addition, obtain "market equity returns" m as the weekly returns on adjusted prices of the SPY ETF.

[1] Not all debt issuers *have* publicly traded equity. I have selected CDS rates from issuers that do.

2 Models

A *predictive* regression is aimed at predicting future behavior in some convenient way. For quantitative investment strategies, such a regression would typically be aimed at predicting asset returns. For predictive segments of the following analysis you will compare boxcar OLS window size 16 against exponentially decaying regression weights with half life 12. Contemporaneous regressions should be boxcar OLS with window size 16.

Begin by forming a CDS "index return" r^{Index} as the arithmetic average of the r^{CDS}.

For each ticker $E = E_1, \ldots, E_N$, you will be working with both contemporaneous and predictive models for its "spread returns". The contemporaneous model is of the form

$$r_E^{\text{CDS}} \sim r_E^{\text{Equity}} + r^{\text{Index}} + \epsilon. \tag{1}$$

Starting from the $K+1$st week of available returns on the ticker, define weekly calibration data for contemporaneous returns as the returns from the K previous weeks, where in our case we use $K = 16$.

Create one contemporaneous model of the (CAPM) form

$$r_E^{\text{Equity}} \sim m + \epsilon \tag{2}$$

using 16 week boxcar OLS and denote its weekly regression coefficients for the n-th data row as $\gamma_{E,n}$.

Also create a contemporaneous model of the form

$$r_E^{\text{CDS}} \sim r_E^{\text{Equity}} + r^{\text{Index}} + \epsilon. \tag{3}$$

using 16 week boxcar OLS.

For the upcoming n-th data row, we attempt to predict its change in CDS spread *hedged by the contemporaneous predictors*. Define the *hedge portfolio return* as the returns on predictors in the contemporaneous model[2]

$$f_{E,n} = \beta_{E,\text{Equity}}^{(n)} r_{E,n}^{\text{Equity}} + \beta_{E,\text{Index}}^{(n)} r_n^{\text{Index}}. \tag{4}$$

As you see $f_{E,n}$ is just a prediction from our contemporaneous model for a single data row, so it is easy to obtain.

[2]Remember that the β values depend on both the ticker and which data row n we are working with.

Now we will define the *residual return*[3] as the residual error in this prediction

$$\rho_{E,n} = r_{E,n}^{\text{CDS}} - f_{E,n}. \tag{5}$$

We also define residual equity return as

$$c_{E,n} = r_{E,n}^{\text{Equity}} - \gamma_{E,n} m_n \tag{6}$$

Once we have our data series of residual returns ρ for all equities and weeks, we will form predictive regression models in exponentially decaying and boxcar forms. Create new models of the form

$$\rho_{E,n} \sim c_{E,n-1} + \epsilon \tag{7}$$

that use the past week's equity return residual to predict novel changes (i.e. residuals) to CDS spread based on previous residual equity returns. and call the regression coefficients $\mu_{E,n}$.

The residuals of the predictive model itself are

$$q_{E,n} = \rho_{E,n} - \mu_{E,n-1} c_{E,n-1} \tag{8}$$

so that, for example, $\mu_{E,45}$ is computed by using data rows[4] $i = 4 - 45$ of $\rho_{E,i}$ and data rows $i = 3 - 44$ of $c_{E,i}$.

Thus we have two contemporaneous regressions per ticker, one for equity and one for CDS. These may both be kept as 16-week boxcar throughout. This gives us ρ and c. For the exponential and boxcar regressions with c predicting ρ, you may choose your own window sizes and half lives, though 16 and 12 are fine starting points.

3 Analysis

Compare performance of predictive regressions in exponentially decaying (i.e. discounted) versus boxcar forms. Are there historical events where they have better or worse performance? How are the tails? What statistical properties differ significantly between them?

In this assignment, you are not expected to consider a trading strategy.

[3] Confusing fact: residual return is often also called "hedged return" which is dangerously similar-sounding to "hedge return".

[4] These start at 4 and 3 rather than 2 and 1 so that we have nontrivial regressions in the early rows.