Diesen Termin sollten Sie sich nicht entgehen lassen!

www.ai-conference.com
Ort: IHK, Frankfurt

Inhalt

Leitartikel ................................................................................................................................................................ 2 - 4

Mitgliederneuvorstellung .................................................................................................................................. 6

Impressionen der BAI AIC 2014 .....................................................................................................................7 - 13

Mit einer Service-KVG in die regulierte Welt
Johannes Glasl, DSC Deutsche SachCapital GmbH ......................................................................................... 14 - 17

Spannungsfelder der neuen Investmentbesteuerung – zwischen Privilegierung und Diskriminierung
Christiane Schönbach, Robert Welzel, WTS ................................................................................................. 18 - 25

Efficient estimation of non-linear Hedge Fund returns
Nicolas Fuchshofen, BAI .................................................................................................................................... 26 - 30

Veranstaltungen ............................................................................................................................................. 31 - 32

Buchvorstellungen ......................................................................................................................................... 34 - 35
Efficient estimation of non-linear Hedge Fund returns

It is broadly known that Hedge Fund returns do either follow a normal distribution or have a strictly linear connection to market returns. Thus they cannot be explained via a simple linear regression like OLS, where only returns of the particular time series are taken into consideration. Below a well-known approach is used to illustrate some issues concerning the explanatory power of statistical models regarding Hedge Fund returns. We use the EDHEC Equity L/S Index as a proxy for a certain Hedge Fund strategy and apply the Fama French / Carhart factor model to the data:

\[ R = \beta_0 + \beta_1 (R_{MK} - R_{RF}) + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 WML_t + \varepsilon_t \]

The model describes the Hedge Fund returns \( R \) through the market risk premium \( (R_{MK} - R_{RF}) \) and three additional variables that correct for differences in book \( (HML) \) and market capitalization \( (SMB) \) values and momentum effects \( (WML) \). First, we need to check, whether the different time series are normally distributed, since otherwise the OLS regression does not hold. For instance, the coefficient of determination \( R^2 \) represents the squared Pearson Correlation Coefficient, which captures variances and covariances of normally distributed variables. For testing the time series the dependent and all independent variables are examined through the Jarque-Bera test \( (JB) \). This test uses information about the higher moments of the distribution (skewness and kurtosis) and is asymptotically chi-square distributed with two degrees of freedom. The corresponding null hypothesis assumes a normal distribution:

\[ R: JB = 20.2096, p-value 4.08824e-005 \]
\[ R_{MK}-R_{RF}: JB = 34.4499, p-value 3.30601e-008 \]
\[ SMB: JB = 0.402049, p-value 0.817892 \]
\[ HML: JB = 1.37831, p-value 0.502 \]
\[ WML: JB = 16.5123, p-value 0.000259655 \]

The normality assumptions can be rejected regarding the Hedge Fund returns, the market risk premium and momentum effects. Beyond the test regarding normal distribution it is crucial to examine, whether the variables are stationary respectively have a unit root. In the case of a unit root, the data are subject to trends and are not mean reverting, which can potentially cause serious problems with the inference taken from the regression output. A so called “spurious regression” can occur, if the variables share a common trend. The Augmented Dickey-Fuller test \( (ADF test) \) is used, which assumes a unit root in its null hypothesis and a “constant and trend” in the application shown below:

\[ R: ADF tau = -4.70799, \text{ asymptotic p-value 0.0006354} \]
\[ R_{MK}-R_{RF}: ADF tau = -3.7916, \text{ asymptotic p-value 0.0169} \]
\[ SMB: ADF tau = -4.64877, \text{ asymptotic p-value 0.0008082} \]
\[ HML: ADF tau = -5.69924, \text{ asymptotic p-value 5.967e-006} \]
\[ WML: ADF tau = -1.87821, \text{ asymptotic p-value 0.6658} \]

The Hedge Fund returns, \( SMB \) and \( HLM \) are stationary in a highly significant way, whereas the market premium can be assumed as stationary on the 5% significance level. The momentum effects seem to have a trend. A common solution for both non-normally distributed and non-stationary variables is the use of the first differences of the corresponding data.

---

1 For a broad discussion see for example Viebig, Jan (2011): What do we know about the risk and return characteristics of hedge funds? Journal of Derivatives & Hedge Funds, 2/18, pp. 167-191.
2 Data obtained from http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/Ff_developed.html
This leads to a loss of information regarding the absolute level of observed returns, but allows – hopefully – the application of the desired model. First, we apply the Jarque-Bera test again:

\[ dR: JB = 0.401516, \text{ p-value } 0.8181 \]
\[ dR_{MK-R_{TF}}: JB = 5.54967, \text{ p-value } 0.0623598 \]
\[ dSMB: JB = 1.23638, \text{ p-value } 0.53892 \]
\[ dHLM: JB = 0.504249, \text{ p-value } 0.777148 \]
\[ dWML: JB = 76.2087, \text{ p-value } 2.82803e-017 \]

The first differences of the market risk premium are only significantly normally distributed on the 10 percent confidence level, whereas the momentum effects are even less normally distributed. One possibility is to relax the variable \(dWML\) from the regression model. A problem can occur, if \(dWML\) has some explanatory power regarding the Hedge Fund returns that remain in the residuals of the regression. The consequent serial correlation could cause inefficiency in the standard errors of the regression coefficients. But before testing on serial correlation we examine, if the first differences of the variable are stationary:

\[ dR: ADF \, \text{tau} = -11.4951, \text{ asymptotic p-value } 3.883e-026 \]
\[ d(R_{MK-R_{TF}}): ADF \, \text{tau} = -13.0647, \text{ asymptotic p-value } 2.097e-033 \]
\[ dSMB: ADF \, \text{tau} = -12.9999, \text{ asymptotic p-value } 4.287e-033 \]
\[ dHLM: ADF \, \text{tau} = -12.8654, \text{ asymptotic p-value } 1.881e-032 \]
\[ dWML: ADF \, \text{tau} = -2.07564, \text{ asymptotic p-value } 0.5589 \]

Now all variables apart from the momentum effects are stationary on the highest significance level. As the Jarque-Bera test statistic already indicates \(dWML\) is dropped from the regression. Effectively the Durbin-Watson statistic (2.78) and the applied Breusch-Godfrey test for autocorrelation (12 lags included) for the regression

\[ R = \beta_0 + \beta_1 d(R_{MK-R_{TF}}) + \beta_2 dSMB_t + \beta_3 dHML_t + \epsilon_t \]

indicate serial correlation. This can be seen in the displayed residual correlogram and the corresponding autocorrelation and partial autocorrelation function below. The difference between both functions is that the partial one eliminates the influence of the variables in between the different lags.

Residual autocorrelation and partial autocorrelation function of OLS: \( R = \beta_0 + \beta_1 d(R_{MK-R_{TF}}) + \beta_2 dSMB_t + \beta_3 dHML_t + \epsilon_t \).
As the Durbin-Watson statistic suggests (2.78), the serial correlation of the first three orders is negative. But even, if \( dWML \) is included into the model, we observe serial correlation (DW statistic 2.68). Again, the coefficients are estimated consistently, but the standard errors may be not efficient. The linear regression with monthly data excluding the non-normal distributed and non-stationary \( dWML \) between February 2005 and March 2014 leads to the following outcome:

\[
\begin{align*}
\text{OLS, using observations 2005:02-2014:03 (T = 110)} \\
\text{Dependent variable: Equity L/S Index} \\
\begin{array}{lllll}
\hline
\text{coefficient} & \text{std. error} & \text{t-ratio} & \text{p-value}\text{\textsuperscript{3}} \\
\hline
\text{const} & -4.49637e-05 & 0.000919445 & -0.04890 & 0.9611 \\
\text{d(RMkt-RFR)} & 0.406904 & 0.0156948 & 25.93 & 1.08e-04\text{***} \\
\text{dSMB} & 0.212727 & 0.0501099 & 4.245 & 4.70e-05\text{***} \\
\text{dHLM} & -0.257740 & 0.0499253 & -5.163 & 1.15e-06\text{***} \\
\hline
\end{array}
\end{align*}
\]

With an \( R^2 \) of 0.869 the linear regression captures 86.9 percent of the variations of the Hedge Fund returns. Apart from the constant all coefficients are significant at the 1% level, whereby \( dHML \) has a negative relationship to the Hedge Fund Index. Besides serial correlation, heteroscedasticity is present in the data (Lagrange multiplier (LM) test statistic: 6.35), as it can only denied at the 10% significance level. Lastly the coefficients have very high common explanatory power, which is given by the very small p-value of the F-statistic.

For the purpose of getting more reliable standard errors, an autoregressive model AR(1) is used in addition to the first three lags (see residual PACF) of the independent variables:

\[
R = \beta_0 + \beta_1 d(RMkt-RFR) + \beta_2 d(RMkt-RFR)(-1) + \beta_3 d(RMkt-RFR)(-2) + \beta_4 dSMB + \beta_5 dSMB(-1)
+ \beta_6 dSMB(-2) + \beta_7 dSMB(-3) + \beta_8 dHLM + \beta_9 dHLM(-1) + \beta_{10} dHLM(-2) + \beta_{11} dHLM(-3) + \epsilon_t
\]

As some lags of the first differences of the market premium and of \( SMB \) are not significant, they could be dropped from the regression. But for the purpose of including observed serial correlation into the model, we retain them:

\[
\begin{align*}
\text{Cochrane-Orcutt, using observations 2005:06-2014:03 (T = 106)} \\
\text{Dependent variable: Equity L/S Index} \\
\text{rho} = -0.407042 \\
\begin{array}{lllll}
\hline
\text{coefficient} & \text{std. error} & \text{t-ratio} & \text{p-value}\text{\textsuperscript{3}} \\
\hline
\text{const} & -2.00646e-05 & 0.000565766 & -0.03546 & 0.9718 \\
\text{d(RMkt-RFR)} & 0.403179 & 0.0167983 & 24.00 & 1.26e-04\text{***} \\
\text{d(RMkt-RFR)}(-1) & 0.0193238 & 0.0171760 & 1.125 & 0.2635 \\
\text{d(RMkt-RFR)}(-2) & 0.0483358 & 0.0171728 & 2.729 & 0.0076\text{***} \\
\text{d(RMkt-RFR)}(-3) & 0.00696461 & 0.0173953 & 0.4004 & 0.6898 \\
\text{dSMB} & 0.184576 & 0.0556454 & 3.317 & 0.0013\text{***} \\
\text{dSMB}(-1) & -0.0500874 & 0.0569203 & -0.8800 & 0.3812 \\
\text{dSMB}(-2) & -0.0587512 & 0.0552840 & -1.063 & 0.2907 \\
\text{dSMB}(-3) & 0.0194468 & 0.0570957 & 0.3406 & 0.7342 \\
\text{dHLM} & -0.245239 & 0.0507281 & -4.834 & 5.26e-06\text{***} \\
\text{dHLM}(-1) & 0.0962727 & 0.0499720 & 1.927 & 0.0571* \\
\text{dHLM}(-2) & -0.140224 & 0.0500058 & -2.804 & 0.0061\text{***} \\
\text{dHLM}(-3) & 0.00563301 & 0.0505570 & 0.1114 & 0.9115 \\
\hline
\end{array}
\end{align*}
\]

\(\text{R-squared} \quad 0.869916 \quad \text{Adj. R-squared} \quad 0.866234 \quad \text{F}(3, 106) \quad 236.2860 \quad \text{P-value}(F) \quad 8.75e-47 \quad \text{Durbin-Watson} \quad 2.787129 \)

\(3\text{ Significance levels: ***(1%), ** (5%), * (10%)}\)
During the Cochrane-Orcutt procedure the coefficients obtained from the OLS regression are transformed through the application of the autoregressive parameter rho, which is calculated from the OLS residuals. The new coefficients as well as the new standard errors are merely slightly different from the OLS regression above. Hence, we are able to capture more than 90% of the variation of the Hedge Fund returns through the AR(1) regression. But nevertheless, we observe persistent autocorrelation problems. The Durbin-Watson statistic is better than before, but this time it is calculated from the transformed model, which uses a lagged dependent variable as well, and therefore not reliable.

As we can see in the graph above, the problem of serial correlation is slightly smaller than in the OLS regression before. The persistence might be solved with an ARMA model, which includes serial correlation of the residuals into the regression through a moving average term in addition to the autoregressive one used above. We use the same regression, but relax the third lags of the differences as they showed no significance and add two moving average terms to the AR(1) one:

\[ R = \beta_0 + \beta_1 d(R_{MK} - R_f) + \beta_2 d(R_{MK} - R_f)(-1) + \beta_3 d(R_{MK} - R_f)(-2) + \epsilon_t \]

\[ R = \beta_0 + \beta_1 d(R_{MK} - R_f) + \beta_2 d(R_{MK} - R_f)(-1) + \beta_3 d(R_{MK} - R_f)(-2) + \beta_4 d(R_{MK} - R_f)(-3) + \beta_5 dSMB_1 + \beta_6 dSMB_2(-1) + \beta_7 dSMB_2(-2) + \beta_8 dSMB_2(-3) + \beta_9 dHLM_1 + \beta_{10} dHLM_2(-1) + \epsilon_t \]

### Residual autocorrelation and partial autocorrelation function of AR(1):

\[ (\beta_0 + \beta_1 d(R_{MK} - R_f) + \beta_2 d(R_{MK} - R_f)(-1) + \beta_3 d(R_{MK} - R_f)(-2)) + \epsilon_t \]

\[ (\beta_0 + \beta_1 d(R_{MK} - R_f) + \beta_2 d(R_{MK} - R_f)(-1) + \beta_3 d(R_{MK} - R_f)(-2) + \beta_4 d(R_{MK} - R_f)(-3) + \beta_5 dSMB_1 + \beta_6 dSMB_2(-1) + \beta_7 dSMB_2(-2) + \beta_8 dSMB_2(-3) + \beta_9 dHLM_1 + \beta_{10} dHLM_2(-1) + \epsilon_t) \]

As we can see in the graph above, the problem of serial correlation is slightly smaller than in the OLS regression before. The persistence might be solved with an ARMA model, which includes serial correlation of the residuals into the regression through a moving average term in addition to the autoregressive one used above. We use the same regression, but relax the third lags of the differences as they showed no significance and add two moving average terms to the AR(1) one:

**ARMA, using observations 2005:04-2014:03 (T = 108)**

- Estimated using Kalman filter (exact ML)
- Dependent variable: Equity L/S Index
- Standard errors based on Hessian

<table>
<thead>
<tr>
<th></th>
<th>coefficient</th>
<th>std. error</th>
<th>z-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>const</td>
<td>-3.79921e-05</td>
<td>5.66787e-05</td>
<td>-0.6703</td>
</tr>
<tr>
<td>AR(1)</td>
<td>-0.938746</td>
<td>0.0746495</td>
<td>-12.58</td>
</tr>
<tr>
<td>MA(1)</td>
<td>0.0480407</td>
<td>0.0739729</td>
<td>0.6494</td>
</tr>
<tr>
<td>MA(2)</td>
<td>-0.899429</td>
<td>0.0710018</td>
<td>-12.67</td>
</tr>
<tr>
<td>d(RMK-RFR)</td>
<td>0.406957</td>
<td>0.0138988</td>
<td>29.28</td>
</tr>
<tr>
<td>d(RMK-RFR)</td>
<td>0.00991096</td>
<td>0.0142877</td>
<td>0.6937</td>
</tr>
<tr>
<td>d(RMK-RFR)</td>
<td>0.0340256</td>
<td>0.0142778</td>
<td>2.383</td>
</tr>
<tr>
<td>dSMB</td>
<td>0.234324</td>
<td>0.0479067</td>
<td>4.891</td>
</tr>
<tr>
<td>dSMB(-1)</td>
<td>-0.00853781</td>
<td>0.0472131</td>
<td>-0.1808</td>
</tr>
<tr>
<td>dSMB(-2)</td>
<td>-0.0452232</td>
<td>0.0455715</td>
<td>-0.9924</td>
</tr>
<tr>
<td>dHLM</td>
<td>-0.211808</td>
<td>0.0449206</td>
<td>-4.715</td>
</tr>
<tr>
<td>dHLM(-1)</td>
<td>0.107849</td>
<td>0.0453329</td>
<td>2.379</td>
</tr>
<tr>
<td>dHLM(-2)</td>
<td>-0.101498</td>
<td>0.0455489</td>
<td>-2.228</td>
</tr>
</tbody>
</table>

**Efficient estimation of non-linear Hedge Fund returns**
As seen in the AR(1) regression the autoregressive parameter is highly significant in the ARMA regression as well. Moreover the p-value of the MA(2) parameter indicates significant moving average issues of the second order. Like in the OLS regression the independent variables without lag are highly significant. In addition to second lags of the market premium and dHLM the first lag of dHML is significant on the 5% level. Thus, the Hedge Fund returns are influenced by the current market premium and the one two months before, whilst the standard error of the market premium is smaller than in the regressions run. The influence of dSMB is about 10% higher than in the first OLS regression and has a smaller standard error as well. Whereas the coefficient of dHML is absolutely speaking smaller, but preserves its negative relationship to the Hedge Fund returns. The sign changes over the subsequent lags, so that the relationship of the first lags coefficient and its corresponding variable is positive. The standard errors are about the same with all three book value variables. The smaller standard errors of the coefficients indicate a higher level of efficiency of the regression.

Furthermore we got rid of the problems of serial correlation and heteroscedasticity. The test for autocorrelation up to lag order twelve leads to an acceptation of the null hypothesis of no serial correlation (Ljung-Box Q' = 6.15). In addition the ARCH test with the same lag length indicates no significant autoregressive or heteroscedastic effects (LM test statistic = 16.27) anymore. Finally we need to check, whether the residuals of the ARMA regression are normally distributed. Otherwise the output may be invalid. With a test statistic of 0.51 and a corresponding p-value of 0.77 we cannot reject the null hypothesis of normally distributed residuals.

So, we were able to estimate the Hedge Fund returns through the use of the Fama French / Carhart factor model consistently and – with the ARMA model – in the most efficient way. Since the ARMA model executes a non-linear regression, no "traditional" R² is displayed. The Rank Correlation Coefficient (Spearman's rho) between the fitted and actual values of the Hedge Fund returns equals 0.96. The graph shows the comparison between the predicted Hedge Fund returns using the ARMA regression and their actual values.

Kontakt:

Nicolas Fuchshofen, Referent Research und Öffentlichkeitsarbeit

Bundesverband Alternative Investments e.V. (BAI), Poppelsdorfer Allee 106, 53115 Bonn
Tel: +49 (0) 228-96987-15, Fax: +49 (0) 228-96987-90
E-Mail: fuchshofen@bvai.de, Internet: www.bvai.de