Introduction to Quantitative Methods
Empirical Evidence

Dr. Aris Kartsaklas
aris.kartsaklas@brunel.ac.uk

Univariate Time Series with Applications in Finance
Autumn Term 2011/2012

1 Financial Time Series and Regression Analysis (EMPIRICAL EVIDENCE)

1.1 Can US mutual funds beat the market?

- Jensen (1968) was the first to systematically test the performance of mutual funds, and in particular examine whether any 'beat the market'.

- He used a sample of annual returns on the portfolios of 115 mutual funds from 1945–64.

  Each of the 115 funds was subjected to a separate OLS time series regression of the form

  (1) $R_{jt} - R_{ft} = \alpha_j + \beta_j (R_{mt} - R_{ft}) + u_{jt}$

  where

  R_{jt} is the return on portfolio j at time t,

  R_{ft} is the return on a risk-free proxy (a 1-year government bond)

  R_{mt} is the return on a market portfolio proxy

  u_{jt} is an error term and

  α_j, β_j are parameters to be estimated

- The quantity of interest is the significance of α_j, since this parameter defines whether the fund **outperforms** or **underperforms** the market index

  A positive and significant α_j for a given fund would suggest that the fund is able to earn significant abnormal returns in excess of the market-required return for a fund of this given riskiness

  This coefficient has become known as 'Jensen's alpha' (thus the null hypothesis is given by H_0 : $\alpha_j = 0$)

- Summarized regression results across the 115 funds are given in table 1

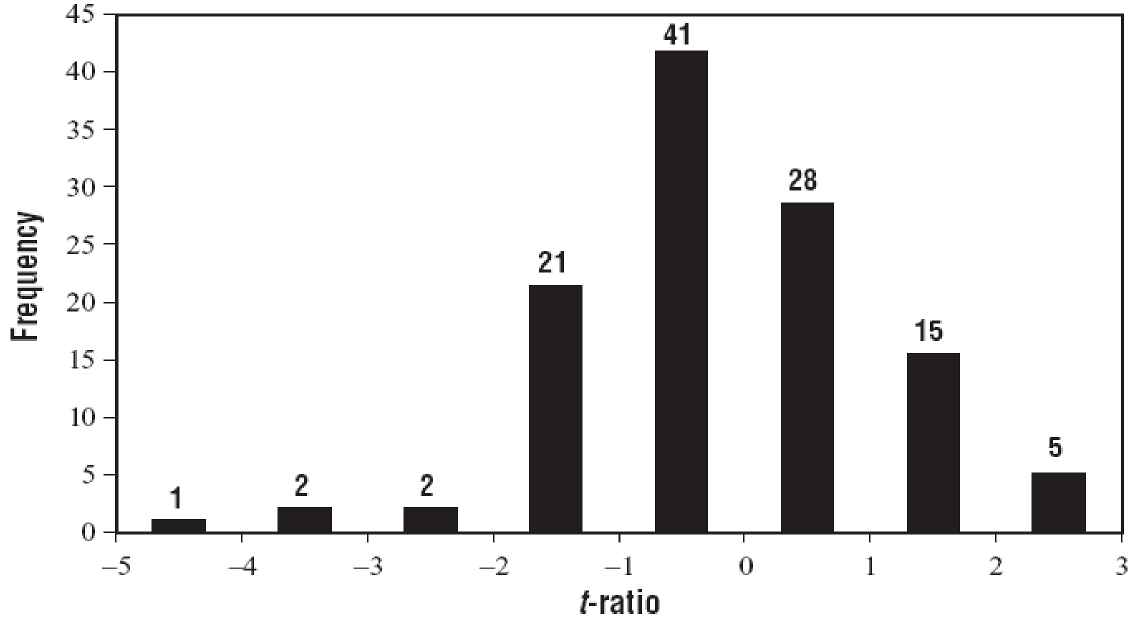
Table 1. Summary statistics for the estimated regression results

			Extremal values	
Item	Mean value	Median value	Minimum	Maximum
$\hat{\alpha}$	−0.011	−0.009	−0.080	0.058
$\hat{\beta}$	0.840	0.848	0.219	1.405
Sample size	17	19	10	20

Source: Jensen (1968). Reprinted with the permission of Blackwell Publishers.

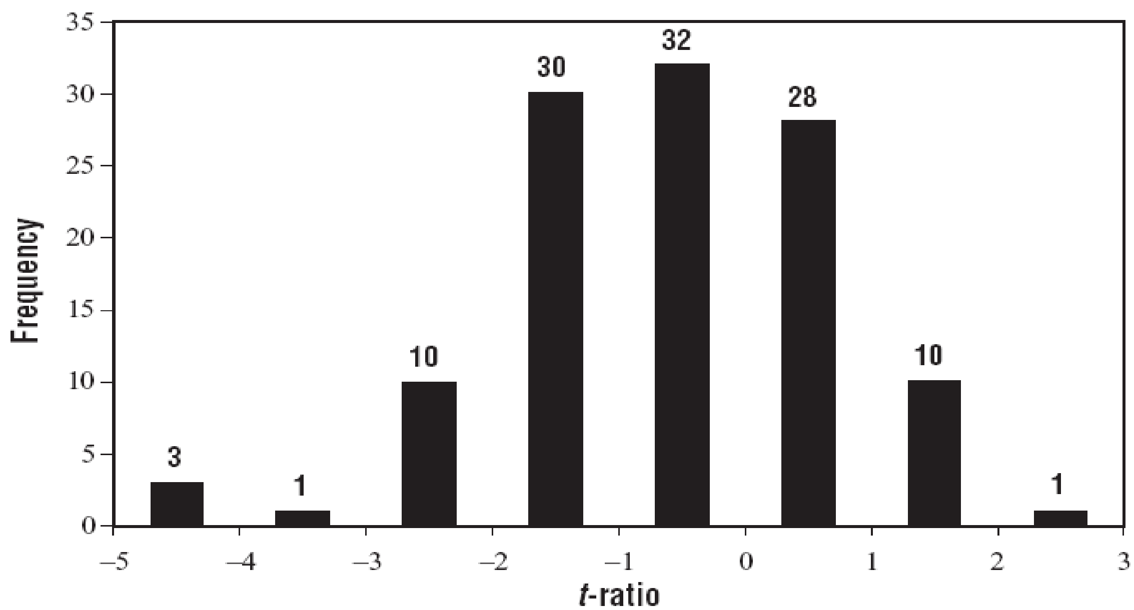
• Frequency distribution of t-ratios of mutual fund alphas (gross of transactions costs)

Figure 1.



• Frequency distribution of t-ratios of mutual fund alphas (net of transactions costs)

Figure 2.



- The appropriate critical value for a two-sided test of $\alpha_j = 0$ is approximately 2.10

  (assuming 20 years of annual data leading to 18 degrees of freedom)

- Only five funds have estimated t-ratios greater than 2

  → Only five funds outperform the market before transactions costs are taken into account.

- Interestingly, five firms have also significantly underperformed the market, with t-ratios of –2 or less

- When transactions costs are taken into account (figure 2)

  Only one fund out of 115 is able to significantly outperform the market

  Fourteen funds significantly underperform it.

- Given that a nominal 5% two-sided size of test is being used, one would expect two or three funds to 'significantly beat the market' by chance alone

  It would thus be concluded that, during the sample period studied, US fund managers appeared unable to systematically generate positive abnormal returns.

1.2 Can UK unit trust managers beat the market?

- Jensen's study has proved pivotal in suggesting a method for conducting empirical tests of the performance of fund managers.

  Criticism: Only between 10 and 20 annual observations were used for each regression

  Such a small number of observations is really insufficient for the asymptotic theory underlying the testing procedure to be validly invoked.

- A variant on Jensen's test is now estimated in the context of the UK market, by considering monthly returns on 76 equity unit trusts. The data cover the period January 1979–May 2000 (257 observations for each fund)

- Summary statistics for unit trust returns, January 1979–May 2000

  Table 2.

	Mean (%)	Minimum (%)	Maximum (%)	Median (%)
Average monthly return, 1979–2000	1.0	0.6	1.4	1.0
Standard deviation of returns over time	5.1	4.3	6.9	5.0

- From these summary statistics, the average continuously compounded return is 1.0% per month, although the most interesting feature is the wide variation in the performances of the funds.

  The worst-performing fund yields an average return of 0.6% per month over the 20-year period, while the best would give 1.4% per month.

- This variability is further demonstrated in figure 3, which plots over time the value of £100 invested in each of the funds in January 1979.

- A regression of the form (1) is applied to the UK data and the summary results presented in table 3

- CAPM regression results for unit trust returns, January 1979–May 2000

  Table 3.

Estimates of	Mean	Minimum	Maximum	Median
$\alpha(\%)$	−0.02	−0.54	0.33	−0.03
β	0.91	0.56	1.09	0.91
t-ratio on α	−0.07	−2.44	3.11	−0.25

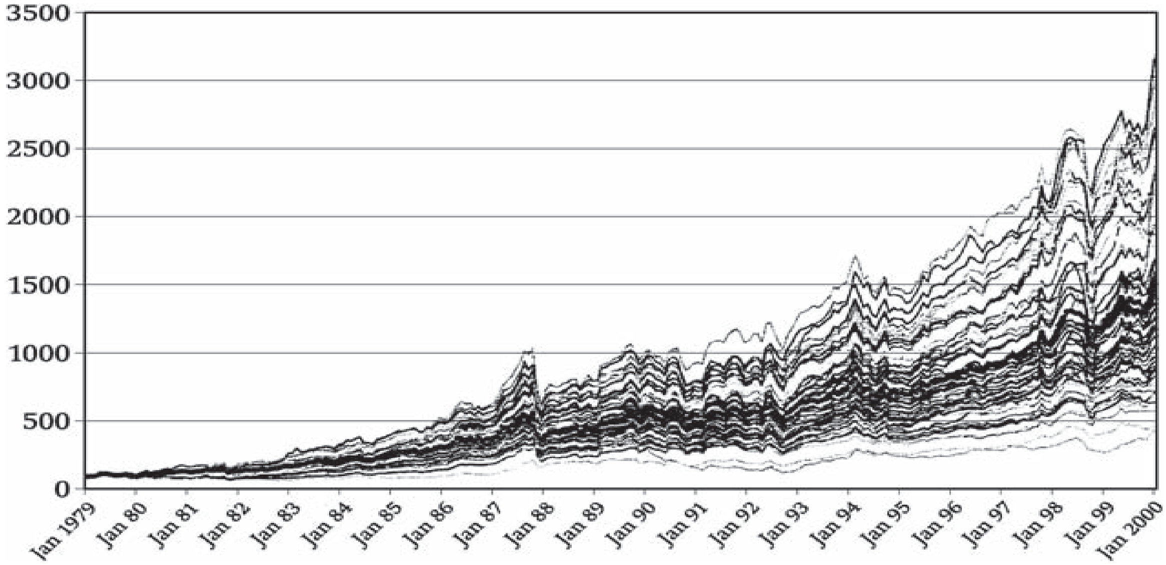
- **First**, most of the funds have estimated betas less than one again, perhaps suggesting that the fund managers have historically been risk-averse or investing disproportionately in blue chip companies in mature sectors

- **Second**, gross of transactions costs, nine funds of the sample of 76 were able to significantly outperform the market by providing a significant positive alpha, while seven funds yielded significant negative alphas.

- **The average fund (where 'average' is measured using either the mean or the median) is not able to earn any excess return over the required rate given its level of risk**

- Performance of UK unit trusts 1979-2000

Figure 3.



1.3 The overreaction hypothesis and the UK stock market

- Two studies by DeBondt and Thaler (1985, 1987) showed that stocks experiencing a poor performance over a 3–5-year period subsequently tend to outperform stocks that had previously performed relatively well.

  This implies that, on average, stocks which are 'losers' in terms of their returns subsequently become 'winners', and vice versa.

- Clare and Thomas (1995) conduct a similar study using monthly UK stock returns from January 1955 to 1990 (36 years) on all firms traded on the London Stock exchange.

- This phenomenon seems at first blush to be inconsistent with the efficient markets hypothesis, and Clare and Thomas propose two explanations

- **First** is that the 'overreaction effect' is just another manifestation of the 'size effect'.

  The size effect is the tendency of small firms to generate on average, superior returns to large firms.

  The argument would follow that the losers were small firms and that these small firms would subsequently outperform the large firms.

- DeBondt and Thaler did not believe this a sufficient explanation, but Zarowin (1990) found that allowing for firm size did reduce the subsequent return on the losers

  Zarowin (1990) also finds that 80% of the extra return available from holding the losers accrues to investors in January, so that almost all of the 'overreaction effect' seems to occur at the start of the calendar year.

- **Second** is that the reversals of fortune reflect changes in equilibrium required returns.

  The losers are argued to be likely to have considerably higher CAPM betas, reflecting investors' perceptions that they are more risky.

  Of course, betas can change over time, and a substantial fall in the firms' share prices (for the losers) would lead to a rise in their leverage ratios, leading in all likelihood to an increase in their perceived riskiness.

  Therefore, the required rate of return on the losers will be larger, and their ex post performance better.

- Ball and Kothari (1989) find the CAPM betas of losers to be considerably higher than those of winners.

- Clare and Thomas take a random sample of 1000 firms and, for each, they calculate the monthly excess return of the stock for the market over a 12-, 24- or 36-month period for each stock i

  $U_{it} = R_{it} - R_{mt} \quad t = 1, ..., n,\ i = 1, ..., 1000,\ n = 12, 24$ or 36

- Then the average monthly return over each stock i for the first 12-, 24-, or 36-month period is calculated

  $\bar{R}_i = \frac{1}{n}\sum_{t=1}^{n} U_{it}$

- The stocks are then ranked from highest average return to lowest average return

  From these average returns 5 portfolios are formed and returns are calculated assuming an equal weighting of stocks in each portfolio

  Portofolio 1 (Best performing 20% of firms)

  ...

  Portofolio 5 (Worst performing 20% of firms)

- The same sample length n is used to monitor the performance of each portfolio.

  Thus, for example, if the portfolio formation period is one, two or three years, the subsequent portfolio tracking period will also be one, two or three years, respectively.

  Then another portfolio formation period follows and so on until the sample period has been exhausted.

- How many samples of length n will there be? First, suppose $n = 1$ year.

  Estimate $\bar{R}_{it}$ for year 1 ... Monitor portfolios for year 2

  Estimate $\bar{R}_{it}$ for year 3 ... Monitor portfolios for year 4

  ...

  Monitor portfolios for year 36

- So if $n = 1$, there are 18 independent (non-overlapping) observation periods and 18 independent tracking periods.

  By similar arguments, $n = 2$ gives 9 independent periods and $n = 3$ gives 6 independent periods.

- The mean return for each month over the 18, 9, or 6 periods for the winner and loser portfolios (the top 20% and bottom 20% of firms in the portfolio formation period) are denoted by $\bar{R}^W_{pt}$ and $\bar{R}^L_{pt}$ respectively.

  Define the difference between these as $\bar{R}_{Dt} = \bar{R}^W_{pt} - \bar{R}^L_{pt}$

- The first regression to be performed is of the excess return of the losers over the winners on a constant only

  (2) $\bar{R}_{Dt} = \alpha_1 + \eta_t$

  where η_t is an error term.

- The test is of whether α_1 is significant and positive. However, a significant and positive α_1 is not a sufficient condition for the overreaction effect to be confirmed because it could be owing to higher returns being required on loser stocks owing to loser stocks being more risky.

- The solution, Clare and Thomas (1995) argue, is to allow for risk differences by regressing against the market risk premium

  $\bar{R}_{Dt} = \alpha_2 + \beta(R_{mt} - R_{ft}) + \eta_t$

  where

  R_{mt} is the return on the FTA All-share

  R_{ft} is the return on a UK government three-month Treasury Bill

- The results for each of these two regressions are presented in table 4.

  As can be seen by comparing the returns on the winners and losers in the first two rows of table 4

  12 months is not a sufficiently long time for losers to become winners

  By the two-year tracking horizon, however, the losers have become winners, and similarly for the three-year samples.

- This translates into an average 1.68% higher return on the losers than the winners at the two-year horizon, and 1.56% higher return at the three-year horizon.

- Recall that the estimated value of the coefficient in a regression of a variable on a constant only is equal to the average value of that variable.

  It can also be seen that the estimated coefficients on the constant terms for each horizon are exactly equal to the differences between the returns of the losers and the winners.

  This coefficient is statistically significant at the two-year horizon, and marginally significant at the three-year horizon.

- Is there an overreaction effect in the UK stock market?

  Table 4.

Panel A: All Months			
	$n = 12$	$n = 24$	$n = 36$
Return on loser	0.0033	0.0011	0.0129
Return on winner	0.0036	−0.0003	0.0115
Implied annualised return difference	−0.37%	1.68%	1.56%
Coefficient for (2.55): $\hat{\alpha}_1$	−0.00031	0.0014**	0.0013
	(0.29)	(2.01)	(1.55)
Coefficients for (2.56): $\hat{\alpha}_2$	−0.00034	0.00147**	0.0013*
	(−0.30)	(2.01)	(1.41)
Coefficients for (2.56): $\hat{\beta}$	−0.022	0.010	−0.0025
	(−0.25)	(0.21)	(−0.06)
Panel B: all months except January			
Coefficient for (2.55): $\hat{\alpha}_1$	−0.0007	0.0012*	0.0009
	(−0.72)	(1.63)	(1.05)

Notes: t-ratios in parentheses; * and ** denote significance at the 10% and 5% levels, respectively.
Source: Clare and Thomas (1995). Reprinted with the permission of Blackwell Publishers.

- In the second test regression, $\bar{\beta}$ represents the difference between the market betas of the winner and loser portfolios

  None of the beta coefficient estimates are even close to being significant, and the inclusion of the risk term makes virtually no difference to the coefficient values or significances of the intercept terms.

- Removal of the January returns from the samples reduces the subsequent degree of overperformance of the loser portfolios, and the significances of the $\hat{\alpha}$ terms is somewhat reduced. It is concluded, therefore, that only a part of the overreaction phenomenon occurs in January.

- Clare and Thomas then proceed to examine whether the overreaction effect is related to firm size, although the results are not presented here.

Conclusions

- The main conclusions from Clare and Thomas' study are:

  (1) There appears to be evidence of overreactions in UK stock returns, as found in previous US studies

  (2) These over-reactions are unrelated to the CAPM beta

  (3) Losers that subsequently become winners tend to be small, so that most of the overreaction in the UK can be attributed to the size effect

1.4 Determinants of sovereign credit ratings

- Sovereign credit ratings are an assessment of the riskiness of debt issued by governments

  They embody an estimate of the probability that the borrower will default on her obligation

- Two famous US ratings agencies, Moody's and Standard and Poor's, provide ratings for many governments.

  Although the two agencies use different symbols to denote the given riskiness of a particular borrower, the ratings of the two agencies are comparable.

- Gradings are split into two broad categories: investment grade and speculative grade

  Investment grade issuers have good or adequate payment capacity

  Speculative grade issuers either have a high degree of uncertainty about whether they will make their payments, or are already in default.

- The highest grade offered by the agencies, for the highest quality of payment capacity, is 'triple A', which Moody's denotes 'Aaa' and Standard and Poor's denotes 'AAA'

  The lowest grade issued to a sovereign in the Cantor and Packer sample was B3 (Moody's) or B- (Standard and Poor's)

- Thus the number of grades of debt quality from the highest to the lowest given to governments in their sample is 16

- The central aim of Cantor and Packer's paper is an attempt to explain and model how the agencies arrived at their ratings.

  Although the ratings themselves are publicly available, the models or methods used to arrive at them are shrouded in secrecy.

  The agencies also provide virtually no explanation as to what the relative weights of the factors that make up the rating are.

- Thus, a model of the determinants of sovereign credit ratings could be useful in assessing whether the ratings agencies appear to have acted rationally.

- Such a model could also be employed to try to predict the rating that would be awarded to a sovereign that has not previously been rated and when a re-rating is likely to occur.

Data

- Cantor and Packer (1996) obtain a sample of government debt ratings for 49 countries as of September 1995 that range between the above gradings.

- The ratings variable is quantified, so that the highest credit quality (Aaa/AAA) in the sample is given a score of 16, while the lowest rated sovereign in the sample is given a score of 1 (B3/B-)

  This score forms the dependent variable.

- The factors that are used to explain the variability in the ratings scores are macroeconomic variables.

  All of these variables embody factors that are likely to influence a government's ability and willingness to service its debt costs.

- Ideally, the model would also include proxies for socio-political factors, but these are difficult to measure objectively and so are not included

  It is not clear in the paper from where the list of factors was drawn.

- The included variables (with their units of measurement) are:

  1. **Per capita income** (in 1994 thousand US dollars).

  Cantor and Packer argue that per capita income determines the tax base, which in turn influences the government's ability to raise revenue.

  2. **GDP growth** (annual 1991–4 average, %).

  The growth rate of increase in GDP is argued to measure how much easier it will become to service debt costs in the future.

  3. **Inflation** (annual 1992–4 average, %).

  Cantor and Packer argue that high inflation suggests that inflationary money financing will be used to service debt when the government is unwilling or unable to raise the required revenue through the tax system.

  4. **Fiscal balance** (average annual government budget surplus as a proportion of GDP 1992–4, %).

  Again, a large fiscal deficit shows that the government has a relatively weak capacity to raise additional revenue and to service debt costs.

5. **External balance** (average annual current account surplus as a proportion of GDP 1992–4, %).

Cantor and Packer argue that a persistent current account deficit leads to increasing foreign indebtedness, which may be unsustainable in the long run.

6. **External debt** (foreign currency debt as a proportion of exports in 1994, %).

Reasoning as for external balance (which is the change in external debt over time).

7. **Dummy for economic development** (=1 for a country classified by the IMF as developed, 0 otherwise).

Cantor and Packer argue that credit ratings agencies perceive developing countries as relatively more risky beyond that suggested by the values of the other factors listed above

8. **Dummy for default history** (=1 if a country has defaulted, 0 otherwise).

It is argued that countries that have previously defaulted experience a large fall in their credit rating.

- The income and inflation variables are transformed to their logarithms. The model is linear and estimated using OLS.

- OLS is not an appropriate technique when the dependent variable can take on only one of a certain limited set of values (in this case, 1, 2, 3, . . . 16).

  In such applications, a technique such as ordered probit (covered by Mauro) would usually be more appropriate.

- Cantor and Packer argue that any approach other than OLS is infeasible given the relatively small sample size (49), and the large number (16) of ratings categories.

  The results from regressing the rating value on the variables listed above are presented in table 5.

- Four regressions are conducted, each with identical independent variables but a different dependent variable.

  Regressions are conducted for the rating score given by each agency separately, with results presented in columns (4) and (5) of table 5.

- Occasionally, the ratings agencies give different scores to a country – for example, in the case of Italy, Moody's gives a rating of 'A1', which would generate a score of 12 on a 16-scale.

  Standard and Poor's (S and P), on the other hand, gives a rating of 'AA', which would score 14 on the 16-scale, two gradings higher.

- Thus a regression with the average score across the two agencies, and with the difference between the two scores as dependent variables, is also conducted, and presented in columns (3) and (6), respectively of table 5.

- Determinants and impacts of sovereign credit ratings

Table 5.

Explanatory variable (1)	Expected sign (2)	Dependent variable Average rating (3)	Moody's rating (4)	S&P rating (5)	Difference Moody's/S&P (6)
Intercept	?	1.442 (0.663)	3.408 (1.379)	−0.524 (−0.223)	3.932** (2.521)
Per capita income	+	1.242*** (5.302)	1.027*** (4.041)	1.458*** (6.048)	−0.431*** (−2.688)
GDP growth	+	0.151 (1.935)	0.130 (1.545)	0.171** (2.132)	−0.040 (0.756)
Inflation	−	−0.611*** (−2.839)	−0.630*** (−2.701)	−0.591*** (−2.671)	−0.039 (−0.265)
Fiscal balance	+	0.073 (1.324)	0.049 (0.818)	0.097* (1.71)	−0.048 (−1.274)
External balance	+	0.003 (0.314)	0.006 (0.535)	0.001 (0.046)	0.006 (0.779)
External debt	−	−0.013*** (−5.088)	−0.015*** (−5.365)	−0.011*** (−4.236)	−0.004*** (−2.133)
Development dummy	+	2.776*** (4.25)	2.957*** (4.175)	2.595*** (3.861)	0.362 (0.81)
Default dummy	−	−2.042*** (−3.175)	−1.63** (−2.097)	−2.622*** (−3.962)	1.159*** (2.632)
Adjusted R^2		0.924	0.905	0.926	0.836

Notes: t-ratios in parentheses; *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.
Source: Cantor and Packer (1996). Reprinted with permission from *Institutional Investor*.

Interpreting the models

- The models are difficult to interpret in terms of their statistical adequacy, since virtually no diagnostic tests have been undertaken.

- The values of the adjusted R^2, at over 90% for each of the three ratings regressions, are high for cross-sectional regressions, indicating that the model seems able to capture almost all of the variability of the ratings about their mean values across the sample.

- There does not appear to be any attempt at reparameterisation presented in the paper, so it is assumed that the authors reached this set of models after some searching.

- In this particular application, the residuals have an interesting interpretation as the difference between the actual and fitted ratings.

  The actual ratings will be integers from 1 to 16, although the fitted values from the regression and therefore the residuals can take on any real value.

- Cantor and Packer argue that the model is working well as no residual is bigger than 3, so that no fitted rating is more than three categories out from the actual rating, and only four countries have residuals bigger than two categories.

- Furthermore, 70% of the countries have ratings predicted exactly (i.e. the residuals are less than 0.5 in absolute value).

- Now, turning to interpret the models from a financial perspective, it is of interest to investigate whether the coefficients have their expected signs and sizes.

  The expected signs for the regression results of columns (3)–(5) are displayed in column (2) of table 5 (as determined by this author).

- As can be seen, all of the coefficients have their expected signs, although the fiscal balance and external balance variables are not significant or are only very marginally significant in all three cases.

- The coefficients can be interpreted as the average change in the rating score that would result from a unit change in the variable.

  So, for example, a rise in per capita income of $1,000 will on average increase the rating by 1.0 units according to Moody's and 1.5 units according to Standard & Poor's.

The development dummy suggests that, on average, a developed country will have a rating three notches higher than an otherwise identical developing country.

And everything else equal, a country that has defaulted in the past will have a rating two notches lower than one that has always kept its obligation.

- By and large, the ratings agencies appear to place similar weights on each of the variables, as evidenced by the similar coefficients and significances across columns (4) and (5) of table 5.

  This is formally tested incolumn (6) of the table, where the dependent variable is the difference between Moody's and Standard and Poor's ratings.

- Only three variables are statistically significantly differently weighted by the two agencies.

- Standard & Poor's places higher weights on income and default history, while Moody's places more emphasis on external debt.

- The paper continues, among other things, to consider whether **ratings add to publicly available information**, and whether it is possible to determine **what factors affect how the sovereign yields react to ratings announcements**.

The relationship between ratings and yields

- Cantor and Packer try to determine whether ratings have any additional information useful for modelling the cross-sectional variability of sovereign yield spreads over and above that contained in publicly available macroeconomic data.

- The dependent variable is now the log of the yield spread, i.e. $\ln(Yield\ on\ the\ sovereign\ bond - Yield\ on\ a\ US\ treasury\ bond)$

  One may argue that such a measure of the spread is imprecise, for the true credit spread should be defined by the entire credit quality curve rather than by just two points on it.

- Three regressions are presented in table 6, denoted specifications (1), (2) and (3).

- **Specification (1)** is a regression of the $\ln(spread)$ on only a constant and the average rating (column (1))

  Column (1) shows that ratings have a highly significant inverse impact on the spread.

- **Specification (2)** is a regression of the $\ln(spread)$ on the macroeconomic variables used in the previous analysis.

  The expected signs are given (as determined by this author) in column (2).

  As can be seen, all coefficients have their expected signs, although now only the coefficients belonging to the external debt and the two dummy variables are statistically significant.

- **Specification (3)** is a regression on both the average rating and the macroeconomic variables.

  When the rating is included with the macroeconomic factors, none of the latter is any longer significant – only the rating coefficient is statistically significantly different from zero.

- This message is also portrayed by the adjusted R^2 values, which are highest for the regression containing only the rating, and slightly lower for the regression containing the macroeconomic variables and the rating.

- One may also observe that, under specification (3), the coefficients on the per capita income, GDP growth and inflation variables now have the wrong sign.

  This is, in fact, never really an issue, for if a coefficient is not statistically significant, it is indistinguishable from zero in the context of hypothesis testing, and therefore it does not matter whether it is actually insignificant and positive or insignificant and negative.

- Only coefficients that are both of the wrong sign and statistically significant imply that there is a problem with the regression.

  It would thus be concluded from this part of the paper that there is no more incremental information in the publicly available macroeconomic variables that is useful for predicting the yield spread than that embodied in the rating.

  The information contained in the ratings encompasses that contained in the macroeconomic variables.

- Do ratings add to public information?

  Table 6.

		Dependent variable: ln (yield spread)		
Variable	Expected sign	(1)	(2)	(3)
Intercept	?	2.105*** (16.148)	0.466 (0.345)	0.074 (0.071)
Average rating	−	−0.221*** (−19.175)		−0.218*** (−4.276)
Per capita income	−		−0.144 (−0.927)	0.226 (1.523)
GDP growth	−		−0.004 (−0.142)	0.029 (1.227)
Inflation	+		0.108 (1.393)	−0.004 (−0.068)
Fiscal balance	−		−0.037 (−1.557)	−0.02 (−1.045)
External balance	−		−0.038 (−1.29)	−0.023 (−1.008)
External debt	+		0.003*** (2.651)	0.000 (0.095)
Development dummy	−		−0.723*** (−2.059)	−0.38 (−1.341)
Default dummy	+		0.612*** (2.577)	0.085 (0.385)
Adjusted R^2		0.919	0.857	0.914

Notes: *t*-ratios in parentheses; *, **and *** indicate significance at the 10%, 5% and 1% levels, respectively.
Source: Cantor and Packer (1996). Reprinted with permission from *Institutional Investor*.

What determines how the market reacts to ratings announcements?

- Cantor and Packer also consider whether it is possible to build a model to predict how the market will react to ratings announcements, in terms of the resulting change in the yield spread.

- The dependent variable for this set of regressions is now the change in the log of the relative spread, i.e. log[(yield – treasury yield)/treasury yield], over a two-day period at the time of the announcement.

- The sample employed for estimation comprises every announcement of a ratings change that occurred between 1987 and 1994

  79 such announcements were made, spread over 18 countries.

  39 of these were actual ratings changes by one or more of the agencies, and 40 were listed as likely in the near future to experience a regrading.

- Moody's calls this a 'watchlist', while Standard and Poor's term it their 'outlook' list.

- The explanatory variables are mainly dummy variables for:

  1. whether the announcement was positive – i.e. an upgrade

  2. whether there was an actual ratings change or just listing for probable regrading

  3. whether the bond was speculative grade or investment grade

  4. whether there had been another ratings announcement in the previous 60 days

  5. the ratings gap between the announcing and the other agency.

- The following cardinal variable was also employed:

  the change in the spread over the previous 60 days.

- The results are presented in table 7, but in this text, only the final specification (numbered 5 in Cantor and Packer's exhibit 11) containing all of the variables described above is included.

- As can be seen from table 4.4, the models appear to do a relatively poor job of explaining how the market will react to ratings announcements.

  The adjusted R^2 value is only 12%, and this is the highest of the five specifications tested by the authors.

  Further, only two variables are significant and one marginally significant of the seven employed in the model.

- It can therefore be stated that yield changes are significantly higher following a ratings announcement for speculative than investment grade bonds, and that ratings changes have a bigger impact on yield spreads if there is an agreement between the ratings agencies at the time the announcement is made.

  Further, yields change significantly more if there has been a previous announcement in the past 60 days than if not.

- On the other hand, neither whether the announcement is an upgrade or downgrade, nor whether it is an actual ratings change or a name on the watchlist, nor whether the announcement is made by Moody's or Standard & Poor's, nor the amount by which the relative spread has already changed over the past 60 days, has any significant impact on how the market reacts to ratings announcements.

- What determines reactions to ratings announcements?

Table 7.

Dependent variable: log relative spread	
Independent variable	Coefficient (t-ratio)
Intercept	−0.02 (−1.4)
Positive announcements	0.01 (0.34)
Ratings changes	−0.01 (−0.37)
Moody's announcements	0.02 (1.51)
Speculative grade	0.03** (2.33)
Change in relative spreads from day −60 to day −1	−0.06 (−1.1)
Rating gap	0.03* (1.7)
Other rating announcements from day −60 to day −1	0.05** (2.15)
Adjusted R^2	0.12

Note: * and ** denote significance at the 10% and 5% levels, respectively.
Source: Cantor and Packer (1996). Reprinted with permission from *Institutional Investor*.

- Conclusions

1. To summarise, six factors appear to play a big role in determining sovereign credit ratings – incomes, GDP growth, inflation, external debt, industrialised or not and default history

2. The ratings provide more information on yields than all of the macroeconomic factors put together

3. One cannot determine with any degree of confidence what factors determine how the markets will react to ratings announcements.

1.5 Covered and uncovered interest parity

- The determination of the price of one currency in terms of another (i.e. the exchange rate) has received a great deal of empirical examination in the international finance literature.

- Three hypotheses in particular are studied – covered interest parity (**CIP**), uncovered interest parity (**UIP**) and purchasing power parity (**PPP**).

  Violation of one or more of the parities may offer the potential for arbitrage, or at least will offer further insights into how financial markets operate.

  See Cuthbertson and Nitsche (2004) or the many references therein for a more comprehensive treatment

- **Covered Interest Parity**

  CIP implies that, if financial markets are efficient, it should not be possible to make a riskless profit by borrowing at a risk-free rate of interest in a domestic currency, switching the funds borrowed into another (foreign) currency, investing them there at a risk free rate and locking in a forward sale to guarantee the rate of exchange back to the domestic currency

- Thus, if CIP holds, it is possible to write

  $f_t - s_t = (r - r^*)_t$

  where

  f_t and s_t are the log of the forward and spot prices of the domestic in terms of the foreign currency at time t,

  r is the domestic interest rate and

  r^* is the foreign interest rate.

- This is an equilibrium condition which must hold otherwise there would exist riskless arbitrage opportunities, and the existence of such arbitrage would ensure that any deviation from the condition cannot hold indefinitely.

  It is worth noting that, underlying CIP are the assumptions that the risk-free rates are truly risk-free – that is, there is no possibility for default risk.

- It is also assumed that there are no transactions costs, such as broker's fees, bid–ask spreads, stamp duty, etc., and that there are no capital controls, so that funds can be moved without restriction from one currency to another

- **Uncovered Covered Interest Parity**

  UIP takes CIP and adds to it a further condition known as 'forward rate unbiasedness' (FRU). Forward rate unbiasedness states that the forward rate of foreign exchange should be an unbiased predictor of the future value of the spot rate.

- If this condition does not hold, again in theory riskless arbitrage opportunities could exist.

  UIP, in essence, states that the expected change in the exchange rate should be equal to the interest rate differential between that available risk-free in each of the currencies.

- Algebraically, this may be stated as

  (3) $s^e_{t+1} - s_t = (r - r^*)_t$

  where the notation is as above and s^e_{t+1} is the expectation, made at time t of the spot exchange rate that will prevail at time $t + 1$.

- Tests of CIP unsurprisingly (for it is a pure arbitrage condition) tend not to reject the hypothesis that the condition holds

  Taylor (1987, 1989) has conducted extensive examinations of CIP, and concluded that there were historical periods when arbitrage was profitable, particularly during periods where the exchange rates were under management.

- Relatively simple **tests of UIP and FRU** take equations of the form (3) and add intuitively relevant additional terms.

  If UIP holds, these additional terms should be insignificant.

- Ito (1988) tests UIP for the yen/dollar exchange rate with the three-month forward rate for January 1973 until February 1985.

  The sample period is split into three as a consequence of perceived structural breaks in the series.

  Strict controls on capital movements were in force in Japan until 1977, when some were relaxed and finally removed in 1980.

  A Chow test confirms Ito's intuition and suggests that the three sample periods should be analysed separately.

- **Two separate regressions** are estimated for each of the three sample sub-periods

  (4) $s_{t+3} - f_{t,3} = a + b_1(s_t - f_{t-3,3}) + b_1(s_{t-1} - f_{t-4,3}) + u_t$

  where

  s_{t+3} is the spot interest rate prevailing at time $t + 3$

  $f_{t,3}$ is the forward rate for three periods ahead available at time t

  u_t is an error term.

- A natural joint hypothesis to test is $H_0 : a = 0$ and $b_1 = 0$ and $b_2 = 0$

  This hypothesis represents the restriction that the deviation of the forward rate from the realised rate should have a mean value insignificantly different from zero ($a = 0$) and it should be independent of any information available at time t ($b_1 = 0$ and $b_2 = 0$).

  All three of these conditions must be fulfilled for UIP to hold.

- The second equation that Ito tests is

  (5) $s_{t+3} - f_{t,3} = a + (s_t - f_{t,3}) + \upsilon_t$

  where υ_t is an error term and the hypothesis of interest in this case is $H_0 : a = 0$ and $b = 0$.

- **Equation (4)** tests whether past forecast errors have information useful for predicting the difference between the actual exchange rate at time $t+3$, and the value of it that was predicted by the forward rate at time t.

- **Equation (5)** tests whether the forward premium has any predictive power for the difference between the actual exchange rate at time $t + 3$, and the value of it that was predicted by the forward rate at time t..

- The results for the three sample periods are presented in Ito's table 3, and are adapted and reported here in table 8.

- Uncovered interest parity test results

Table 8.

Sample period	1973M1–1977M3	1977M4–1980M12	1981M1–1985M2
Panel A: Estimates and hypothesis tests for $S_{t+3} - f_{t,3} = a + b_1(s_t - f_{t-3,3}) + b_2(s_{t-1} - f_{t-4,3}) + u_t$			
Estimate of a	0.0099	0.0031	0.027
Estimate of b_1	0.020	0.24	0.077
Estimate of b_2	−0.37	0.16	−0.21
Joint test $\chi^2(3)$	23.388	5.248	6.022
P-value for joint test	0.000	0.155	0.111
Panel B: Estimates and hypothesis tests for $S_{t+3} - f_{t,3} = a + b(s_t - f_{t,3}) + v_t$			
Estimate of a	0.00	−0.052	−0.89
Estimate of b	0.095	4.18	2.93
Joint test $\chi^2(2)$	31.923	22.06	5.39
p-value for joint test	0.000	0.000	0.07

Source: Ito (1988). Reprinted with permission from MIT Press Journals.

Conclusion

- **The main conclusion is that UIP clearly failed to hold throughout the period of strictest controls, but there is less and less evidence against UIP as controls were relaxed.**

References

- Chris Brooks, Introductory Econometrics for Finance, Cambridge University Press, Second Edition. **Chapters 1-5**
- Jensen, M. C. (1968) The Performance of Mutual Funds in the Period 1945–1964, Journal of Finance 23, 389–416
- Cantor, R. and Packer, F. (1996) Determinants and Impacts of Sovereign Credit Ratings, Journal of Fixed Income 6, 76–91
- Clare, A. D. and Thomas, S. H. (1995) The Overreaction Hypothesis and the UK Stock Market, Journal of Business Finance and Accounting 22(7), 961–73
- Ito, T. (1988) Use of (Time-Domain) Vector Autoregressions to Test Uncovered Interest Parity, Review of Economics and Statistics 70(2), 296–305