Lecture 4 Fama and Macbeth Regression Analysis

Tao Zeng

SOE & AFR, Zhejiang University

Nov 1, 2023

FM Regression

- Portfolio analysis is a **nonparametric** technique which does not make any assumptions about the nature of the relation between the variables.
- Portfolio analysis is difficult to include a large set of controls when examining the relation.
- (Fama and MacBeth 1973, FM hereafter) regression analysis is able to control for a large set of other variables when examining the relation of interest.
- FM makes the assumption that the relation between each control variable and the outcome variable of interest is **linear**.

FM Regression

- FM regression analysis is implemented using a two-step procedure:
  - Run periodic cross-sectional regressions of the dependent variable of interest, which we denote Y, on one or more independent variables $X1, X2$, etc., using data from each time period t.
    - Get slope coefficients and an intercept coefficient on each independent variable for each period.
  - Analyze the time series of each of the regression coefficients to determine whether the average coefficient differs from zero.
- Can we interpret the slope coefficient as factor return or risk premium? **NO.**
- **But the average coefficient differs from zero is the same as the average factor risk premium differs from zero.**

Periodic Cross-Sectional Regressions

- The cross-sectional regression specification at time period t is

$$Y_{i,t} = \delta_{0,t} + \delta_{1,t} X1_{i,t} + \delta_{2,t} X2_{i,t} + \cdots + \epsilon_{i,t}.$$

- The result is a time series of intercept and slope coefficients $\delta_{0,t}, \delta_{1,t}, \delta_{2,t}$, R-squared, adjusted R-squared, and number of observations, etc.
- The type of cross-sectional regression used here can be ordinary-least-squares (OLS) regression, weighted-least-squares regression or even a logistic or probit regression if the dependent variable is discrete.
- The full specification of the cross-sectional regression is

$$r_{i,t+1} = \delta_{0,t} + \delta_{1,t}\beta_{i,t} + \delta_{2,t} Size_{i,t} + \delta_{3,t} BM_{i,t} + \epsilon_{i,t+1}.$$

where $r_{i,t+1}$ is the one-year-ahead excess return of stock i.

Periodic Cross-Sectional Regressions

- Periodic FM Regression Results

Panel A I

$$r_{i,t+1} = \delta_{0,t} + \delta_{1,t}\beta_{i,t} + \epsilon_{i,t}$$

$t/t+1$	$\delta_{0,t}$	$\delta_{1,t}$	R_t^2	Adj. R_t^2	n_t
1988/1989	1.68	5.53	0.002	0.002	5646
1989/1990	−29.52	2.28	0.001	0.000	5470
1990/1991	41.11	11.66	0.002	0.002	5360
1991/1992	36.36	−17.58	0.008	0.008	5265
1992/1993	28.09	−9.24	0.009	0.009	5353
1993/1994	−5.21	−0.52	0.000	−0.000	5634
1994/1995	25.42	2.87	0.001	0.000	6108
1995/1996	18.27	−5.41	0.004	0.004	6234
1996/1997	31.02	−17.29	0.024	0.024	6528
1997/1998	−9.34	4.41	0.001	0.001	6796
1998/1999	1.49	45.88	0.017	0.017	6520

Periodic Cross-Sectional Regressions

- Periodic FM Regression Results

Panel A II					
$t/t+1$	$\delta_{0,t}$	$\delta_{1,t}$	R_t^2	Adj. R_t^2	n_t
1999/2000	−0.85	−14.61	0.010	0.010	6036
2000/2001	37.13	−24.58	0.028	0.028	5817
2001/2002	7.81	−27.26	0.114	0.114	5449
2002/2003	72.17	2.28	0.000	−0.000	5038
2003/2004	28.59	−13.70	0.018	0.018	4698
2004/2005	7.83	−5.93	0.008	0.008	4537
2005/2006	12.73	−1.54	0.000	0.000	4466
2006/2007	−12.06	4.39	0.004	0.004	4412
2007/2008	−41.07	−0.76	0.000	−0.000	4310
2008/2009	62.76	−1.42	0.000	−0.000	4229
2009/2010	18.21	7.70	0.008	0.008	3949
2010/2011	0.68	−6.91	0.009	0.009	3782
2011/2012	25.92	−5.87	0.003	0.003	3661

Periodic Cross-Sectional Regressions

- Periodic FM Regression Results

Panel D I

$$r_{i,t+1} = \delta_{0,t} + \delta_{1,t}\beta_{i,t} + \delta_{2,t}\text{ Size }_{i,t} + \delta_{3,t}BM_{i,t} + \epsilon_{i,t}$$

$t/t+1$	$\delta_{0,t}$	$\delta_{1,t}$	$\delta_{2,t}$	$\delta_{3,t}$	R_t^2	Adj. R_t^2	n_t
1988/1989	−11.10	−1.18	3.60	4.65	0.016	0.015	4301
1989/1990	−28.76	0.21	0.85	−0.34	0.002	0.001	4239
1990/1991	80.00	25.11	−10.50	−9.59	0.024	0.023	4176
1991/1992	42.12	−7.59	−5.53	6.15	0.025	0.025	4176
1992/1993	40.71	−3.83	−4.56	3.48	0.024	0.023	4166
1993/1994	−9.61	2.46	−0.70	7.93	0.013	0.012	4464
1994/1995	28.95	4.12	−1.53	1.92	0.002	0.001	4826
1995/1996	12.87	−4.09	−0.82	10.97	0.013	0.012	5009
1996/1997	6.39	−12.53	2.06	11.34	0.024	0.024	5203
1997/1998	−21.03	2.24	1.83	7.34	0.006	0.005	5475
1998/1999	69.93	61.01	−14.24	−9.63	0.036	0.035	5288

Periodic Cross-Sectional Regressions

- Periodic FM Regression Results

Panel D II							
$t/t+1$	$\delta_{0,t}$	$\delta_{1,t}$	$\delta_{2,t}$	$\delta_{3,t}$	R_t^2	Adj. R_t^2	n_t
1999/2000	−31.70	−23.89	5.93	5.00	0.029	0.029	4885
2000/2001	70.83	−20.93	−6.95	−3.40	0.044	0.043	4617
2001/2002	−2.91	−22.14	−0.11	3.73	0.091	0.091	4444
2002/2003	152.96	30.13	−18.95	−0.46	0.062	0.061	4097
2003/2004	30.12	−15.71	−0.38	4.03	0.032	0.031	3831
2004/2005	−7.37	−7.93	2.52	6.35	0.022	0.022	3697
2005/2006	12.72	−0.50	−0.71	6.72	0.005	0.004	3893
2006/2007	−23.73	1.09	2.68	−0.24	0.013	0.013	3849
2007/2008	−49.72	−3.17	1.91	−0.42	0.007	0.007	3740
2008/2009	133.77	24.19	−17.69	8.21	0.029	0.029	3699
2009/2010	24.38	9.40	−1.49	1.17	0.014	0.013	3516
2010/2011	−15.02	−11.60	3.33	0.59	0.032	0.031	3372
2011/2012	16.92	−1.90	−0.28	8.33	0.012	0.011	3298

Average Cross-Sectional Regression Results

- Compute the time-series averages of the periodic cross-sectional regression coefficients and other regression results (R-squared, adjusted R-squared, and number of observations).
- Examine whether the average coefficient is statistically different from zero.
  - Calculate the standard errors and the associated t-statistics and p-values to test the null hypothesis that the average coefficient is equal to zero.

Average Cross-Sectional Regression Results

- Summarized FM Regression Results

Coefficient	Value	(1)	(2)	(3)	(4)
δ_0	Average	14.97	23.23	9.83	21.74
	Standard error	2.70	4.36	1.65	4.58
	t-statistic	5.55	5.32	5.94	4.75
	p-value	0.00	0.00	0.00	0.00
δ_1	Average	-2.73			0.96
	Standard error	1.86			1.68
	t-statistic	**-1.47**			**0.57**
	p-value	0.16			0.57
δ_2	Average		-2.29		-2.49
	Standard error		0.62		0.57
	t-statistic		**-3.69**		**-4.37**
	p-value		0.00		0.00

Average Cross-Sectional Regression Results

- Summarized FM Regression Results

Coefficient	Value	(1)	(2)	(3)	(4)
δ_3	Average			4.77	3.08
	Standard error			0.73	0.79
	t-statistic			**6.51**	**3.89**
	p-value			0.00	0.00
R^2		0.011	0.010	0.005	0.024
Adj.R^2		0.011	0.009	0.005	0.023
n		5221	5584	4270	4261

Interpreting FM Regressions

- A statistically significant average slope coefficient indicates a cross sectional relation between the given independent variable X and the dependent variable Y in the average time period.
- When the regression specification includes more than one independent variable, statistical significance indicates that a relation between X and Y exists after controlling for the effects of the other independent variables included in the regression specification.
- If the coefficient of interest is statistically significant in one specification but insignificant when additional controls are added to the specification, then the relation between X and Y appears to be explained by some linear combination of the added control variables.
- If a statistically significant relation appears after including additional controls, this indicates that it is necessary to control for other effects captured by the newly added control variables in order to detect the relation of interest.

Interpreting FM Regressions

- There is no indication of a relation between β and future stock returns, as the average coefficient on $\beta\,(\delta_1)$ is statistically indistinguishable from zero in all specifications that include β as an independent variable.
- The average slope on β $((\delta_1))$ in the univariate regression is -2.73 with a corresponding *t*-statistic of -1.47 and 0.96 with corresponding *t*-statistic of 0.57 in the full specification.

Interpreting FM Regressions

- There exists a strong negative relation between *Size* and future stock returns, which is robust to the inclusion of β and *BM* as independent variables.
- The average slope on *Size* ((δ_2)) in the univariate regression is -2.29 with a corresponding t-statistic of -3.69 and -2.49 with corresponding t-statistic of -4.37 in the full specification.

Interpreting FM Regressions

- There exists a strong positive relation between *BM* and future stock returns which is also robust to the inclusion of β and *Size* as independent variables.
- The average slope on *BM* ((δ_3)) is 4.77 with a corresponding *t*-statistic of 6.51 in the univariate regression and 3.08 with corresponding *t*-statistic of 3.89 in the full specification.
- The average *R*-squared and adjusted *R*-squared values that include all three variables are 0.024 and 0.023, respectively, indicating that only a little more than 2% of the total cross-sectional variation in future stock returns is explained by β, Size, and BM.