Asset Pricing Tests, Endogeneity issues and Fama-French factors

Allen, David

University of Sydney, Asia University, Edith Cowan University

30 June 2022

Asset Pricing Tests, Endogeneity issues and Fama-French factors

David E. Allen[a,*]

[a] *School of Mathematics and Statistics, University of Sydney, Department of Finance, Asia University, Taiwan, and School of Business and Law, Edith Cowan University, Australia*

Abstract

This paper features a statistical analysis of the independence of the core Fama/French factors; SMB and HML, using daily data, of the factor return series, for the USA, Developed Markets and Japan, using a sample taken from the data-sets that are available on French's website. The various series and their inter-relationships are analysed using rolling OLS regressions, so as to explore their independence and issues related to their endogeneity. The OLS analysis incorporates Ramsey's RESET tests of functional form misspecification. The empirical results suggest that these factors, when combined in OLS regression analysis, as suggested by Fama and French (2018), and generally in the empirical asset pricing literature featuring time-series tests, are frequently not independent, and thus likely to suffer from endogeneity. The rolling regression analysis suggests significant and time-varying relationships between the core factors and rejects their independence for long periods of time within the samples. A significant non-linear relationship exists between some of the series, as indicated by the exployment of squared terms, which are frequently significant. The empirical results suggest that using these factors in linear regression analysis, such as suggested by Fama and French (2018), as a method of screening factor relevance, is likely to be problematic, in that the estimated standard errors are likely to be sensitive to the non-independence of factors. This is also likely to be a potential problem for asset pricing tests that use the popular time-series approach, as first suggested by Fama and Macbeth (1973).

Keywords: Fama-French Factors, Correct specification, Ramsey's RESET, Endogeneity, Strong Endogeneity, Consistent standard errors
JEL Codes: C13, C14, G12.

---

*Corresponding author
Email address: profallen2007@gmail.com (David E. Allen)

June 30, 2022

1. Introduction

In a fundamental paper, Fama and French (1993, p3), stated that: "there are three stock-market factors: an overall market factor and factors related to firm size and book-to-market equity". French generously provides estimates of these original factors, and more recently suggested additions, on his personal website (see http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/f-f_factors.html). The original 1993 paper triggered the development of a virtual global industry in the testing for the effects of various factors on various portfolios selected from global markets.

Cochrane (2011, p.1047), in a Presidential Address to the American Finance Association, observed that: "we also thought that the cross-section of expected returns came from the CAPM. Now we have a zoo of new factors." Harvey, Liu, and Zhu (2015) list 316 anomalies proposed as potential factors in asset-pricing models, and comment that there are others that do not make their list.

Fama and French (2018) propose a method for screening competing factors, and explain that previous approaches can be described under two main headings. The left-hand-side (LHS) approach judges competing models on the intercepts (unexplained average returns) left in time series regressions to explain excess returns on sets of LHS portfolios. A drawback is that different sets of LHS portfolios can lead to different intercepts and, therefore, to different inferences.

An alternative right-hand-side (RHS) approach uses spanning regressions to judge whether individual factors contribute to the explanation of average returns provided by a model. Each candidate factor is regressed on the model's other factors. If the intercept in a spanning regression is non-zero, the factor adds to the model's explanation of average returns in that sample period. Fama and French (2018) note that the GRS statistic of Gibbons, Ross, and Shanken (GRS 1989), hereafter GRS, produces a test of whether multiple factors add to a base model's explanatory power.

The GRS test test is based on the strong assumptions of linearity, independence and a Gaussian distribution. GRS on the assumption that there is a given riskless rate of interest, R_{ft}, for each time period. Excess returns are computed by subtracting R_{ft}, from the total rates of return. GRS consider the following multivariate linear regression:

$$\tilde{r}_{it} = \alpha_{ip} + \beta_{ip}\tilde{r}_{pt} + \tilde{\varepsilon}_{it} \quad \forall i = 1, ..., N, \tag{1}$$

where $\tilde{r}_{it} \equiv$ excess return on asset i in period t, $\tilde{r}_{pt} \equiv$ excess return on the portfolio whose efficiency is being tested, and $\tilde{\epsilon}_{it} \equiv$ disturbance term for asset i in period t. The disturbances are assumed to be jointly normally distributed in each period, with mean zero and nonsingular covariance matrix $\sum$, conditional on the excess returns for portfolio p. They also assume independence of the disturbances over time. In order that $\sum$ be non-singular, $\tilde{r}_{pt}$ and the N left-hand-side assets must be linearly independent.

GRS suggest that if a particular portfolio is mean-variance efficient, (that is it minimizes variance for a given level of expected return), then the following first-order condition must be satisfied for the given N assets:

$$E(\tilde{r}_{it}) = \beta_{ip} E(\tilde{r}_{pt}). \tag{2}$$

GRS combine the first-order condition in (2) with the distributional assumption suggested by (1), and obtain the following parametric restriction, which they state in the form of a null hypothesis:

$$H_o \ a_{ip} = 0, \ \ \forall i = 1, ..., N. \tag{3}$$

The GRS test is based on a null hypothesis that the intercept in the above regression, as shown in expressions (1) and (2), is zero. There are several assumptions required for this test to be valid, namely linearity, independence, and Gaussian distributions.

Fama and French (2018) adopt a test proposed by Barillas and Shanken (2016). Barillas and Shanken (2016) assume that the factors of competing models are among the LHS returns that each model is supposed to explain. Formally, let R be the target set of non-factor LHS excess returns, f_i the factors of model i, and F_{Ai} the union of the factors of model i's competitors. In the BS approach, the set of LHS returns for model i, Π_i, combines R and F_{Ai}, with linearly dependent components deleted. Competing models are assessed on the maximum (max) squared Sharpe ratio for the intercepts from time series regressions of LHS returns on a model's factors.

Define a_i as the vector of intercepts from regressions of Π_i on f_i, and $\sum_i$ as the residual covariance matrix. The maximum squared Sharpe ratio for the intercepts is given by:

$$Sh^2 a_i = a_i^{'} {\sum}_i^{-1} a^i, \tag{4}$$

and the superior model is judged to be the one with the smallest $Sh^2 a_i$.

Gibbons et al. (1989) show that $a_i^{'} \sum_i^{-1} a^i$, is the difference between the max squared Sharpe ratio constructed from f_i and Π_i together, and the max for f_i individually:

$$Sh^2 a_i = Sh^2 \Pi_i f_i - Sh^2 f_i. \tag{5}$$

Fama and French (2018) suggest that since Π_i includes the factors of all model i's competitors, the union of Π_i and f_i, which they call Π, does not depend on i. This means that equation (5) can be simplified to:

$$Sh^2 a_i = Sh^2 \Pi_i f_i - Sh^2 f_i \tag{6}$$

Fama and French (2018) assume that R is the target set of non-factor LHS excess returns, and that the best model is the one which produces the highest $Sh^2 f$. They suggest that there is bias when comparing non-nested models, and conduct a bootstrap simulation of in - and out - of - sample results to compensate. What Fama and French (2018) do not mention is a potential problem with endogeneity of the RHS variables that is integral to their suggested metric.

Fama and French (2020) advance and refine their argument by comparing the cross-section regression approach of Fama and MacBeth (1973) to construct cross-section factors corresponding to the time-series factors of Fama and French (2015). They suggest that time-series models that use only cross-section factors provide better descriptions of average returns than time-series models that use time-series factors.

Fama and French (2020) suggest that Fama and Macbeth (1973) cross-section regressions are a type of factor model, and write the cross-section regression of stock returns for month t, R_{it}, $i = 1, ..., n$, on observed values of size (MC_{it-1}),the book-to-market ratio (BM_{it-1}), operating profitability (OP_{it-1}),and the rate of growth of assets INV_{it-1}).

$$R_{it} = R_{zt} + R_{MCt}MC_{it-1} + R_{BMt}BM_{it-1} + R_{OPt}OP_{it-1} + R_{INVt}INV_{it-1} + e_{it}. \quad (7)$$

They suggest that the slope estimates in Equation (1) are portfolio returns that, as indicated by the notation, can be interpreted as being factors.

Fama (1976, ch. 9) shows that the slope for each variable in an Fama Macbeth (1973) cross-section regression is the return on a portfolio of the left-hand-side (LHS) assets with weights for the assets that set the month t -1 portfolio value of that variable to one and zero out other explanatory variables. The intercept in an Fama Macbeth (1973) cross-section regression (R_{zt} in (1)) is the month t return on a standard portfolio of the LHS assets with weights that sum to one and zero out each explanatory variable.

They further suggest that when the cross-section regression in Equation(1) is stacked across t, it becomes an asset pricing model that can be used in time-series applications. In this perspective, it is natural to move R_{zt} to the left side of the equation so LHS returns are in excess of R_{zt}. This is shown in equation (8):

$$R_{it} - R_{zt} = RMC_{it-1} + R_{BMt}BM_{it-1} + R_{OPt}OP_{it-1} + R_{INVt}INV_{it-1} + e_{it}. \quad (8)$$

Equation (8) is a four-factor model in which four factors used to explain asset returns in excess of R_{zt}. They use Equation (8) as a time-series model (model, not regression) to describe average returns for a wide range of left-hand-side assets, and they compare the performance of (8) in this task to that of the model of FF (2015) that uses time-series factors. This approach can be written as:

$$R_{it} - R_{ft} = a_i + b_i(R_{mt} - R_{ft}) + s_iSMB_t + h_iHML_t + r_iRMW_t + c_iCMA_t + e_{it}. \quad (9)$$

Equation (9) represents a five factor model.

R_{ft} is the risk-free rate (one-month U.S. Treasury bill rate observed at the beginning of month t), and R_{mt} is the value-weight (VW) stock market return for month t. The remaining four factors are differences between returns on diversified portfolios of small and big stocks (SMB_t), high and low BM stocks(HML_t), stocks with robust and weak profitability (RMW_t), and stocks of low and high investment firms (CMA_t), conservative minus aggressive). The intercept a_i is the pricing error for LHS asset i in the time-series regression (9). The average across t of the residual e_{it} in model (8) is the pricing error for asset i.

Fama and French (2020) point out that there are important differences between Equations (8) and (9). In the time-series regression (9) the factors are prespecified. In equation (9), a least squares time-series regression optimizes an asset's factor loadings on the prespecified factors, subject to the constraint that the factor loadings are constant and assuming the disturbances in (9) are independent and identically distributed (iid) across time. In short, the time-series regression (9) optimizes loadings on factors that are not themselves optimized.

In this paper I concentrate on the time-series approach to asset pricing tests, as featured in equation (9). I limit my attention to the SMB and HML factors, given that there are more extensive data sets featuring these two 'original' factors on French's website. I apply simple tests of endogeneity, contemporaneously and with lags, by examining the independence of factors in sets of daily data, taken from Kenneth French's website, featuring the Fama/French estimates of the excess return on the market portfolio, and estimates of SMB and HML. I also explore whether the factors are statistically related in a linear fashion.

The paper is divided into four section sections, this introduction is followed by section 2, which introduces the data and statistical and econometric methods employed, section 3 presents the results and section four concludes.

2. Research Methods

Allen and McAleer (2018), suggest that the method proposed for 'choosing factors', by Fama and French (2018), is likely to suffer from an endogeneity problem and recommend the use of instrumental variables to address this issue. The presence of endogeneity in the regressors causes OLS estimators to be biased and inconsistent. Endogeneity may be the result of measurement error, reverse casualty/simultaneity, omitted variable or unobserved variables, omitted selection, and lagged dependent variables. These are the main reasons why the R.H.S. regressor and the error term may be correlated. Wu (1973) discusses issues related to the problem of endogeneity, and Hausman (1978) discusses various specification tests and the use of instruments to address the problem.

Anatolyev and Mikusheva (2022) further explore the problems associated with estimating risk premia in unconditional linear factor pricing models. They suggest that, typically, the data used in the empirical literature are characterized by weakness of some pricing factors, strong cross-sectional dependence in the errors, and (moderately) high cross-sectional dimensionality. They posit that the conventional two-pass estimation procedure delivers inconsistent estimates of the risk premia and propose a new estimation procedure based on sample-splitting instrumental variables regression.

The previous section mentioned the two-pass estimation procedure, Fama and MacBeth (1973) in which the first pass regression estimates risk exposures (betas) for each asset, and then, at the second pass, those estimates are used as regressors to estimate the risk premia. It was also mentioned that asymptotic justification of this procedure, frequently relies on assumptions that often do not hold up in realistic circumstances. The empirical analysis in this paper provides evidence of the violation of these assumptions in the Fama French data sets, as made available on French's website.

Anatolyev and Mikusheva (2022) note that two types of violations of the idealistic setting have been noted in previous literature and, as the first, they mention the problem of weak (but priced) observed factors. Kan and Zhang (1999) examine circumstances where a factor is useless, defined as being independent of all the asset returns, and provide theoretical results and simulation evidence that the second-pass cross-sectional regression tends to find the beta risk of the useless factor priced more often than it should. Raponi et al. (2020) remarked that risk exposures (or betas) to some observed factors tend to be small to such an extent that their estimation errors are of the same order of magnitude as the betas themselves and report that firm characteristics are found to explain a much larger proportion of variation in estimated expected returns than betas.

The second issue, according to Anatolyev and Mikusheva (2022), is the problem related to strong cross-sectional dependence in error terms, which in many cases can be modeled as a factor structure with an unaccounted or non-included factor, as suggested by Kleibergen (2009) and Kleibergen and Zhan (2015). Anatolyev and Mikusheva (2022) demonstrate that within a dimension-asymptotic framework the presence of small betas leads to a failure of the classical two-pass procedure, while in addition, the presence of missing factors exacerbates the problem.

Onatski (2015), provides asymptotic approximations to the squared error of the least squares estimator of the common component in large approximate factor models with a possibly misspecified number of factors. The approximations are derived under both strong and weak factors asymptotics assuming that the cross-sectional and temporal dimensions of the data are comparable. He employs simulations and obtains results that suggest that the consistency under the weak factors asymptotics requires either no cross-sectional or no temporal correlation in the idiosyncratic terms. The assessment of the relationship between the base factors is the focus of attention in this paper. In the next section I explore the behaviour of these using factor data taken directly from French's website.

3. Endogeneity tests on Fama-French 3 factors

3.1. Some preliminaries

The factor data sets used feature daily three factor Fama/French return series, for the USA, Developed Markets and Japan, taken from Ken French's website. (See: https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html).

The Fama/French factors are constructed using the 6 value-weight portfolios formed on size and book-to-market.

SMB (Small Minus Big) is the average return on the three small portfolios minus the average return on the three big portfolios.

SMB = 1/3 (Small Value + Small Neutral + Small Growth) - 1/3 (Big Value + Big Neutral + Big Growth).

HML (High Minus Low) is the average return on the two value portfolios minus the average return on the two growth portfolios.

HML = 1/2 (Small Value + Big Value) - 1/2 (Small Growth + Big Growth).

Rm-Rf, the excess return on the market, value-weight return of all CRSP firms incorporated in the USA and listed on the NYSE, AMEX, or NASDAQ that have a CRSP share code of 10 or 11 at the beginning of month t, good shares and price data at the beginning of t, and good return data for t minus the one-month Treasury bill rate (from Ibbotson Associates).

The daily US data sets run from 1926-07-01 to 2022-04-29 and feature a total of 25,230 observations. Descriptive statistics for the daily US factor series are shown in Table 1, whilst plots of the US daily series are provided in Figure 1. The US daily excess return for the market has a mean of 0.03, a standard deviation of 1.08, is negatively skewed and displays excess kurtosis 0f 16.79. The market factor SMB has a mean of of 0.004, has a standard deviation of 0.62, is positively skewed, and has excess kurtosis of 15.8. The factor HML has a mean of 0.015, a standard deviation of 0.62, is positively skewed and has excess kurtosis of 15.8. Finally the US riskfree rate has a mean of 0.01, a standard deviation of 0.01, is positively skewed and has excess kurtosis of 1.41.

The Japanese daily 3 factor sample runs from 02/07/1990 up to 29/04/2022 comprising 8305 observations and descriptive statistics of the Japanese daily series are provided in Table 2. The excess Japanese daily market return is only 0.006, but it must be borne in mind that the Japanese 'great recession' started in 1990, triggered by a collapse in stock and land prices. Its standard deviation was 1.35, with positive skewness and excess kurtosis of 5.37.

The Japanese SMB factor had a mean of -0.004, a standard deviation of 0.66, negative skewness and excess kurtosis of 7.95. The Japanese HML factor had a mean of 0.013, a standard deviation of 0.58, positive skewness, and excess kurtosis of 5.21. The Japanese daily RF had a mean of 0.009, a standard deviation of 0.008, positive skewness, and negative excess kurtosis. Plots of the Japanese daily series are provided in Figure 2. It is apparent in Figure 2 that there were relatively infrequent changes in daily interest rates in Japan within this period and they remained at 'low' levels.

A further daily series for developed markets, ex the USA, was taken from French's website. This data set includes the following countries: Austria, Belgium, Canada, Switzerland, Germany, Denmark, Spain, Finland, France, Great Britain, Greece, Hong Kong, Ireland, Italy, Japan, Netherlands, Norway , New Zealand, Portugal , Sweden, and Singapore. This series features 8261 observations drawn from 02/07/1990 to 28/02/2022. Descriptive statistics for this series are shown in Table 3.

The website provides a description of how the factors are constructed. It states that all returns are in U.S. dollars, include dividends and capital gains, and are not continuously compounded. The market is the return on a region's

Table 1: Descriptive statistics Basic US Daily Factor Series 1926-07-01--2022-04-29

Summary Statistics, using the observations 1926-07-01–2022-04-29
for the variable MktRF (25230 valid observations)

Mean	Median	Minimum	Maximum
0.030188	0.060000	−17.440	15.760
Std. Dev.	C.V.	Skewness	Ex. kurtosis
1.0780	35.711	−0.16535	16.792
5% perc.	95% perc.	IQ Range	Missing obs.
−1.5800	1.4900	0.90000	0

Summary Statistics, using the observations 1926-07-01–2022-04-29
for the variable SMB (25230 valid observations)

Mean	Median	Minimum	Maximum
0.0044744	0.010000	−11.670	8.1800
Std. Dev.	C.V.	Skewness	Ex. kurtosis
0.59132	132.15	−0.70261	21.394
5% perc.	95% perc.	IQ Range	Missing obs.
−0.83000	0.81000	0.52000	0

Summary Statistics, using the observations 1926-07-01–2022-04-29
for the variable HML (25230 valid observations)

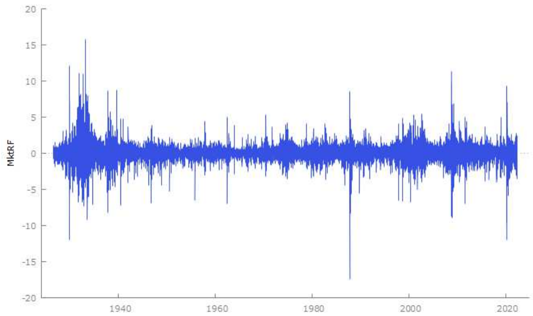
Mean	Median	Minimum	Maximum
0.015196	0.010000	−6.0200	9.0400
Std. Dev.	C.V.	Skewness	Ex. kurtosis
0.61989	40.792	0.71858	15.809
5% perc.	95% perc.	IQ Range	Missing obs.
−0.83000	0.88000	0.51000	0

Summary Statistics, using the observations 1926-07-01–2022-04-29
for the variable RF (25230 valid observations)

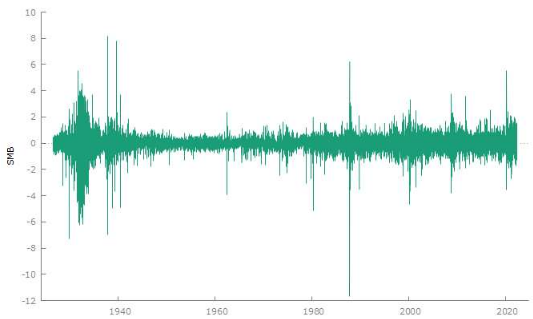
Mean	Median	Minimum	Maximum
0.012113	0.010000	−0.0030000	0.061000
Std. Dev.	C.V.	Skewness	Ex. kurtosis
0.011923	0.98428	1.1674	1.4173
5% perc.	95% perc.	IQ Range	Missing obs.
0.00000	0.034000	0.019000	0

Figure 1: Time Series Plots US Daily Series

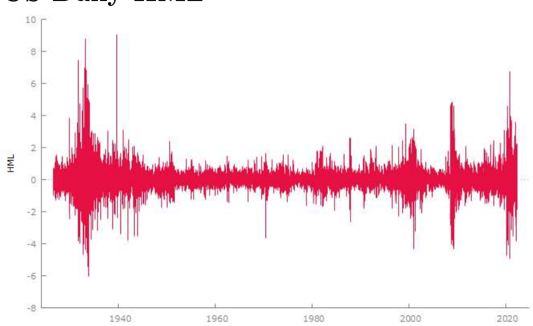
US Daily MktRF



US Daily SMB



US Daily HML



US Daily RF

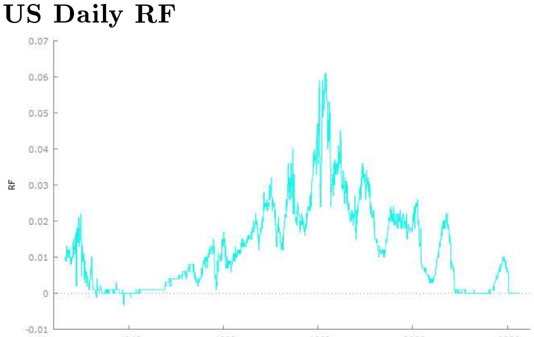


Table 2: Descriptive statistics Basic Japanese Daily Factor Series 1990-07-02--2022-04-29

Summary Statistics, using the observations 1990-07-02–2022-04-29
for the variable MktRF (8305 valid observations)

Mean	Median	Minimum	Maximum
0.0060205	0.010000	−10.850	13.020
Std. Dev.	C.V.	Skewness	Ex. kurtosis
1.3544	224.96	0.11221	5.3683
5% perc.	95% perc.	IQ Range	Missing obs.
−2.1270	2.0970	1.4350	0

Summary Statistics, using the observations 1990-07-02–2022-04-29
for the variable SMB (8305 valid observations)

Mean	Median	Minimum	Maximum
−0.0040638	0.00000	−9.1800	4.6000
Std. Dev.	C.V.	Skewness	Ex. kurtosis
0.66291	163.12	−0.63929	7.9480
5% perc.	95% perc.	IQ Range	Missing obs.
−1.0600	0.98000	0.70000	0

Summary Statistics, using the observations 1990-07-02–2022-04-29
for the variable HML (8305 valid observations)

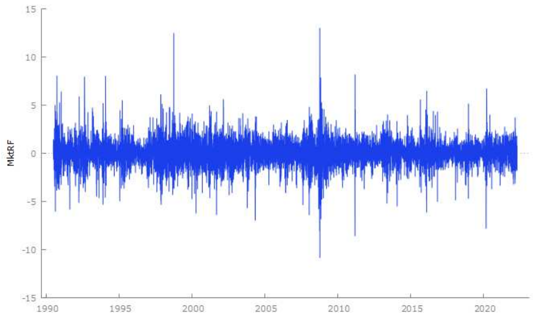
Mean	Median	Minimum	Maximum
0.013002	0.00000	−4.8800	4.5700
Std. Dev.	C.V.	Skewness	Ex. kurtosis
0.57601	44.303	0.21467	5.2102
5% perc.	95% perc.	IQ Range	Missing obs.
−0.89000	0.96000	0.52000	0

Summary Statistics, using the observations 1990-07-02–2022-04-29
for the variable RF (8305 valid observations)

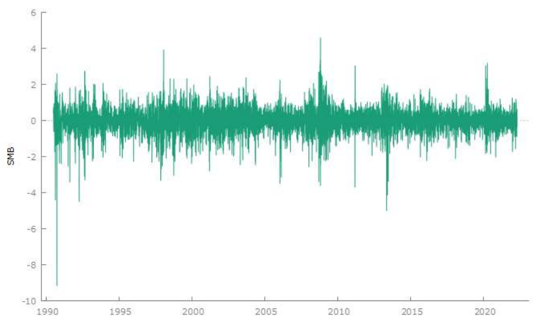
Mean	Median	Minimum	Maximum
0.0092414	0.010000	0.00000	0.030000
Std. Dev.	C.V.	Skewness	Ex. kurtosis
0.0087874	0.95087	0.31075	−1.2618
5% perc.	95% perc.	IQ Range	Missing obs.
0.00000	0.020000	0.020000	0

Figure 2: Time Series Plots Japanese Daily Series

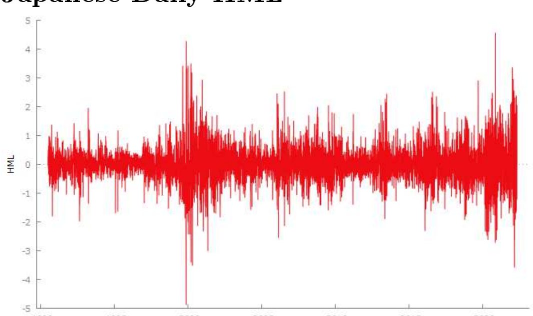
Japanese Daily MktRF



Japanese SMB



Japanese Daily HML



Japanese Daily RF

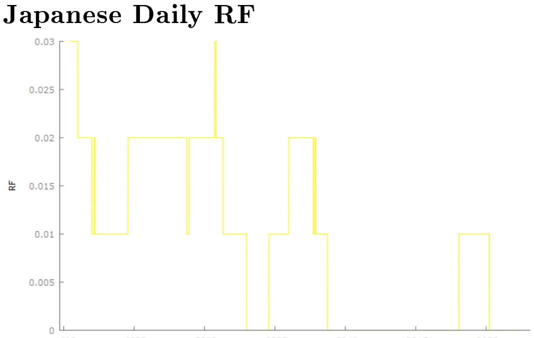


Table 3: Descriptive statistics Developed Markets Daily Factor Series 1990-07-02--2022-04-29

Summary Statistics, using the observations 1990-07-02–2022-02-28
for the variable MktRF (8261 valid observations)

Mean	Median	Minimum	Maximum
0.024009	0.050000	−9.6200	9.2000
Std. Dev.	**C.V.**	**Skewness**	**Ex. kurtosis**
0.90747	37.797	−0.43616	10.596
5% perc.	**95% perc.**	**IQ Range**	**Missing obs.**
−1.4000	1.2900	0.84000	0

Summary Statistics, using the observations 1990-07-02–2022-02-28
for the variable SMB (8261 valid observations)

Mean	Median	Minimum	Maximum
−0.0031001	0.010000	−5.3800	2.3700
Std. Dev.	**C.V.**	**Skewness**	**Ex. kurtosis**
0.42678	137.67	−0.87842	8.6114
5% perc.	**95% perc.**	**IQ Range**	**Missing obs.**
−0.66000	0.62000	0.45000	0

Summary Statistics, using the observations 1990-07-02–2022-02-28
for the variable HML (8261 valid observations)

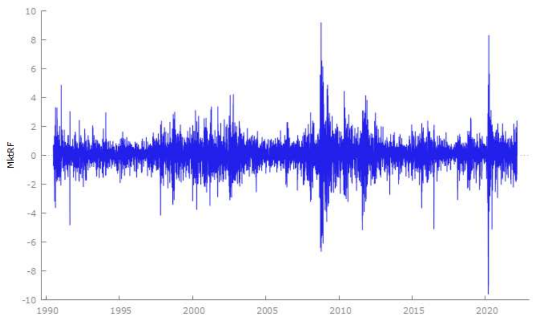
Mean	Median	Minimum	Maximum
0.010836	0.00000	−3.1000	4.1600
Std. Dev.	**C.V.**	**Skewness**	**Ex. kurtosis**
0.41856	38.625	0.44092	8.0936
5% perc.	**95% perc.**	**IQ Range**	**Missing obs.**
−0.60000	0.63000	0.36000	0

Summary Statistics, using the observations 1990-07-02–2022-02-28
for the variable RF (8261 valid observations)

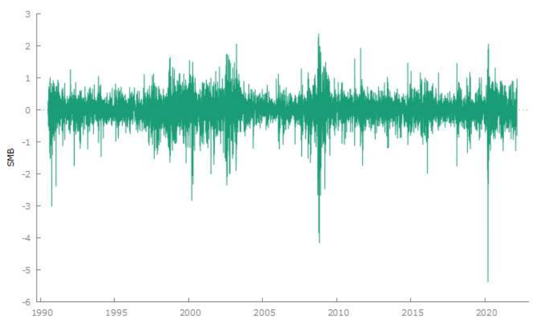
Mean	Median	Minimum	Maximum
0.0092906	0.010000	0.00000	0.030000
Std. Dev.	**C.V.**	**Skewness**	**Ex. kurtosis**
0.0087848	0.94555	0.30208	−1.2639
5% perc.	**95% perc.**	**IQ Range**	**Missing obs.**
0.00000	0.020000	0.020000	0

Figure 3: Time Series Plots Developed Markets Daily Series

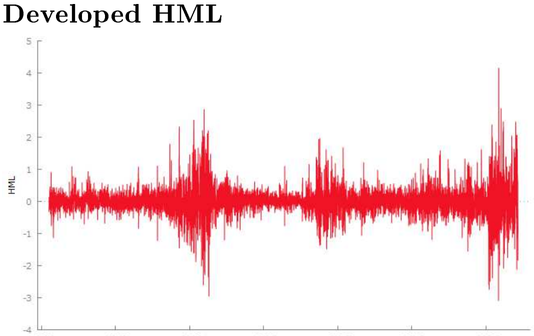
Developed MktRF



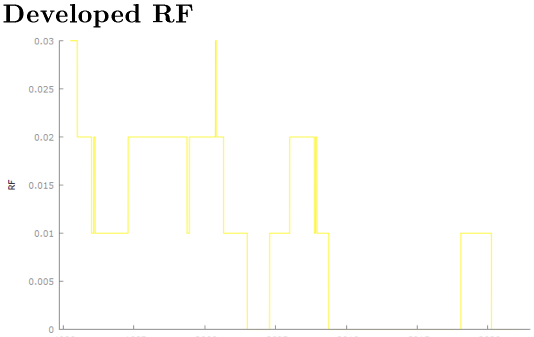
Developed SMB



Developed HML



Developed RF



value-weight market portfolio minus the U.S. one month T-bill rate. The SMB and HML factors, are constructed by sorting stocks in a region into two market cap and three book-to-market equity (B/M) groups at the end of each June. Big stocks are those in the top 90% of June market cap for the region, and small stocks are those in the bottom 10%. The B/M breakpoints for a region are the 30th and 70th percentiles of B/M for the big stocks of the region.

The mean return on the daily excess return in the developed markets, (MktRF) is 0.024, its standard deviation is 0.91, the skewness is -0.44 and the excess kurtosis is 10.60. The factor SMB has a mean value of -0.003, its standard deviation is 0.43, the skewness is -0.88, and the excess kurtosis is 8.61. Plots of the daily series for developed markets are shown in Figure 3.

3.2. Regression Analysis

I set up simple tests of the independence of the base factors using biviariate regressions to explore the relationships between them, and applied rolling regressions using a 250 day window, or roughly one year's worth of daily data, to explore the time varying relationships between them. For each of the three data sets, the USA, Japan and developed markets, SMB and HML were regressed on the excess market return $R_M - R_F$ (MktRF), plus SMB and HML were regressed on one another. A key assumption in asset pricing tests is that these market factors are independent. The results shown in Figures 4 to 6 reveal that, more often than not, this is not the case.

3.2.1. US Results

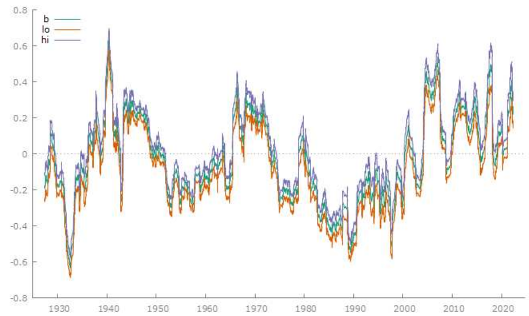
The first panel ot Figure 4 plots the results of regressing SMB and HML on the market factor $R_M - R_F$, in the US market using a 250 day window from July 1926 all the way up to the end of April 2022. The estimated slope coefficient is plotted in green, two standard deviations above and below the estimate, are plotted as 'lo' in brown and 'hi' in blue, representing the confidence limits of two standard errors.

The key factor in the interpretation of the graphs is whether the error bands overlap 0 in the middle of the diagrams. If They are above or below it, with no overlaps, then the slope coefficient is significantly positive or negative, as the case may be.

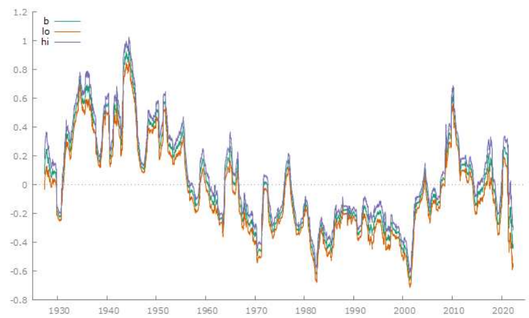
The results of these regressions, for SMB and $R_M - R_F$, shown in the first panel of Figure 4, reveal a significant negative relationship between SMB and $R_M - R_F$, from 1926 up to about 1928, when the relationship becomes briefly insignificant in the case of the US daily series, before returning to a significant negative relationship from 1929 to about 1935. After a period of indeterminacy, the relationship then switches to being significantly positive, for a period from around 1935 to 1948. From around 1950 to 1965 the relationship is significantly negative again, apart from a brief period of indeterminacy in the early 1960s. The relationship continues to oscillate, and is significantly negative for most of the 70's, 80's and 90's, before switching sign in the early 2000's and remaining significantly positive for the the bulk of the remaining time period. This does not support the assumption that this pair of factors is independent.

Figure 4: Rolling Regressions US US Daily Factor Series 1926-07-01--2022-04-29 (250 day window)

SMB on MktRF



HML on MktRF



HML on SMB

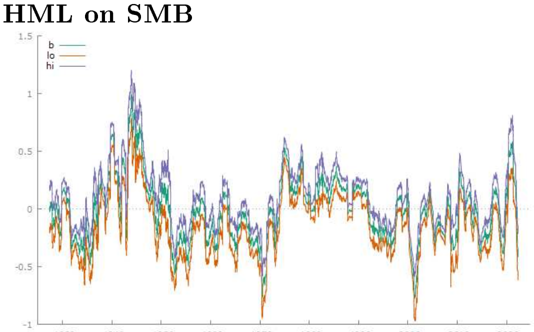


Table 4: OLS Regression Analysis USA SMB regressed on MktRF

OLS, using observations 1926-07-01–2022-04-29 ($T = 25230$)
Dependent variable: SMB

	Coefficient	Std. Error	t-ratio	p-value
const	0.00710378	0.00367700	1.932	0.0534
MktRF	−0.0870982	0.00340954	−25.55	0.0000

Mean dependent var	0.004474	S.D. dependent var	0.591316	
Sum squared resid	8598.997	S.E. of regression	0.583825	
R^2	0.025215	Adjusted R^2	0.025176	
$F(1, 25228)$	652.5702	P-value(F)	4.0e–142	RESET
Log-likelihood	−22221.18	Akaike criterion	44446.37	
Schwarz criterion	44462.64	Hannan–Quinn	44451.63	
$\hat{\rho}$	−0.013377	Durbin–Watson	2.026724	

test for specification –
Null hypothesis: specification is adequate
Test statistic: $F(2, 25226) = 107.129$
with p-value $= P(F(2, 25226) > 107.129) = 4.6874\text{e-}47$

OLS, using observations 1926-07-01–2022-04-29 ($T = 25230$)
Dependent variable: SMB

	Coefficient	Std. Error	t-ratio	p-value
const	0.0164715	0.00376539	4.374	0.0000
MktRF	−0.0880447	0.00340251	−25.88	0.0000
sq_MktRF	−0.00802988	0.000728390	−11.02	0.0000

Mean dependent var	0.004474	S.D. dependent var	0.591316
Sum squared resid	8557.770	S.E. of regression	0.582435
R^2	0.029888	Adjusted R^2	0.029811
$F(2, 25227)$	388.6099	P-value(F)	6.0e–167
Log-likelihood	−22160.56	Akaike criterion	44327.11
Schwarz criterion	44351.52	Hannan–Quinn	44335.01
$\hat{\rho}$	−0.021036	Durbin–Watson	2.042031

Table 5: OLS Regression Analysis USA HML regressed on MktRF

OLS, using observations 1926-07-01–2022-04-29 (T = 25230)
Dependent variable: HML

	Coefficient	Std. Error	t-ratio	p-value
const	0.0122841	0.00384887	3.192	0.0014
MktRF	0.0964636	0.00356890	27.03	0.0000

Mean dependent var	0.015196	S.D. dependent var	0.619886	
Sum squared resid	9421.639	S.E. of regression	0.611113	
R^2	0.028143	Adjusted R^2	0.028105	
$F(1, 25228)$	730.5624	P-value(F)	1.2e–158	RESET
Log-likelihood	−23373.73	Akaike criterion	46751.47	
Schwarz criterion	46767.74	Hannan–Quinn	46756.73	
$\hat{\rho}$	0.086478	Durbin–Watson	1.826962	

test for specification –
Null hypothesis: specification is adequate
Test statistic: $F(2, 25226) = 88.8704$
with p-value = $P(F(2, 25226) > 88.8704)$ = 3.46258e-39

OLS, using observations 1926-07-01–2022-04-29 (T = 25230)
Dependent variable: HML

	Coefficient	Std. Error	t-ratio	p-value
const	0.00331795	0.00394294	0.8415	0.4001
MktRF	0.0973695	0.00356295	27.33	0.0000
sq_MktRF	0.00768570	0.000762737	10.08	0.0000

Mean dependent var	0.015196	S.D. dependent var	0.619886
Sum squared resid	9383.871	S.E. of regression	0.609899
R^2	0.032039	Adjusted R^2	0.031963
$F(2, 25227)$	417.5045	P-value(F)	4.1e–179
Log-likelihood	−23323.06	Akaike criterion	46652.12
Schwarz criterion	46676.53	Hannan–Quinn	46660.02
$\hat{\rho}$	0.084612	Durbin–Watson	1.830706

Table 6: OLS Regression Analysis USA HML regressed on SMB

OLS, using observations 1926-07-01–2022-04-29 (T = 25230)
Dependent variable: HML

	Coefficient	Std. Error	t-ratio	p-value
const	0.0155291	0.00389294	3.989	0.0001
SMB	−0.0743954	0.00658347	−11.30	0.0000

Mean dependent var	0.015196	S.D. dependent var	0.619886
Sum squared resid	9645.651	S.E. of regression	0.618336
R^2	0.005036	Adjusted R^2	0.004997
$F(1, 25228)$	127.6975	P-value(F)	1.54e–29
Log-likelihood	−23670.16	Akaike criterion	47344.32
Schwarz criterion	47360.59	Hannan–Quinn	47349.59
$\hat{\rho}$	0.087324	Durbin–Watson	1.825308

RESET test for specification –
Null hypothesis: specification is adequate
Test statistic: $F(2, 25226) = 82.9036$
with p-value = $P(F(2, 25226) > 82.9036) = 1.29788\text{e-}36$

OLS, using observations 1926-07-01–2022-04-29 (T = 25230)
Dependent variable: HML

	Coefficient	Std. Error	t-ratio	p-value
const	0.00691708	0.00396923	1.743	0.0814
SMB	−0.0644351	0.00663656	−9.709	0.0000
sq_SMB	0.0245021	0.00232153	10.55	0.0000

Mean dependent var	0.015196	S.D. dependent var	0.619886
Sum squared resid	9603.247	S.E. of regression	0.616987
R^2	0.009410	Adjusted R^2	0.009332
$F(2, 25227)$	119.8249	P-value(F)	1.61e–52
Log-likelihood	−23614.58	Akaike criterion	47235.16
Schwarz criterion	47259.57	Hannan–Quinn	47243.06
$\hat{\rho}$	0.086148	Durbin–Watson	1.827663

The middle panel of Figure 4 depicts the relations between HML and $RM-RF$, in the case of the US daily series, and shows a similar varying pattern. From 1926 to about 1955, the relationship is significantly positive apart from a brief period of indeterminacy, and a brief period of a negative relationship in 1929-30. From around 2005 onwards, the relationship between the two series oscillates between significant positive and significant relationships, changing sign five times.

The last panel of Figure 4 shows the relationship between SMB and HML in the US daily series. This relationship is very variable and it oscillates between being significantly positive and significantly negative over most of the period, before becoming significantly negative or indeterminate from the 1990's onwards.

These results suggest the the US daily factors are not typically independent and furthermore, that they are likely to suffer from an endogeneity problem if they are used as explanatory variables in time series regressions.

I next explored the relations between the US factors across the entire time-period period using OLS regression. I examined whether the factors followed a linear relationship. The factors were regressed pairwise on each other, then RESET tests, Ramsey (1969), were applied to the regression equations to test whether a non-linear specification was merited.

The results of these regressions are presented in Tables 4, 5 and 6, and all support the existence of a significant non-linear relationship between the three factors in the sample period. The problems that can follow are examined by Andrews and Mikusheva (2016), who explore the issues related to the problems encountered by conventional tests for composite hypotheses based on linearity in minimum distance models and demonstrate that they can be unreliable when the relationship between the structural and reduced-form parameters is highly nonlinear.

3.2.2. Japanese Results

The rolling regression results using a 250 day window for the three Japanese daily factor series are presented in Figures 5. The results for SMB regressed on MktRF are particularly striking in Figure 5, where it can be seen that for the bulk of the entire period, there is a significant negative relationship between these two factors.

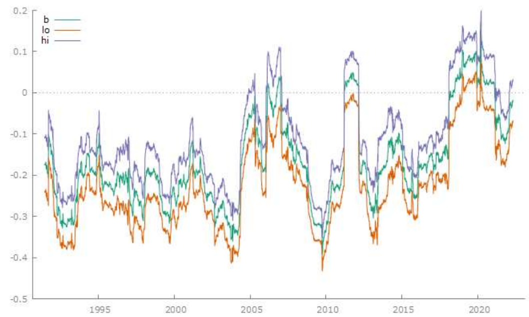
The results for HMP regressed on MktRF are also notable, but in this case, the relationship is significantly negative for most of the period from 1990 until around 2008, then it switches sign and becomes significantly positive for several small intervals, before switching back to being significantly negative for several years from around 2017.

The final diagram in Figure 5 shows the relationship between HML and SMB in Japan. This relationship is significantly positive for several periods up to 2003, from which time it switches sign and becomes significantly negative for the bulk of the remaining period.

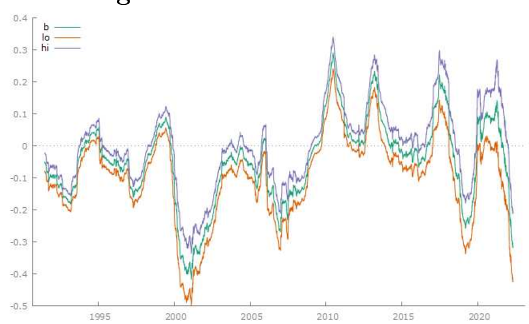
Tables 7, 8 and 9 shows the results of regressions across the whole of the Japanese sample period which parallel those just reported for the US market

Figure 5: Rolling Regressions Japan 250 Day Window

SMB on MktRF



HML Regressed on MktRF



HML Regressed on SMB

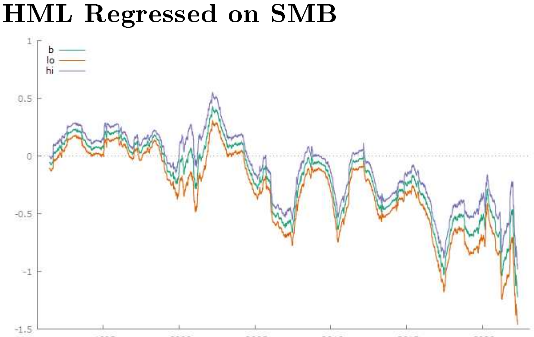


Table 7: Regression of SMB on MktRF Japanese daily sample.

OLS, using observations 1990-07-02–2022-04-29 ($T = 8305$)
Dependent variable: SMB

	Coefficient	Std. Error	t-ratio	p-value
const	−0.00292297	0.00670732	−0.4358	0.6630
MktRF	−0.189495	0.00495251	−38.26	0.0000

Mean dependent var	−0.004064	S.D. dependent var	0.662905	
Sum squared resid	3102.157	S.E. of regression	0.611244	
R^2	0.149893	Adjusted R^2	0.149791	
$F(1, 8303)$	1464.008	P-value(F)	3.7e–295	RESET
Log-likelihood	−7695.069	Akaike criterion	15394.14	
Schwarz criterion	15408.19	Hannan–Quinn	15398.94	
$\hat{\rho}$	0.153232	Durbin–Watson	1.693437	

test for specification –
Null hypothesis: specification is adequate
Test statistic: $F(2, 8301) = 62.2617$
with p-value = $P(F(2, 8301) > 62.2617) = 1.44842\text{e-}27$

OLS, using observations 1990-07-02–2022-04-29 ($T = 8305$)
Dependent variable: SMB

	Coefficient	Std. Error	t-ratio	p-value
const	0.0225161	0.00710295	3.170	0.0015
MktRF	−0.187219	0.00492602	−38.01	0.0000
sq_MktRF	−0.0138768	0.00133978	−10.36	0.0000

Mean dependent var	−0.004064	S.D. dependent var	0.662905
Sum squared resid	3062.583	S.E. of regression	0.607369
R^2	0.160738	Adjusted R^2	0.160536
$F(2, 8302)$	795.0131	P-value(F)	0.000000
Log-likelihood	−7641.754	Akaike criterion	15289.51
Schwarz criterion	15310.58	Hannan–Quinn	15296.71
$\hat{\rho}$	0.151223	Durbin–Watson	1.697460

Table 8: Regression of SMB on MktRF Japanese daily sample.

OLS, using observations 1990-07-02–2022-04-29 ($T = 8305$)
Dependent variable: HML

	Coefficient	Std. Error	t-ratio	p-value
const	0.0133648	0.00625727	2.136	0.0327
MktRF	−0.0602943	0.00462021	−13.05	0.0000

Mean dependent var	0.013002	S.D. dependent var	0.576014	
Sum squared resid	2699.822	S.E. of regression	0.570230	
R^2	0.020099	Adjusted R^2	0.019981	
$F(1, 8303)$	170.3061	P-value(F)	1.52e–38	RESET
Log-likelihood	−7118.238	Akaike criterion	14240.48	
Schwarz criterion	14254.52	Hannan–Quinn	14245.27	
$\hat{\rho}$	0.130228	Durbin–Watson	1.739543	

test for specification –
Null hypothesis: specification is adequate
Test statistic: $F(2, 8301) = 7.07425$
with p-value $= P(F(2, 8301) > 7.07425) = 0.000851738$

OLS, using observations 1990-07-02–2022-04-29 ($T = 8305$)
Dependent variable: HML

	Coefficient	Std. Error	t-ratio	p-value
const	0.00586076	0.00666478	0.8794	0.3792
MktRF	−0.0609657	0.00462214	−13.19	0.0000
sq_MktRF	0.00409339	0.00125713	3.256	0.0011

Mean dependent var	0.013002	S.D. dependent var	0.576014
Sum squared resid	2696.379	S.E. of regression	0.569901
R^2	0.021349	Adjusted R^2	0.021113
$F(2, 8302)$	90.55275	P-value(F)	1.25e–39
Log-likelihood	−7112.938	Akaike criterion	14231.88
Schwarz criterion	14252.95	Hannan–Quinn	14239.08
$\hat{\rho}$	0.130306	Durbin–Watson	1.739387

Table 9: Regression of HML on SMB Japanese daily sample.

OLS, using observations 1990-07-02–2022-04-29 ($T = 8305$)
Dependent variable: HML

	Coefficient	Std. Error	t-ratio	p-value
const	0.0124149	0.00623323	1.992	0.0464
SMB	−0.144431	0.00940329	−15.36	0.0000

Mean dependent var	0.013002	S.D. dependent var	0.576014
Sum squared resid	2679.078	S.E. of regression	0.568035
R^2	0.027628	Adjusted R^2	0.027511
$F(1, 8303)$	235.9164	P-value(F)	1.61e–52
Log-likelihood	−7086.208	Akaike criterion	14176.42
Schwarz criterion	14190.46	Hannan–Quinn	14181.22
$\hat{\rho}$	0.136472	Durbin–Watson	1.727056

RESET test for specification –
Null hypothesis: specification is adequate
Test statistic: $F(2, 8301) = 1.41518$
with p-value = $P(F(2, 8301) > 1.41518) = 0.24294$

factors, and test whether the relationship between factors is linear. The results in Table 7 show that the relationship between SMB and MktRF is non-linear. Similarly, Table 8 reports the relationship between HML and MktRF is also non-linear. However, Table 9 demonstrates that there is no evidence of non-linearity in the relationship between HML and SMB in the Japanese daily series.

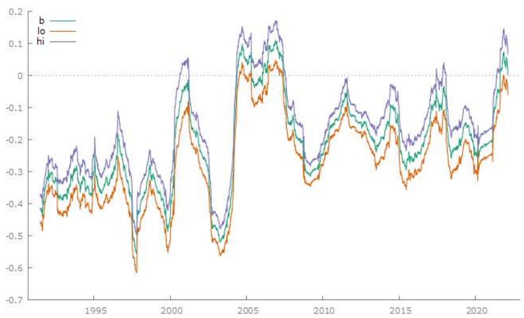
3.2.3. Regression Analysis Developed Markets

Figure 6 displays the results of the rolling regressions analysis in developed markets between the daily factor return series using a 250 day window. The first diagram in Figure 6 shows that for almost the entire period, there is a significant negative relationship between SMB and MktRF in these developed markets. The relationship between HML and MktRF is more variable. It is predominantly negative until around 2003, then its has several small periods featuring a positive relationship, it becomes significantly negative again around 2017, and then switches back to being significantly positive around 2020. The relationship between HML and SMB in developed markets is less consistent. It starts the sample period significantly positive, becomes insigificant around 1993, switches back to being signicantly positive around 1997, remains so apart from a brief switch up to 2006, and then becomes significantly negative all the way up to 2016, and then becomes insignificant.

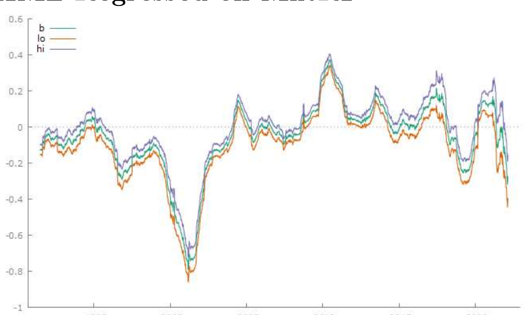
Tables 10, 11 and 12 report regression results that examine the linearity of the relationship between the three factors in the daily sample for developed markets. The results for SMB and HML, when regressed on MktRF, support the hypothesis of a non-linear relationship. However, there is no evidence of a non-linear relationship between HML and SMB.

Figure 6: Rolling Regressions Developed Markets 250 day window

SMB regressed on MktRF



HML Regressed on MktRF



HML Regressed on SMB

Table 10: Regression of SMB on MktRF Developed Markets daily sample.

OLS, using observations 1990-07-02–2022-02-28 ($T = 8261$)
Dependent variable: SMB

	Coefficient	Std. Error	t-ratio	p-value
const	0.00252491	0.00407316	0.6199	0.5353
MktRF	−0.234286	0.00448720	−52.21	0.0000

Mean dependent var	−0.003100	S.D. dependent var	0.426784
Sum squared resid	1131.149	S.E. of regression	0.370081
R^2	0.248164	Adjusted R^2	0.248073
$F(1, 8259)$	2726.106	P-value(F)	0.000000
Log-likelihood	−3509.130	Akaike criterion	7022.260
Schwarz criterion	7036.299	Hannan–Quinn	7027.058
$\hat{\rho}$	0.102398	Durbin–Watson	1.794546

RESET test for specification –
Null hypothesis: specification is adequate
Test statistic: $F(2, 8257) = 82.2727$
with p-value = $P(F(2, 8257) > 82.2727) = 4.17624e\text{-}36$

OLS, using observations 1990-07-02–2022-02-28 ($T = 8261$)
Dependent variable: SMB

	Coefficient	Std. Error	t-ratio	p-value
const	0.0159800	0.00420357	3.802	0.0001
MktRF	−0.239908	0.00447757	−53.58	0.0000
sq_MktRF	−0.0161657	0.00139271	−11.61	0.0000

Mean dependent var	−0.003100	S.D. dependent var	0.426784
Sum squared resid	1112.991	S.E. of regression	0.367120
R^2	0.260233	Adjusted R^2	0.260054
$F(2, 8258)$	1452.489	P-value(F)	0.000000
Log-likelihood	−3442.284	Akaike criterion	6890.568
Schwarz criterion	6911.626	Hannan–Quinn	6897.765
$\hat{\rho}$	0.084062	Durbin–Watson	1.831197

Table 11: Regression of HML on MktRF Developed Markets daily sample.

OLS, using observations 1990-07-02–2022-02-28 ($T = 8261$)
Dependent variable: HML

	Coefficient	Std. Error	t-ratio	p-value
const	0.0116734	0.00459382	2.541	0.0111
MktRF	−0.0348597	0.00506077	−6.888	0.0000

Mean dependent var	0.010836	S.D. dependent var	0.418558
Sum squared resid	1438.811	S.E. of regression	0.417386
R^2	0.005712	Adjusted R^2	0.005592
$F(1, 8259)$	47.44751	P-value(F)	6.06e–12
Log-likelihood	−4502.858	Akaike criterion	9009.716
Schwarz criterion	9023.754	Hannan–Quinn	9014.513
$\hat{\rho}$	0.216750	Durbin–Watson	1.566026

RESET test for specification –
Null hypothesis: specification is adequate
Test statistic: $F(2, 8257) = 32.7717$
with p-value = $P(F(2, 8257) > 32.7717) = 6.66196\text{e-}15$

OLS, using observations 1990-07-02–2022-02-28 ($T = 8261$)
Dependent variable: HML

	Coefficient	Std. Error	t-ratio	p-value
const	0.0158364	0.00477653	3.315	0.0009
MktRF	−0.0365991	0.00508787	−7.193	0.0000
sq_MktRF	−0.00500166	0.00158254	−3.161	0.0016

Mean dependent var	0.010836	S.D. dependent var	0.418558
Sum squared resid	1437.073	S.E. of regression	0.417159
R^2	0.006913	Adjusted R^2	0.006673
$F(2, 8258)$	28.74402	P-value(F)	3.63e–13
Log-likelihood	−4497.865	Akaike criterion	9001.729
Schwarz criterion	9022.787	Hannan–Quinn	9008.925
$\hat{\rho}$	0.215910	Durbin–Watson	1.567701

Table 12: Regression of HML on SMB Developed Markets daily sample.

OLS, using observations 1990-07-02–2022-02-28 ($T = 8261$)
Dependent variable: HML

	Coefficient	Std. Error	t-ratio	p-value
const	0.0108193	0.00460543	2.349	0.0188
SMB	−0.00552954	0.0107914	−0.5124	0.6084

Mean dependent var	0.010836	S.D. dependent var	0.418558
Sum squared resid	1447.031	S.E. of regression	0.418577
R^2	0.000032	Adjusted R^2	-0.000089
$F(1, 8259)$	0.262557	P-value(F)	0.608382
Log-likelihood	−4526.388	Akaike criterion	9056.776
Schwarz criterion	9070.815	Hannan–Quinn	9061.574
$\hat{\rho}$	0.218349	Durbin–Watson	1.562827

RESET test for specification –

Null hypothesis: specification is adequate
Test statistic: $F(2, 8257) = 1.13528$
with p-value $= P(F(2, 8257) > 1.13528) = 0.321383$

3.3. Time Series Tests

Engle, et al (1983) drew attention to the issues related to the fact that precise definitions of "exogeneity" are elusive and their direct implications for concept for inference, given that a certain variable is "exogenous" or not, on any given definition. They proposed definitions for weak and strong exogeneity in terms of the distributions of observable variables. They suggested that essentially, a variable z_t , in a model is defined to be weakly exogenous for estimating a set of parameters λ , if inference on λ conditional on z_t , involves no loss of information. Heuristically, given that the joint density of random variables (y_t, z_t) always can be written as the product of y_t conditional on z_t, times the marginal of z_t, the weak exogeneity of z_t, entails that the precise specification of the latter density is irrelevant to the analysis, and, in particular that all parameters which appear in this marginal density are nuisance parameters.

If in addition to being weakly exogenous, z_t , is not Granger-caused in the sense of Granger (1969), by any of the endogenous variables in the system, then z, is defined to be strongly exogenous.

The results in this subsection, in Tables 13 to 15 report Granger causality tests of the three factors for the US daily series, when a lagged value of a second factor is added to the auto-regression on lags of the factor in question; thus constituting a test of strong exogenity. The results show that all three factors reject the hypothesis of strict exogeneity, as their Adjusted R squares increase when one lag of a further factor is included, but the effect is most pronounced in the SMB case, when a lagged value of the MktRF factor is added to the estimation.

Tables 16 to 18 show the results of a similar analysis for Japan. All the cases of the factors taken pairwise show evidence of Granger causality and reject the null hypothesis of strong exogeneity. Finally, Tables 19 to 21 report the results of strong exogeneity tests for the developed markets, and in all 3 pairs of cases, reject the null hypothesis of strong exogeneity.

4. Conclusion

The results in the paper suggest that there is a significant potential issue of endogeneity between the three base Fama-French factors, when asset pricing tests are conducted in a time series context, as set out in section 1 of the paper in equation (9). These daily sample series for the USA, Japan, and developed markets, show significant relationships between the factors, when they are explored using rolling regression analysis and a 250 day window. For the bulk of the time, there is a statistically significant relationship between the factors. This is further compounded by the fact that the sign of this relationship frequently switches.

Tests of the linearity of the relationship between SMB, HML, and the excess return on the market MktRF, uniformly reject the null of a linear relationship. This is not the case in the relationship between SMB and HML.

Further analyses in a time-series context and tests of Granger causality between the factors reject the Engle et al. (1983) concept of 'strong exogeneity'.

The results suggest that caution should be oberved in the application of time series asset pricing tests to avoid potential pitfalls outlined by Anatolyev and Mikusheva (2022), Andrews and Mikusheva (2016), Mikusheva and Sun (2022), and Onatski (2015). Allen and McAleer (2018) point out that the issue of potential endogeneity undermines the approach suggesting for choosing factors by Fama and French (2018). Perhaps they support the alternative interpetation of factor models, suggested by Fama and French (2020), as set out in equation (7) in section 1 of this paper?

Table 13: Granger Causality tests US Daily Data

SMB and MktRF

OLS, using observations 1926-07-02–2022-04-29 (T = 25229)
Dependent variable: SMB

	Coefficient	Std. Error	t-ratio	p-value
const	0.00470680	0.00371835	1.266	0.2056
SMB_1	−0.0501871	0.00628828	−7.981	0.0000

Mean dependent var	0.004484	S.D. dependent var	0.591326
Sum squared resid	8799.154	S.E. of regression	0.590592
R^2	0.002519	Adjusted R^2	0.002479
$F(1, 25227)$	63.69711	P-value(F)	1.51e–15
Log-likelihood	−22511.06	Akaike criterion	45026.12
Schwarz criterion	45042.39	Hannan–Quinn	45031.39
$\hat{\rho}$	0.000343	Durbin's h	1.117031

OLS, using observations 1926-07-02–2022-04-29 (T = 25229)
Dependent variable: SMB

	Coefficient	Std. Error	t-ratio	p-value
const	8.90393e–005	0.00358890	0.02481	0.9802
MktRF_1	0.146063	0.00337094	43.33	0.0000
SMB_1	−0.00794355	0.00614453	−1.293	0.1961

Mean dependent var	0.004484	S.D. dependent var	0.591326
Sum squared resid	8189.627	S.E. of regression	0.569781
R^2	0.071615	Adjusted R^2	0.071542
$F(2, 25226)$	972.9617	P-value(F)	0.000000
Log-likelihood	−21605.50	Akaike criterion	43217.00
Schwarz criterion	43241.41	Hannan–Quinn	43224.90
$\hat{\rho}$	−0.009459	Durbin's h	−6.895978

Table 14: Granger Causality tests US Daily Data

HML and MktRF

OLS, using observations 1926-07-02–2022-04-29 (T = 25229)
Dependent variable: HML

	Coefficient	Std. Error	t-ratio	p-value
const	0.0139421	0.00389037	3.584	0.0003
HML_1	0.0834228	0.00627417	13.30	0.0000

Mean dependent var	0.015208	S.D. dependent var	0.619896
Sum squared resid	9626.923	S.E. of regression	0.617747
R^2	0.006959	Adjusted R^2	0.006920
$F(1, 25227)$	176.7901	P-value(F)	3.32e–40
Log-likelihood	−23645.21	Akaike criterion	47294.41
Schwarz criterion	47310.68	Hannan–Quinn	47299.68
$\hat{\rho}$	−0.001393	Durbin's h	−2.671962

OLS, using observations 1926-07-02–2022-04-29 (T = 25229)
Dependent variable: HML

	Coefficient	Std. Error	t-ratio	p-value
const	0.0137208	0.00389118	3.526	0.0004
MktRF_1	0.00854229	0.00366008	2.334	0.0196
HML_1	0.0809287	0.00636397	12.72	0.0000

Mean dependent var	0.015208	S.D. dependent var	0.619896
Sum squared resid	9624.844	S.E. of regression	0.617693
R^2	0.007174	Adjusted R^2	0.007095
$F(2, 25226)$	91.13418	P-value(F)	3.66e–40
Log-likelihood	−23642.48	Akaike criterion	47290.96
Schwarz criterion	47315.37	Hannan–Quinn	47298.86
$\hat{\rho}$	−0.001632	Durbin's h	NA

Table 15: Granger Causality tests US Daily Data

HML and SMB

OLS, using observations 1926-07-02–2022-04-29 (T = 25229)
Dependent variable: HML

	Coefficient	Std. Error	t-ratio	p-value
const	0.0139421	0.00389037	3.584	0.0003
HML_1	0.0834228	0.00627417	13.30	0.0000

Mean dependent var	0.015208	S.D. dependent var	0.619896
Sum squared resid	9626.923	S.E. of regression	0.617747
R^2	0.006959	Adjusted R^2	0.006920
$F(1, 25227)$	176.7901	P-value(F)	3.32e–40
Log-likelihood	−23645.21	Akaike criterion	47294.41
Schwarz criterion	47310.68	Hannan–Quinn	47299.68
$\hat{\rho}$	−0.001393	Durbin's h	−2.671962

OLS, using observations 1926-07-02–2022-04-29 (T = 25229)
Dependent variable: HML

	Coefficient	Std. Error	t-ratio	p-value
const	0.0139930	0.00389046	3.597	0.0003
SMB_1	−0.00930734	0.00659393	−1.412	0.1581
HML_1	0.0827923	0.00628992	13.16	0.0000

Mean dependent var	0.015208	S.D. dependent var	0.619896
Sum squared resid	9626.162	S.E. of regression	0.617735
R^2	0.007038	Adjusted R^2	0.006959
$F(2, 25226)$	89.39467	P-value(F)	2.06e–39
Log-likelihood	−23644.21	Akaike criterion	47294.42
Schwarz criterion	47318.83	Hannan–Quinn	47302.32
$\hat{\rho}$	−0.001395	Durbin's h	−5.133830

Table 16: Granger Causality tests Japanese Daily Data

SMB and MktRF

OLS, using observations 1990-07-03–2022-04-29 ($T = 8304$)
Dependent variable: SMB

	Coefficient	Std. Error	t-ratio	p-value
const	−0.00389581	0.00726533	−0.5362	0.5918
SMB_1	0.0527169	0.0109596	4.810	0.0000

Mean dependent var	−0.004110	S.D. dependent var	0.662932
Sum squared resid	3638.850	S.E. of regression	0.662050
R^2	0.002779	Adjusted R^2	0.002659
$F(1, 8302)$	23.13709	P-value(F)	1.54e–06
Log-likelihood	−8357.176	Akaike criterion	16718.35
Schwarz criterion	16732.40	Hannan–Quinn	16723.15
$\hat{\rho}$	0.001714	Durbin's h	3.074560

OLS, using observations 1990-07-03–2022-04-29 ($T = 8304$)
Dependent variable: SMB

	Coefficient	Std. Error	t-ratio	p-value
const	−0.00431043	0.00695825	−0.6195	0.5356
MktRF_1	0.152595	0.00557209	27.39	0.0000
SMB_1	0.173422	0.0113842	15.23	0.0000

Mean dependent var	−0.004110	S.D. dependent var	0.662932
Sum squared resid	3337.333	S.E. of regression	0.634066
R^2	0.085409	Adjusted R^2	0.085189
$F(2, 8301)$	387.5957	P-value(F)	1.2e–161
Log-likelihood	−7998.046	Akaike criterion	16002.09
Schwarz criterion	16023.17	Hannan–Quinn	16009.29
$\hat{\rho}$	−0.025249	Durbin's h	NA

Table 17: Granger Causality tests Japanese Daily Data

HML and MktRF

OLS, using observations 1990-07-03–2022-04-29 ($T = 8304$)
Dependent variable: HML

	Coefficient	Std. Error	t-ratio	p-value
const	0.0111871	0.00626102	1.787	0.0740
HML_1	0.140140	0.0108668	12.90	0.0000

Mean dependent var	0.013009	S.D. dependent var	0.576048
Sum squared resid	2701.086	S.E. of regression	0.570398
R^2	0.019639	Adjusted R^2	0.019521
$F(1, 8302)$	166.3110	P-value(F)	1.09e–37
Log-likelihood	−7119.823	Akaike criterion	14243.65
Schwarz criterion	14257.69	Hannan–Quinn	14248.45
$\hat{\rho}$	0.001886	Durbin's h	1.233569

OLS, using observations 1990-07-03–2022-04-29 ($T = 8304$)
Dependent variable: HML

	Coefficient	Std. Error	t-ratio	p-value
const	0.0114026	0.00625395	1.823	0.0683
MktRF_1	−0.0209856	0.00466340	−4.500	0.0000
HML_1	0.133144	0.0109650	12.14	0.0000

Mean dependent var	0.013009	S.D. dependent var	0.576048
Sum squared resid	2694.512	S.E. of regression	0.569738
R^2	0.022025	Adjusted R^2	0.021789
$F(2, 8301)$	93.47362	P-value(F)	7.17e–41
Log-likelihood	−7109.706	Akaike criterion	14225.41
Schwarz criterion	14246.49	Hannan–Quinn	14232.61
$\hat{\rho}$	0.002701	Durbin's h	6.142548

Table 18: Granger Causality tests Japanese Daily Data

HML and SMB

OLS, using observations 1990-07-03–2022-04-29 ($T = 8304$)
Dependent variable: HML

	Coefficient	Std. Error	t-ratio	p-value
const	0.0111871	0.00626102	1.787	0.0740
HML_1	0.140140	0.0108668	12.90	0.0000

Mean dependent var	0.013009	S.D. dependent var	0.576048
Sum squared resid	2701.086	S.E. of regression	0.570398
R^2	0.019639	Adjusted R^2	0.019521
$F(1, 8302)$	166.3110	P-value(F)	1.09e–37
Log-likelihood	−7119.823	Akaike criterion	14243.65
Schwarz criterion	14257.69	Hannan–Quinn	14248.45
$\hat{\rho}$	0.001886	Durbin's h	1.233569

OLS, using observations 1990-07-03–2022-04-29 ($T = 8304$)
Dependent variable: HML

	Coefficient	Std. Error	t-ratio	p-value
const	0.0114026	0.00625395	1.823	0.0683
MktRF_1	−0.0209856	0.00466340	−4.500	0.0000
HML_1	0.133144	0.0109650	12.14	0.0000

Mean dependent var	0.013009	S.D. dependent var	0.576048
Sum squared resid	2694.512	S.E. of regression	0.569738
R^2	0.022025	Adjusted R^2	0.021789
$F(2, 8301)$	93.47362	P-value(F)	7.17e–41
Log-likelihood	−7109.706	Akaike criterion	14225.41
Schwarz criterion	14246.49	Hannan–Quinn	14232.61
$\hat{\rho}$	0.002701	Durbin's h	6.142548

Table 19: Granger Causality tests Developed Markets Daily Data SMB and MktRF

OLS, using observations 1990-07-03–2022-04-29 ($T = 8304$)
Dependent variable: SMB

	Coefficient	Std. Error	t-ratio	p-value
const	−0.00117414	0.00575189	−0.2041	0.8383
SMB_1	−0.0261558	0.0109714	−2.384	0.0171

Mean dependent var	−0.001144	S.D. dependent var	0.524295
Sum squared resid	2280.808	S.E. of regression	0.524147
R^2	0.000684	Adjusted R^2	0.000564
$F(1, 8302)$	5.683438	P-value(F)	0.017148
Log-likelihood	−6417.620	Akaike criterion	12839.24
Schwarz criterion	12853.29	Hannan–Quinn	12844.04
$\hat{\rho}$	−0.002194	Durbin's h	−9.582817

OLS, using observations 1990-07-03–2022-04-29 ($T = 8304$)
Dependent variable: SMB

	Coefficient	Std. Error	t-ratio	p-value
const	−0.00302319	0.00564993	−0.5351	0.5926
MktRF_1	0.129226	0.00738302	17.50	0.0000
SMB_1	0.120656	0.0136549	8.836	0.0000

Mean dependent var	−0.001144	S.D. dependent var	0.524295
Sum squared resid	2199.628	S.E. of regression	0.514765
R^2	0.036253	Adjusted R^2	0.036020
$F(2, 8301)$	156.1260	P-value(F)	2.75e–67
Log-likelihood	−6267.145	Akaike criterion	12540.29
Schwarz criterion	12561.36	Hannan–Quinn	12547.49
$\hat{\rho}$	−0.012961	Durbin's h	NA

Table 20: Granger Causality tests Developed Markets Daily Data HML and MktRF

OLS, using observations 1990-07-03–2022-04-29 ($T = 8304$)
Dependent variable: HML

	Coefficient	Std. Error	t-ratio	p-value
const	0.0114264	0.00390141	2.929	0.0034
HML_1	0.242362	0.0106473	22.76	0.0000

Mean dependent var	0.015075	S.D. dependent var	0.366116
Sum squared resid	1047.560	S.E. of regression	0.355221
R^2	0.058745	Adjusted R^2	0.058632
$F(1, 8302)$	518.1430	P-value(F)	2.6e–111
Log-likelihood	−3187.089	Akaike criterion	6378.177
Schwarz criterion	6392.226	Hannan–Quinn	6382.977
$\hat{\rho}$	0.006376	Durbin's h	2.400031

OLS, using observations 1990-07-03–2022-04-29 ($T = 8304$)
Dependent variable: HML

	Coefficient	Std. Error	t-ratio	p-value
const	0.0115709	0.00390129	2.966	0.0030
MktRF_1	−0.00824968	0.00402405	−2.050	0.0404
HML_1	0.241321	0.0106574	22.64	0.0000

Mean dependent var	0.015075	S.D. dependent var	0.366116
Sum squared resid	1047.030	S.E. of regression	0.355152
R^2	0.059222	Adjusted R^2	0.058995
$F(2, 8301)$	261.2729	P-value(F)	9.1e–111
Log-likelihood	−3184.987	Akaike criterion	6375.974
Schwarz criterion	6397.047	Hannan–Quinn	6383.173
$\hat{\rho}$	0.006825	Durbin's h	2.608833

Table 21: Granger Causality tests Developed Markets Daily HML and SMB

OLS, using observations 1990-07-03–2022-04-29 ($T = 8304$)
Dependent variable: HML

	Coefficient	Std. Error	t-ratio	p-value
const	0.0114264	0.00390141	2.929	0.0034
HML_1	0.242362	0.0106473	22.76	0.0000

Mean dependent var	0.015075	S.D. dependent var	0.366116
Sum squared resid	1047.560	S.E. of regression	0.355221
R^2	0.058745	Adjusted R^2	0.058632
$F(1, 8302)$	518.1430	P-value(F)	2.6e–111
Log-likelihood	−3187.089	Akaike criterion	6378.177
Schwarz criterion	6392.226	Hannan–Quinn	6382.977
$\hat{\rho}$	0.006376	Durbin's h	2.400031

OLS, using observations 1990-07-03–2022-04-29 ($T = 8304$)
Dependent variable: HML

	Coefficient	Std. Error	t-ratio	p-value
const	0.0114107	0.00389965	2.926	0.0034
SMB_1	0.0217697	0.00746027	2.918	0.0035
HML_1	0.245070	0.0106829	22.94	0.0000

Mean dependent var	0.015075	S.D. dependent var	0.366116
Sum squared resid	1046.486	S.E. of regression	0.355060
R^2	0.059710	Adjusted R^2	0.059483
$F(2, 8301)$	263.5637	P-value(F)	1.1e–111
Log-likelihood	−3182.832	Akaike criterion	6371.663
Schwarz criterion	6392.737	Hannan–Quinn	6378.863
$\hat{\rho}$	0.005872	Durbin's h	2.339282

References

[1] Anatolyev, S., and A. Mikusheva, (2022) Factor models with many assets: strong factors, weak factors, and the two-pass procedure, *Journal of Econometrics*, 229(1), 103-126.

[2] Andrews, I., and A. Mikusheva, (2016) A geometric approach to non-linear econometric models, Econometrica, 84(3): 1249-1264.

[3] Allen, D. E. and M. McAleer (2018) 'Choosing Factors' by Fama and French (2018): A Comment,
Available at SSRN: https://ssrn.com/abstract=3272608 or http://dx.doi.org/10.2139/ssrn.3272608.

[4] Barillas, F., and J. Shanken, (2018) Comparing asset pricing models, *Journal of Finance*, 73(2), 715-754.

[5] Cochrane, J.H. (2011) Presidential Address: Discount rates, *Journal of Finance*, 66(4), 1047-1108.

[6] Durbin, J. (1954) Errors in variables, *Review of the International Statistical Institute,* 22(1/3), 23–32.

[7] Engle, R.F., D.F. Hendry, and J-F. Richard, (1983) Exogeneity, *Econometrica*, 51(2): 277-304.

[8] Fama, E.F., and K.R. French (1993) Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics,* 33, 3-56.

[9] Fama, E.F., and K.R. French (2018) Choosing factors, *Journal of Financial Economics*, 128(2), 234-252.

[10] Fama, E. F. (1976) *Foundations of finance*, Basic Books: New York.

[11] Fama, E.F., and K.R French (2020) Comparing Cross-Section and Time-Series Factor Models, *The Review of Financial Studies,* 33: 1891–1926.

[12] Fama, E.F., and K.R French (1993) Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics,* 33:3–56.

[13] Fama, E.F., and K.R French (2015) A five-factor asset pricing model, *Journal of Financial Economics,* 116:1–22.

[14] Fama, E. F., and J. D. MacBeth (1973) Risk return, and equilibrium: empirical tests, *Journal of Political Economy,* 81:607–36.

[15] Gibbons, M.R., S.A. Ross, and J. Shanken (1989) A test of the efficiency of a given portfolio, *Econometrica*, 57(5), 1121-1152.

[16] Granger, C.W. J. (1969) Investigating causal relations by econometric models and Cross-Spectral methods, *Econometrica*, 37: 424-438.

[17] Harvey, C.R., Y. Liu, and H. Zhu, (2015) ...and the cross-section of expected returns, *Review of Financial Studies,* 29, 5-68.

[18] Hausman, J.A. (1978) Specification tests in econometrics, *Econometrica*, 46(6), 1251–1271.

[19] Kan, R., and C. Zhang, (1999) Two-Pass Tests of Asset Pricing Models with Useless Factors, *Journal of Finance*, 54(1): 203-235.

[20] Kleibergen, F. (2009) Tests of Risk Premia in Linear Factor Models, *Journal of Econometrics*, 149(2), 149-173.

[21] Kleibergen, F. and Z. Zhan, (2015) Unexplained Factors and Their Effects on Second Pass R-squared's, *Journal of Econometrics,* 189, 101-116.

[22] Mikusheva, A., and L. Sun (2022) Inference with many weak instruments, *Review of Economic Studies* (forthcoming), https://academic.oup.com/restud/advance-article-abstract/doi/10.1093/restud/rdab097/6482756?redirectedFrom=fulltext.

[23] Nakamura, A. and M. Nakamura (1981) On the relationships among several specification error tests presented by Durbin, Wu, and Hausman, *Econometrica,* 49(6), 1583-1588.

[24] Onatski, A. (2015) Asymptotic analysis of the squared estimation error in misspecified factor models, *Journal of Econometrics*, 186:386-406.

[25] Petersen, M. (2009) Estimating standard errors in finance panel data sets: comparing approaches, *Review of Financial Studies,* 22(1), 435-480.

[26] Ramsey, J. B. (1969) Tests for specification errors in classical linear least squares regression analysis, Journal of the Royal Statistical Society, Series B. 31 (2): 350–371.

[27] Raponi, V, Robotti, C., and P. Zaffaroni (2020) Testing Beta-Pricing Models Using Large Cross-Sections, *Review of Financial Studies*, 33(6):2796–2842.

[28] Wu, D-M. (1973) Alternative tests of independence between stochastic regressors and disturbances, *Econometrica*, 41(4), 733–750.