# Drawbacks in the 3-Factor Approach of Fama and French (2018)$^{\star}$

David E. Allen[a,*], and Michael McAleer[b]

[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*
[b] *Department of Finance, College of Management, Asia University, Taiwan, Discipline of Business Analytics, University of Sydney Business School, Australia, Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam, The Netherlands, Department of Economic Analysis and ICAE, Complutense University of Madrid, Spain, Department of Mathematics and Statistics, University of Canterbury, New Zealand, and Institute of Advanced Sciences, Yokohama National University, Japan*

---

## Abstract

This paper features a statistical analysis of the monthly three factor Fama/French return series. We apply rolling OLS regressions to explore the relationship between the 3 factors, using monthly and weekly data from July 1926 to June 2018, that are freely available on French's website. The results suggest there are significant and time-varying relationships between the factors. This is conirmed by non-parametric tests. We then switch to a sub-sample from July 1990 to July 2018, also taken from French's website. The three series and their interrelationships are analysed using two stage least squares and the Hausman test to check for issues related to endogeneity, the Sargan over-identification test and the Cragg-Donald weak instrument test. The relationship between factors is also examined using OLS, incorporating Ramsey's RESET tests of functional form misspecification, plus Naradaya-Watson kernel regression techniques. The empirical results suggest that the factors, when combined in OLS regression analysis, as suggested by Fama and French (2018), are likely to suffer from endogeneity. OLS regression analysis and the application of Ramsey's RESET tests suggest a non-linear relationship exists between the three series, in which cubed terms are significant. This non-linearity is also confirmed by the kernel regression analysis. We use two instruments to estimate the market betas, and then use the factor estimates in a second set of panel data tests using a small sample of monthly returns for US firms that are drawn from the online data source "tingo". These issues are analysed using methods suggested by

---

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

Petersen (2009) to permit clustering in the panels by date and firm. The empirical results suggest that using an instrument to capture endogeneity reduces the standard error of market beta in subsequent cross-sectional tests, but that clustering effects, as suggested by Petersen (2009), will also impact on the estimated standard errors. 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 problematic in that the estimated standard errors are highly sensitive to the correct model specification.

## 1. Introduction

In a fundamental paper, Fama and French (1993, p.3), 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 testing the effects of various factors on various portfolios selected from global markets. Both Fama and French are Directors and advisors to a set of corporate entities under the rubric, Dimensional Fund Advisors, which applies factor models in a managed fund and investment advisory setting.

Cochrane (2011, p.1047), in his Presidential Address, delivered 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) respond to these new challenges by suggesting how to choose among 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) 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 (1989), hereafter GRS, produces a test of whether multiple factors add to a base model's explanatory power.

A perusal of GRS reveals that their test is based on the strong assumptions of linearity, independence and a Gaussian distribution. They proceed 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. Then they 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:

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

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

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

Thus, the 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.

In this comment, we apply simple tests of endogeneity, and independence to a set of monthly data taken from French's website featuring the Fama/French estimates of the excess return on the market portfolio, estimates of SMB and HML. 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).

For the purpose of providing an example, we use a sample of capitalization change adjusted company prices from the free on-line data source "tingo" (see: https://www.tiingo.com). We employed an R library package interface riingo, which provides an interface to the database, (see: https://cran.r-project.org/web/packages/riingo/index.html), and downloaded adjusted monthly price data for 21 companies. This data set of three time series of market factors, consisting of 220 monthly observations from January 2000 through to July 2018, and use a subset of the monthly data from January 2000 to the end of December 2010 comprising 132 observations, to estimate market factors. Tests of endogeneity using two stage least squares and the Hausman test are used, as Fama and French (2018) adopt a test proposed by Barillas and Shanken (BS, 2018).

Barillas and Shanken (2018) 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 $i$, $\Pi_i$, 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 $\Pi_i$ 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 $\Pi_i$ 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 $\Pi_i$ includes the factors of all model $i$'s competitors, the union of $\Pi_i$ and $f_i$, which they call $\Pi$, does not depend on $i$. This means that equation (5) can be simplified to:

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

They 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.

## 2. Endogeneity and related tests

Endogeneity broadly refers to situations in which an explanatory variable is correlated with the error term. The distinction between endogenous and exogenous variables originated in simultaneous equations models, where one separates variables whose values are determined by the model from variables which are determined outside the model. If simultaneity is ignored, then estimation will lead to biased estimates as it violates the exogeneity assumption of the Gauss–Markov theorem. Instrumental variable techniques are commonly used to address this problem. The Hausman test (also called the Durbin-Wu-Hausman specification test) is a statistical hypothesis test which evaluates the consistency (see Nakamura and Nakamura, 1981), of an estimator when compared to an alternative, less efficient but consistent estimator.

Consider the linear regression model $y = bX + e$, in which $y$ is the dependent variable, and $X$ is a vector of regressors, with error term, $e$. Under the null hypothesis, while both estimators are consistent, $b_0$ is efficient (with the smallest asymptotic variance), in the class of estimators containing $b_1$. Under the alternative hypothesis, $b_1$ is consistent, whereas $b_0$ is not.

The Hausman test statistic is given by:

$$H = (b_1 - b_0)^{'}(Var(b_1) - Var(b_0))^{\dagger}(b_1 - b_0), \qquad (7)$$

where $\dagger$ denotes the Moore–Penrose pseudo-inverse. The test statistic is distributed as a chi-squared distribution with the number of degrees of freedom equal to the rank of the matrix $Var(b_1 - b_0) = Var(b_1) - Var(b_0)$. Rejection of the null hypothesis suggests that $b_0$ is inconsistent.

Sargan (1958) developed a test based on the assumption that model parameters are identified via a priori restrictions on the coefficients, and tests the validity of over-identifying restrictions. The residuals from instrumental variables estimation can be used to form a test. This is done by constructing a quadratic form based on the cross-product of the residuals and exogenous variables. Under the null hypothesis that the over-identifying restrictions are valid, the statistic is asymptotically distributed as a chi-square variable with $(m - k)$ degrees of freedom $m$, is the number of instruments and $k$ is the number of endogenous variables. We apply the Sargan (1958) test.

The Cragg-Donald (1993) weak instrument test is constructed by considering the basic model shown below:

$$y = Y\beta + X\gamma + u$$

$$Y = Z\Pi + X + V,$$

where $y$ is the dependent variable of interest, $Y$ is an $N \times T$ matrix of endogenous variables, $Z$ is a matrix of $K_2$ excluded variables, and $X$ is a matrix of $K_1$ included instruments. The main concern is that the explanatory power of $Z$ may be insufficient to permit inference on $\beta$.

Let $P_W = W(W'W)^{-1}W'$ and $M_W = I - P_W$ for any matrix $W$. Let $W^\perp$ be the residuals from the projection on $X$, and so $W^\perp = M_X W$. If we define $\underline{Z} = [XZ]$ to be the matrix of all instruments (included and excluded), then the Cragg-Donald (1993) test statistic can be defined as:

$$G_T = (Y'\underline{Z}Y)^{-1/2}Y^{\perp'}P_{Z\perp}Y^\perp(Y'M_{\underline{Z}}Y)^{-1/2}\left(\frac{T-K_1-K_2}{K_2}\right)^2. \qquad (8)$$

The minimum eigenvalue of $G_T$ is the statistic used for testing for weak instruments.

We run a series of tests in which we explore the relationship between the factors themselves. We regress $SMB$ and $HML$ on $R_M - R_F$. As instruments, we use monthly OECD surveys of expected US manufacturing production, plus the monthly return on the VIX. This was the older version of the VIX (VXO), based on implied volatilities, rather than the new 'model free' VIX, which was introduced in 2003. We run individual time series regressions in which the RHS variables are the other two factors in the 3-factor model. The key issue is the endogeneity of the factors. If they are found to be endogenous, then OLS estimates in a multiple regression are likely to be biased and inconsistent. This would make the validity of the test recommended by Fama and French (2018) for choosing factors more sensitive to the validity of the regression specification used in estimating the factor loadings.

## 3. Endogeneity tests on Fama-French 3 factors

### 3.1. Some preliminaries

As a preliminary examination of the data set we downloaded the entire monthly series for the three basic factors, the excess return on the market $R_M - R_F$, $SMB$, and $HML$, from French's website. These comprised some 1102 observations from July 1926 to June 2018. We then ran rolling regressions using 60-month windows through the entire data set, applying OLS regressions of $SMB$ on $R_M - R_F$, $HML$ on $R_M - R_F$, and $HML$ on $SMB$. The results of these are shown in Figure 1.

The rolling OLS regression results, shown in Figure 1, have error bands plus and minus two standard deviations plotted above and below the regression estimate, which is indicated as "b". The upper band is marked "hi" and the lower band marked "lo".

The estimates of the rolling regression of $SMB$ on $R_M - R_F$ are shown in Figure 1 (a). An interesting feature in the diagram is the dynamic time-varying nature of this relationship, and all of the three plots shown in Figure 1. There is a significant relationship between $SMB$ and $R_M - R_F$ for the period from 1931/32 to 1951/52, and then the relationship becomes insignificant until 1962. It then becomes positive again, and significant, until around 1985. It is then insignificant for a period up to 2003, when it becomes positive and significant again, and remains so, until the end of the estimation period.

# Figure 1: Plots of rolling OLS regressions of factors $SMB$ on $R_M - R_F$, $HML$ on $R_M - R_F$, and $HML$ on $SMB$ using 60 month windows

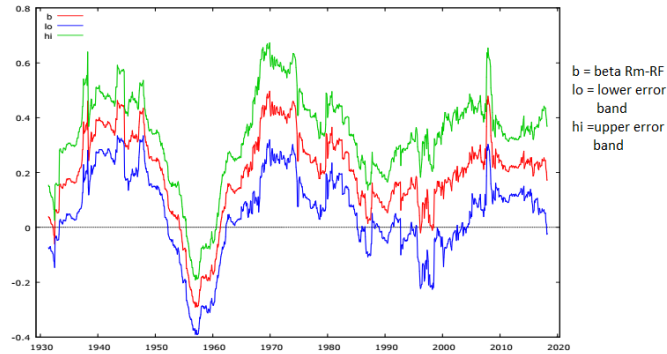Figure 1 (a): plot of regression of $SMB$ on $R_M - R_F$ using 60 month rolling window



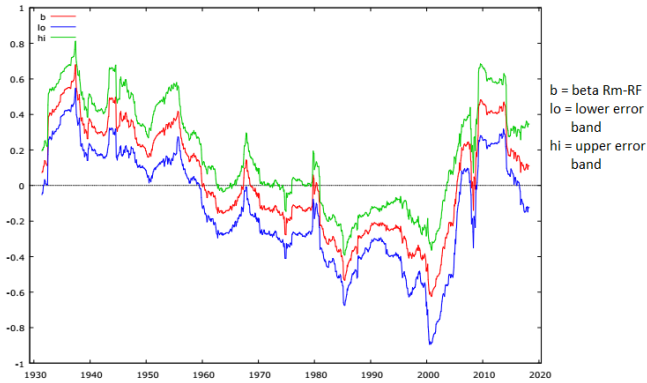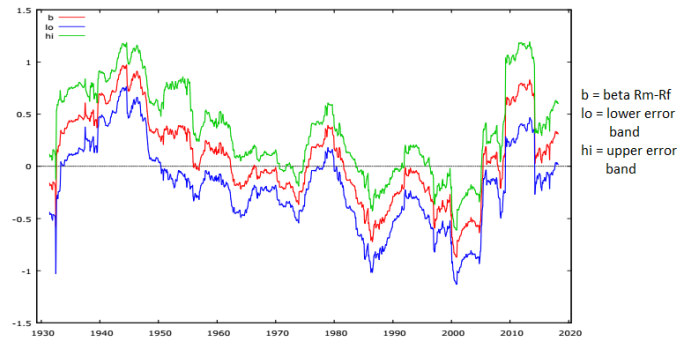Figure 1 (b): plot of regression of $HML$ on $R_M - R_F$ using 60 month rolling window



Figure 1(c): plot of regression of $HML$ on $SMB$ using 60 month rolling window

The OLS regression relationship between $HML$ and $R_M - R_F$, shown in Figure 1 (b), is even more variable. The plot reveals a significant positive relationship between these two variables between 1932 and 1960. The relationship between the two is then insignificant until 1980, at which point it becomes significant and negative. It remains significantly negative until around 2005, when it becomes significantly positive. In the period after 2005, it is predominantly significantly positively related to $R_M - R_F$, but there are short periods in which the relationship beomes insgnificant, and these occur around 2009 and 2012.

Figure 1 (c), the third diagram in the set of three, shows the OLS relationship between $HML$ and $SMB$. This relationship is significantly positive from 1932 to 1950. The relationship is then insignificant until about 1973, when there is a brief spell when it becomes significantly negative. The significant negative relationship re-occurs between 1982 and 1991, and then again from 1995 to 2004. Then from 2009 to 2014, it becomes significantly positive, which is followed by a period of insignificance.

We then repeated the exercise using weekly data for the same three factor series, which again was downloaded from French's website, for a period from the first week in July 1926 to the last week in October 2016, comprising a total of 4817 observations. We again ran bivariate rolling OLS regressions using a window of 52 weeks between the three factors and the results of these are shown in Figure 2. The results, using a one year window of weekly data, for the regression of $SMB$ on $R_M - R_F$, shown in Figure 2(a) reveal that there are significant positive relationships between these two factors in the late 1930s and for a long period in the 1940s. The relationship then becomes significant and negative for periods in the 1950s and 1960s. It becomes significant and positive again in the 1960s, and then switches signs to being negative significant followed by positive significant in the late 1970s and early 1980s. It then becomes negative and significant in the early and late 1990s. In the mid 2000s, it is positive and significant, and this significant positive relationship re-emerges twice in the period between 2010 and 2018.

The regression of $HML$ on $R_M - R_F$ using 52 week windows reveals a similar pattern of periods of prolonged significant positive and negative relationships. Figure 2(b) shows that from 1930 through to 1960 there is an almost unbroken significant positive relationship between these two factors. The significance and sign of the relationship is then reversed for prolonged intervals between the late 1950s and 2000. For an interval of several years leading up to 2010, a significant positive relationship emerges.

The relationship between $HML$ on $SMB$, as shown in Figure 2(c) is less pronounced. However, there are periods from the late 1930s to the late 1940s when a significant positive relationship emerges, whilst a significant relationship briefly emerges in the late 1960's early 1970s. The relationship switches to being positive and significant in the late 1970s and then becomes significantly negative in the mid 1990s and early 2000s.

## Figure 2: Plots of rolling OLS regressions of factors $SMB$ on $R_M - R_F$, $HML$ on $R_M - R_F$, and $HML$ on $SMB$ using 52 week windows July 1926 - Oct 2018

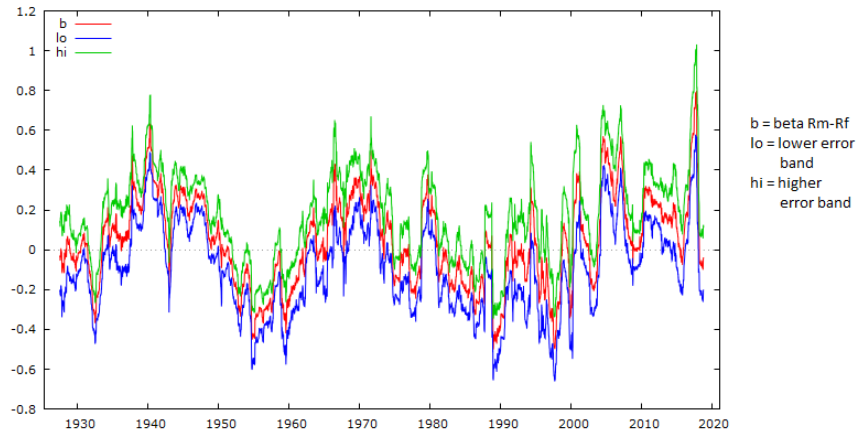Figure 2 (a): plot of regression of $SMB$ on $R_M - R_F$ using 52 week rolling window



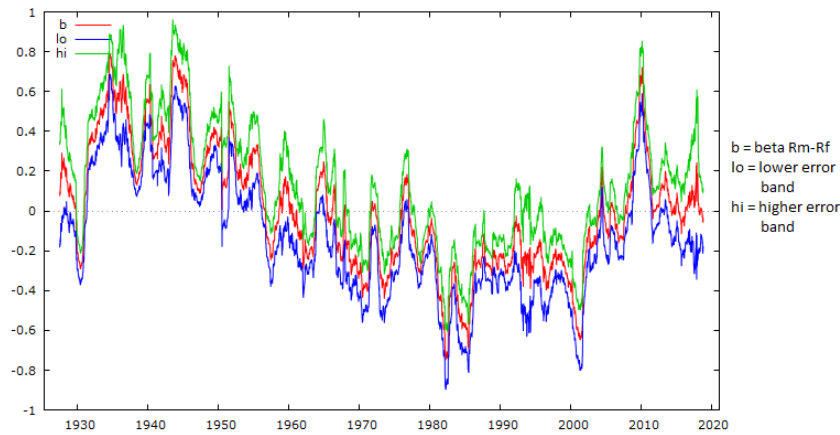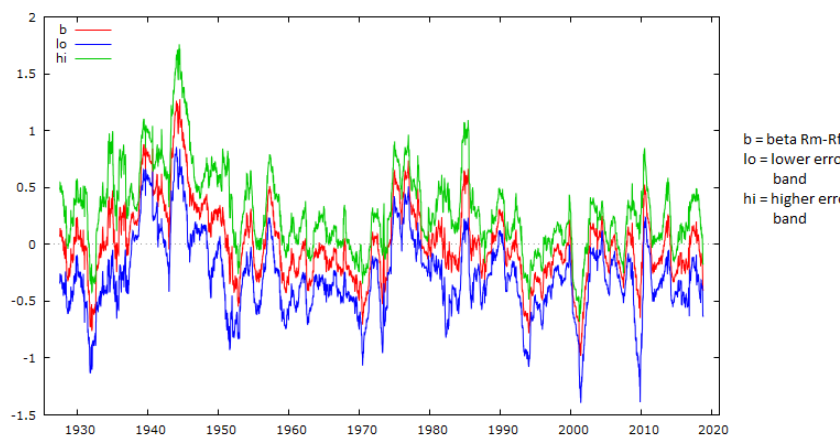Figure 2 (b): plot of regression of $HML$ on $R_M - R_F$ using 52 week rolling window



Figure 2(c): plot of regression of $HML$ on $SMB$ using 52 week rolling window

The results shown in Figures 1 and 2 suggest that, if the 3 factors are employed jointly in a time series regression, to estimate factor loadings then great care must be taken to check the relationships between the factors. Figures 1 and 2 show that the factors are not independent for long periods of time between 1926 and 2018. If they are employed as independent variables in a time series regression, they are likely to suffer from endogeneity.

As a further check we examined the relationship using a non-parametric measure for testing non-linear pairwise independence suggested by Massoumi and Racine (2002) which is available in the R library package 'np' as set out by Hayfield and Racine (2008). This tests the null of pairwise independence of two univariate density (or probability) functions. In the case of continuous variables we construct:

$$S_\rho = \frac{1}{2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (f_1^{1/2} - f_2^{1/2})^2 dxdy$$

$$= \frac{1}{2} \int \int \left(1 - \frac{f_2^{1/2}}{f_1^{1/2}}\right)^2 dF_1(x,y), \tag{9}$$

where $f_1 = f(x_i, y_i)$ is the joint density and $f_2 = g(x_i) \times h(y_i)$ is the product of the marginal densities of the random variables $X_i$ and $Y_i$. The unknown density/probability functions are replaced with nonparametric kernel estimates.

The bootstrap distribution is obtained by resampling with replacement from the empirical distribution of $X$ delivering $\{X_i, Y_i\}$ pairs under the null generated as $\{X_i^*, Y_i\}$ where $X^*$ is the bootstrap resample (i.e. we 'shuffle' $X$ leaving $Y$ unchanged thereby breaking any pairwise dependence to generate resamples under the null). Bandwidths are obtained via likelihood cross-validation by default for the marginal and joint densities.

We implemented this test using a measure of "predictability" for variable $Y$ and its predicted values $\hat{Y}$ (from our implemented model). In our case, our three models implemented were the linear OLS regressions estimated pairwise of the three Fama-French factors on one-another. The results of these tests using a bootstrap with 999 replications are shown in Table 1.

## Table 1: Non-parametric tests of the predictions of pairwise OLS regressions of $SMB$, $HML$ and $R_M - R_F$ on each other using monthly data 1926-2018

| Consistent metric entropy tests for dependence | | |
|---|---|---|
| | Test statistic $Srho$ | Probability |
| $lm(SMB \sim Rm - Rf)$ | 0.007442087 | < 2.22e-16 *** |
| $lm(HML \sim Rm - Rf)$ | 0.00769992 | < 2.22e-16 *** |
| $lm(HML \sim SMB)$ | 0.007699925 | < 2.22e-16 *** |

NB:*** indicates null of independence is rejected at the 0.1% level

The results in Table 1 reveal that non-parametric bootstrapped tests of the null of independence between the three Fama-French factors series, using the

predictions obtained via OLS and the full sample period from July 1926 to June 2018, comprising some 1102 observations, reject the null of indepencence, in all three pairwise cases, at better than the 1 per cent level..

*3.2. Further tests using a subset of the data*

We then set up further simple tests, using a more recent sub-set of the data, using monthly 3 factor Fama-French return series, from July 1990 to July 2018 available on French's website, together with the monthly excess return on the US market, $R_M - R_F$, employed as the independent variable, in a set of time series regressions, also taken from French's website. To check for endogeneity and to estimate the regression equations using two stage least squares, we need suitable instrumental variables that are independently related to some of the factors. We chose "Business Tendency Surveys for Manufacturing: Confidence Indicators: Composite Indicators: European Commission and National Indicators for the United States", (BSCICP02USM460S), which is an OECD monthly indicator series. This series is available on the Federal Reserve Bank of St. Louis (FRED) database, and features the results of surveys of confidence in US Manufacturing. We also used the return on the older version of the VIX (VXO), based on implied volatilities, rather than the new 'model free' VIX, which was introduced in 2003. This was because the data set commences in 2000.

Given the evidence discussed in Figures 1 and 2, we first explored whether the factors used in the typical Fama-French regression are related, for this smaller sample period, by regressing $SMB$ and $HML$ on the market factor $R_M - R_F$. The results of these regressions are shown in Table 2, which reveal a significant relationship between $SMB$ and $R_M - R_F$. Simlar to the results in Figures 1 and 2, this suggests that they are likely to suffer from an endogeneity problem if they are used as explanatory variables in time series regressions.

## Table 2: Results of a Regression of $SMB$ and $HML$ on $R_M - R_F$

OLS, using observations 2000:02–2010:12'($T = 132$)
Dependent variable: SMB

|        | Coefficient | Std. Error | $t$-ratio | p-value |
|--------|-------------|------------|-----------|---------|
| const  | 0.554839    | 0.283955   | 1.954     | 0.0529  |
| RM     | 0.243333    | 0.0581558  | 4.184     | 0.0001*** |

| | | | |
|---|---|---|---|
| Mean dependent var | 0.567786 | S.D. dependent var | 3.449989 |
| Sum squared resid | 1362.415 | S.E. of regression | 3.249824 |
| $R^2$ | 0.119497 | Adjusted $R^2$ | 0.112672 |
| $F(1, 129)$ | 17.50725 | P-value($F$) | 0.000053 |
| Log-likelihood | $-339.2700$ | Akaike criterion | 682.5399 |
| Schwarz criterion | 688.2903 | Hannan–Quinn | 684.8766 |
| $\hat{\rho}$ | $-0.245089$ | Durbin–Watson | 2.354554 |

OLS, using observations 2000:02–2010:12 ($T = 131$)
Dependent variable: HML

|        | Coefficient | Std. Error | $t$-ratio | p-value |
|--------|-------------|------------|-----------|---------|
| const  | 0.609319    | 0.402326   | 1.514     | 0.1323  |
| RM     | 0.151964    | 0.0823989  | 1.844     | 0.0674  |

| | | | |
|---|---|---|---|
| Mean dependent var | 0.617405 | S.D. dependent var | 4.646895 |
| Sum squared resid | 2735.059 | S.E. of regression | 4.604564 |
| $R^2$ | 0.025689 | Adjusted $R^2$ | 0.018136 |
| $F(1, 129)$ | 3.401238 | P-value($F$) | 0.067442 |
| Log-likelihood | $-384.9165$ | Akaike criterion | 773.8330 |
| Schwarz criterion | 779.5834 | Hannan–Quinn | 776.1697 |
| $\hat{\rho}$ | 0.050809 | Durbin–Watson | 1.722185 |

| Ramsey RESET Test | | | | |
|---|---|---|---|---|
|  | Coefficient | Std.Error | t-ratio | p-value |
| Const. | 0.0716255 | 0.609149 | 0.1176 | 0.9066 |
| RM | -0.158218 | 0.137932 | $-1.147$ | 0.2535 |
| y-hat^2 | $-0.690156$ | 0.589792 | $-1.170$ | 0.2441 |
| y-hat^3 | 1.15914 | 0.391449 | 2.961 | 0.0037*** |
| Test statistic: | F = 4.531926 | - | - | 0.0126** |

Note: ***Significant at 1%, **Significant at 5%

The results in Table 2 show that $HML$ and $R_M - R_F$ are not independent, but the relationship is significant at the 0.06 level. Figure 1 (b), shows that there are periods between 2000 and 2010 in which the relationship is significant,

when using rolling regressions.

However, a further check was undertaken by estimating the regression using squared terms and cubed terms, which provided evidence of an even more significant relationship. The original presumption was that the relationship was linear. The results, shown in Table 3, indicate the contrary, and suggest there is a significant relationship between $HML$ and $R_M - R_F$, cubed, and that contemporaneous and cubed lags of $R_M - R_F$ at one month, four months, and six months, are significant. The square terms of $R_M - R_F$ , run in a separate regression, were not significant until the sixth lag, and had a much lower adjusted R-square value. These results are not reported. The results shown in Table 3 suggest that there is a considerable potential endogeneity problem in the typical Fama-French time series regression, at least for this sub-sample of the data. They also suggest that relationships between the factors are not necessarily linear.

## Table 3: Results of the regression of $HML$ on $R_M - R_F$ using cubed terms, together with lags

OLS, using observations 2000:08–2010:12 ($T = 125$)
Dependent variable: HML

|           | Coefficient    | Std. Error   | $t$-ratio | p-value     |
|-----------|----------------|--------------|-----------|-------------|
| const     | 0.661977       | 0.291508     | 2.271     | 0.0250**    |
| CUBRM     | 0.00109952     | 0.000549371  | 2.001     | 0.0477**    |
| CUBRM_1   | 0.00140986     | 0.000558294  | 2.525     | 0.0129**    |
| CUBRM_2   | −0.000834069   | 0.000551494  | −1.512    | 0.1331      |
| CUBRM_3   | 0.000755538    | 0.000544106  | 1.389     | 0.1676      |
| CUBRM_4   | 0.000991569    | 0.000558401  | 1.776     | 0.0784      |
| CUBRM_5   | 5.74668e−005   | 0.000564860  | 0.1017    | 0.9191      |
| CUBRM_6   | −0.00263963    | 0.000559278  | −4.720    | 0.0000***   |

|                    |            |                   |            |
|--------------------|------------|-------------------|------------|
| Mean dependent var | 0.626320   | S.D. dependent var | 3.751008   |
| Sum squared resid  | 1151.345   | S.E. of regression | 3.136966   |
| $R^2$              | 0.340085   | Adjusted $R^2$    | 0.300603   |
| $F(7, 117)$        | 8.613654   | P-value($F$)      | 1.71e–08   |
| Log-likelihood     | −316.1406  | Akaike criterion  | 648.2812   |
| Schwarz criterion  | 670.9077   | Hannan–Quinn      | 657.4732   |
| $\hat{\rho}$       | 0.256632   | Durbin–Watson     | 1.477293   |

Note: ***Significant at 1%, **Significant at 5%

## Table 4: Regressions of single factors on the instrument USPROD

OLS, using observations 2000:02–2010:12 ($T = 131$)
Dependent variable: RM-RF

|  | Coefficient | Std. Error | $t$-ratio | p-value |
|---|---|---|---|---|
| const | −0.395759 | 0.426653 | −0.9276 | 0.3554 |
| USPROD | 0.122581 | 0.0332649 | 3.685 | 0.0003*** |

| Mean dependent var | 0.053206 | S.D. dependent var | 4.901120 |
|---|---|---|---|
| Sum squared resid | 2825.319 | S.E. of regression | 4.679925 |
| $R^2$ | 0.095240 | Adjusted $R^2$ | 0.088226 |
| $F(1, 129)$ | 13.57921 | P-value($F$) | 0.000335 |
| Log-likelihood | −387.0432 | Akaike criterion | 778.0864 |
| Schwarz criterion | 783.8368 | Hannan–Quinn | 780.4230 |
| $\hat{\rho}$ | 0.102879 | Durbin–Watson | 1.784133 |

OLS, using observations 2000:02–2010:12 ($T = 131$)
Dependent variable: SMB

|  | Coefficient | Std. Error | $t$-ratio | p-value |
|---|---|---|---|---|
| const | 0.550760 | 0.315697 | 1.745 | 0.0834 |
| USPROD | 0.00464879 | 0.0246140 | 0.1889 | 0.8505 |

| Mean dependent var | 0.567786 | S.D. dependent var | 3.449989 |
|---|---|---|---|
| Sum squared resid | 1546.888 | S.E. of regression | 3.462857 |
| $R^2$ | 0.000276 | Adjusted $R^2$ | -0.007473 |
| $F(1, 129)$ | 0.035671 | P-value($F$) | 0.850493 |
| Log-likelihood | −347.5875 | Akaike criterion | 699.1751 |
| Schwarz criterion | 704.9255 | Hannan–Quinn | 701.5117 |
| $\hat{\rho}$ | −0.160016 | Durbin–Watson | 2.192858 |

OLS, using observations 2000:02–2010:12 ($T = 131$)
Dependent variable: HML

|  | Coefficient | Std. Error | $t$-ratio | p-value |
|---|---|---|---|---|
| const | 0.607982 | 0.425270 | 1.430 | 0.1552 |
| USPROD | 0.00257252 | 0.0331571 | 0.07759 | 0.9383 |

| Mean dependent var | 0.617405 | S.D. dependent var | 4.646895 |
|---|---|---|---|
| Sum squared resid | 2807.041 | S.E. of regression | 4.664762 |
| $R^2$ | 0.000047 | Adjusted $R^2$ | -0.007705 |
| $F(1, 129)$ | 0.006020 | P-value($F$) | 0.938278 |
| Log-likelihood | −386.6181 | Akaike criterion | 777.2361 |
| Schwarz criterion | 782.9865 | Hannan–Quinn | 779.5728 |
| $\hat{\rho}$ | 0.075600 | Durbin–Watson | 1.680310 |

Note: ***Significant at 1%, **Significant at 5%

## Table 5: Regressions of single factors on the instrument return on VXO

OLS, using observations 2000:02–2010:12 ($T = 131$)
Dependent variable: RM

|  | Coefficient | Std. Error | $t$-ratio | p-value |
|---|---|---|---|---|
| const | 0.00109581 | 0.299529 | 0.003658 | 0.9971 |
| VRET | −20.0490 | 1.71443 | −11.69 | 0.0000*** |

| | | | |
|---|---|---|---|
| Mean dependent var | 0.053206 | S.D. dependent var | 4.901120 |
| Sum squared resid | 1515.799 | S.E. of regression | 3.427882 |
| $R^2$ | 0.514591 | Adjusted $R^2$ | 0.510829 |
| $F(1, 129)$ | 136.7555 | P-value($F$) | 5.51e–22 |
| Log-likelihood | −346.2577 | Akaike criterion | 696.5155 |
| Schwarz criterion | 702.2658 | Hannan–Quinn | 698.8521 |
| $\hat{\rho}$ | 0.131040 | Durbin–Watson | 1.736595 |

OLS, using observations 2000:02–2010:12 ($T = 131$)
Dependent variable: SMB

|  | Coefficient | Std. Error | $t$-ratio | p-value |
|---|---|---|---|---|
| const | 0.554278 | 0.291876 | 1.899 | 0.0598 |
| VRET | −5.19730 | 1.67063 | −3.111 | 0.0023*** |

| | | | |
|---|---|---|---|
| Mean dependent var | 0.567786 | S.D. dependent var | 3.449989 |
| Sum squared resid | 1439.330 | S.E. of regression | 3.340298 |
| $R^2$ | 0.069789 | Adjusted $R^2$ | 0.062578 |
| $F(1, 129)$ | 9.678239 | P-value($F$) | 0.002296 |
| Log-likelihood | −342.8671 | Akaike criterion | 689.7342 |
| Schwarz criterion | 695.4846 | Hannan–Quinn | 692.0709 |
| $\hat{\rho}$ | −0.210599 | Durbin–Watson | 2.288881 |

OLS, using observations 2000:02–2010:12 ($T = 131$)
Dependent variable: HML

|  | Coefficient | Std. Error | $t$-ratio | p-value |
|---|---|---|---|---|
| const | 0.614042 | 0.407131 | 1.508 | 0.1339 |
| VRET | −1.29366 | 2.33032 | −0.5551 | 0.5798 |

| | | | |
|---|---|---|---|
| Mean dependent var | 0.617405 | S.D. dependent var | 4.646895 |
| Sum squared resid | 2800.482 | S.E. of regression | 4.659309 |
| $R^2$ | 0.002383 | Adjusted $R^2$ | -0.005350 |
| $F(1, 129)$ | 0.308182 | P-value($F$) | 0.579759 |
| Log-likelihood | −386.4648 | Akaike criterion | 776.9297 |
| Schwarz criterion | 782.6801 | Hannan–Quinn | 779.2663 |
| $\hat{\rho}$ | 0.067245 | Durbin–Watson | 1.696344 |

Note: ***Significant at 1%, **Significant at 5%

As a potential instrument, the OECD monthly "Business Tendency Surveys for Manufacturing: Confidence Indicators: Composite Indicators: European Commission and National Indicators for the United States", (BSCICP02USM460S), was chosen.

We establish that this is a relevant instrument by time series regressions of the Fama-French factors on this series. The results are shown in Table 4. This instrument is related to $R_M - R_F$, but not to $SMB$ and $HML$. As a second instrument, we used the return on the VXO. The results of regressing the three factors on this instrument are shown in Table 5 and show that there is a significant relationship between $R_M - R_F$ and $SMB$ and the return on VXO. However, $HML$ is not related to the return on VXO.

We can investigate the endogeneity problem using these instruments. We ignore the potential non-linearity of the relationship in subsequent analysis to preserve consistency with attempts to test asset pricing models, though the issue of non-linearity is significant, as suggested by kernel regressions that are discussed in the next paragraph.

The results also suggest that there is not necessarily a linear relationship between the explanatory factors. Furthermore, we ran a Naradaya and Watson kernel regression. Nadaraya (1964) and Watson (1964) proposed to estimate $m$ as a locally weighted average, using a kernel as a weighting function. The Nadaraya–Watson estimator is:

$$\hat{M}_h(x) = \frac{\sum_{i=1}^{N} K_h(x - x_i)y_i}{\sum_{j=1}^{N} K_h(x - x_j)}, \tag{10}$$

where $K$ is a kernel with a bandwidth $h$. The denominator is a weighting term with sum 1. A kernel is a weighting function used in non-parametric estimation techniques. Kernels are used in kernel density estimation to estimate random variables' density functions, or in kernel regression to estimate the conditional expectation of a random variable. We apply kernel estimation as a further check on the linearity of the relationship between the 3 factors. The results are shown in Figure 3.

### 3.3.  OLS and Two Stage Least Squares results, and Hausman tests

The next step is to compare the customary method adopted in asset pricing tests: namely time series regressions using OLS as a means to estimate betas, with two stage least squares, including the use of instruments, and tests of endogeneity using the Hausman test. This next step requires some company return series.

These preliminary results suggest there is a potential endogeneity problem with an OLS time series model, and the regression of returns on a stock or portfolio to estimate their factor loadings or betas in a 3-factor setting.

To assess the extent of the problem, we downloaded a sample of capitalization change adjusted company prices from the free on-line data source

"tingo" (see: https://www.tiingo.com). We used an R library package interface riingo which provides an interface to the database (see: https://cran.r-project.org/web/packages/riingo/index.html). We downloaded adjusted monthly price data for 21 companies, which consisted of Apple and IBM prices, plus the first 19 series listed on their index of companies. The companies used are shown in Table 5. We constructed continuously compounded return series using this price data, where the sample is from January 31, 2011 to the end of December 2017.

We use standard time series linear regressions to estimate a 3-factor model by OLS using data from February 2000 until the end of December 2010. The results of these estimates are shown in Tables 7, 8, 9, and 10. The results appear to be satisfactory, in that of the total of 21 time series regressions, the coefficients on $R_M - R_F$ are significant at the 5 per cent level or better on 16 occasions. $SMB$ has 5 significant coefficients and $HML$ 7 in total. However, the previous analysis shows that these regressions suffer from an endogeneity problem.

We re-estimated the time series beta estimates for the 3-factor model using instrumental variables. The results are shown in Tables 11, 12, 13, and 14. These regressions, undertaken using two stage least squares, with the lagged instrument based on expectations of US Production, while not biased, are even stronger.

If we consider the time series estimates, and the beta coefficients estimated on the market factor, $R_M - R_F$, of the 21 regression estimates, 17 are significant at the 5% level or better. $SMB$ has 7 significant coefficients and $HML$ has 9. Thus, there are more significant coefficients than in the simple time series regressions.

In Table 15 we report the results of tests of the power of the instruments, as applied to the instrumented regressions. Only 3 of the 21 regressions fail the Hausman test. All of the regressions, except 2, pass the Sargan over-identification test. The Cragg-Donald tests uniformly suggest that the TSLS regressions have bias less than 5 per cent in relation to OLS. These results suggest that the instruments are satisfactory.

However, we compared the estimated slope coefficients from the time series estimation of factor loadings using OLS, and those from the estimates using two instrumental variables with one lag, to adjust for the endogeneity problem, plus the application of two stage least squares, and used non-parametric sign tests to examine whether there are any significant differences between the two sets of estimates.

The results, which are reported in Table 16, suggest that there are significant differences in the estimates of the loadings on the excess market return $R_M - R_F$, and on $SMB$, while there is no significant difference in the loading on $HML$. This is reassuring, in that the use of the instruments focused on the excess market return $R_M - R_F$ and $SMB$, while $HML$ was on the borderline of being endogenous.

## Figure 3: Plot of Naradaya/Watson kernel regression analysis of the relationship between factors
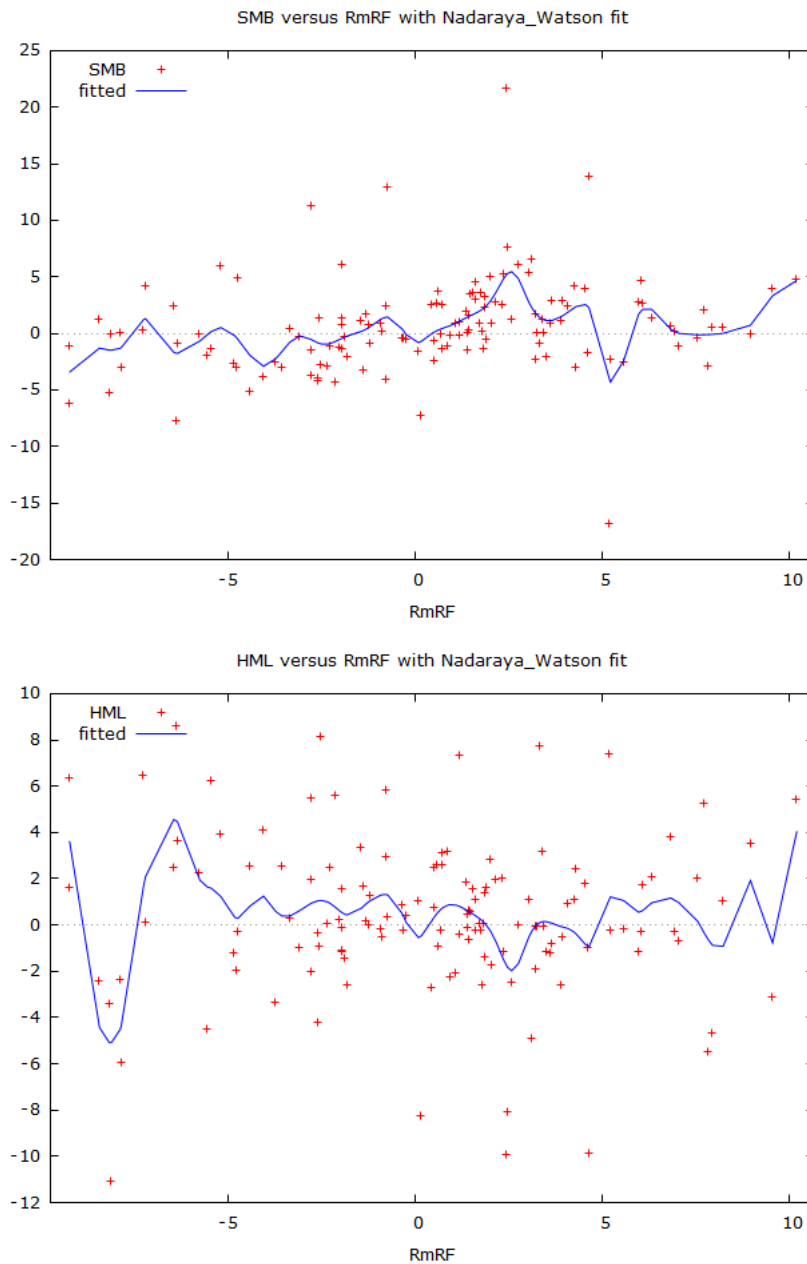
**Table 6: Company Sample**

|    | Company | Code |
|----|---------|------|
|    |         |      |
| 1  | APPLE.INC | AAPL |
| 2  | INERNATIONAL BUSINESS MACHINES CO | IBM |
| 3  | AGILENT TECHNOLOGIES INC | A |
| 4  | YAHOO.INC | AABA |
| 5  | ALABAMA AIRCRAFT INDUSTRIES | AAIIQ |
| 6  | ATLANTIC AMERICA CORP | AAME |
| 7  | ARMADA MERCANTILE LTD | AAMTF |
| 8  | AARON'S INC | AAN |
| 9  | AAON. INC | AAON |
| 10 | AMER-PETRO HUNTER.INC | AAPH |
| 11 | ALL-AMERICAN SPORTPARK.INC | AASP |
| 12 | ALLIANCEBERNSTEIN HOLDING L.P. | AB |
| 13 | ABAXIS.INC | ABAX |
| 14 | AMERIS BANCORP | ABCB |
| 15 | ABEO | ABEO |
| 16 | AMBEV SA | ABEV |
| 17 | ARKANSAS BEST CORP | ABFS |
| 18 | ARCA BIOPHARMA INC | ABIO |
| 19 | ABM INDUSTRIES INC | ABM |
| 20 | ABBOTT LABORATORIES | ABT |
| 21 | AUTOBYTEL INC | ABTL |

**Table 7: Results of Time Series Factor Estimates Contd**

| | Dependent variable: | | | | | |
|---|---|---|---|---|---|---|
| | i ~RM + SMB + HML | | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| RM | 0.020*** | 0.012*** | 0.019*** | 0.021*** | -0.0002 | 0.010*** |
| | (0.003) | (0.001) | (0.002) | (0.003) | (0.004) | (0.003) |
| | | | | | | |
| SMB | -0.0005 | -0.004*** | 0.007** | -0.005* | 0.005 | 0.001 |
| | (0.003) | (0.001) | (0.003) | (0.003) | (0.004) | (0.003) |
| | | | | | | |
| HML | -0.011*** | -0.006*** | -0.006* | -0.015*** | 0.012** | 0.006 |
| | (0.003) | (0.002) | (0.003) | (0.003) | (0.005) | (0.004) |
| | | | | | | |
| Constant | 0.014 | 0.002 | -0.015 | -0.013 | -0.021 | -0.009 |
| | (0.011) | (0.006) | (0.010) | (0.011) | (0.017) | (0.013) |
| | | | | | | |
| Observations | 131 | 131 | 131 | 131 | 131 | 131 |
| $R^2$ | 0.366 | 0.389 | 0.422 | 0.400 | 0.040 | 0.094 |
| Adjusted $R^2$ | 0.351 | 0.375 | 0.408 | 0.386 | 0.018 | 0.073 |
| Residual Std. Error (df = 127) | 0.124 | 0.065 | 0.115 | 0.122 | 0.190 | 0.142 |
| F Statistic (df = 3; 127) | 24.449*** | 27.004*** | 30.900*** | 28.240*** | 1.773 | 4.392*** |

Note:      *p<0.1; **p<0.05; ***p<0.01

**Table 8: Results of Time Series Factor Estimates Contd**

|  | *Dependent variable:* | | | | | |
|---|---|---|---|---|---|---|
|  | i ~RM + SMB + HML | | | | | |
|  | (7) | (8) | (9) | (10) | (11) | (12) |
| RM | 0.006 | -0.001 | 0.007*** | 0.024** | 0.001 | 0.014*** |
|  | (0.006) | (0.002) | (0.002) | (0.011) | (0.010) | (0.002) |
| SMB | 0.001 | -0.001 | 0.003 | 0.019 | 0.015 | 0.002 |
|  | (0.006) | (0.002) | (0.002) | (0.012) | (0.010) | (0.002) |
| HML | 0.006 | -0.003 | 0.004 | 0.010 | 0.006 | 0.007*** |
|  | (0.007) | (0.002) | (0.003) | (0.014) | (0.012) | (0.002) |
| Constant | -0.010 | 0.013 | 0.008 | -0.021 | -0.025 | -0.009 |
|  | (0.024) | (0.008) | (0.008) | (0.048) | (0.041) | (0.007) |
| Observations | 131 | 131 | 131 | 131 | 131 | 131 |
| $R^2$ | 0.014 | 0.017 | 0.127 | 0.065 | 0.017 | 0.425 |
| Adjusted $R^2$ | -0.010 | -0.007 | 0.107 | 0.043 | -0.007 | 0.411 |
| Residual Std. Error (df = 127) | 0.270 | 0.091 | 0.093 | 0.536 | 0.462 | 0.078 |
| F Statistic (df = 3; 127) | 0.585 | 0.720 | 6.177*** | 2.948** | 0.715 | 31.258*** |

*Note:*      $^*$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01

**Table 9: Results of Time Series Factor Estimates Contd**

|  | *Dependent variable:* | | | | | |
|---|---|---|---|---|---|---|
|  | i ~RM + SMB + HML | | | | | |
|  | (13) | (14) | (15) | (16) | (17) | (18) |
| RM | 0.013*** | 0.007*** | 0.018*** | 0.010*** | 0.012*** | 0.023*** |
|  | (0.003) | (0.002) | (0.005) | (0.002) | (0.002) | (0.006) |
|  |  |  |  |  |  |  |
| SMB | 0.002 | 0.010*** | -0.003 | 0.001 | -0.0001 | 0.021*** |
|  | (0.003) | (0.002) | (0.006) | (0.002) | (0.002) | (0.006) |
|  |  |  |  |  |  |  |
| HML | 0.0003 | 0.016*** | -0.009 | 0.002 | 0.014*** | -0.011 |
|  | (0.003) | (0.002) | (0.007) | (0.003) | (0.003) | (0.007) |
|  |  |  |  |  |  |  |
| Constant | 0.002 | -0.011 | -0.021 | 0.016* | -0.003 | -0.068*** |
|  | (0.011) | (0.008) | (0.023) | (0.009) | (0.010) | (0.024) |
|  |  |  |  |  |  |  |
| Observations | 131 | 131 | 131 | 131 | 131 | 131 |
| $R^2$ | 0.180 | 0.370 | 0.091 | 0.180 | 0.303 | 0.263 |
| Adjusted $R^2$ | 0.161 | 0.355 | 0.069 | 0.160 | 0.286 | 0.246 |
| Residual Std. Error (df = 127) | 0.125 | 0.091 | 0.259 | 0.098 | 0.106 | 0.268 |
| F Statistic (df = 3; 127) | 9.322*** | 24.829*** | 4.218*** | 9.266*** | 18.371*** | 15.126*** |

*Note:*        *p<0.1; **p<0.05; ***p<0.01

**Table 10: Results of Time Series Factor Estimates Contd.**

|  | *Dependent variable:* | | |
|---|---|---|---|
|  | i ~RM + SMB + HML | | |
|  | (19) | (20) | (21) |
| RM | 0.007*** | 0.002* | 0.011*** |
|  | (0.002) | (0.001) | (0.004) |
| SMB | 0.009*** | -0.002 | 0.006 |
|  | (0.002) | (0.001) | (0.004) |
| HML | 0.008*** | 0.0002 | 0.006 |
|  | (0.002) | (0.002) | (0.005) |
| Constant | -0.002 | 0.006 | -0.030* |
|  | (0.007) | (0.005) | (0.017) |
| Observations | 131 | 131 | 131 |
| $R^2$ | 0.336 | 0.031 | 0.090 |
| Adjusted $R^2$ | 0.321 | 0.008 | 0.069 |
| Residual Std. Error (df = 127) | 0.075 | 0.061 | 0.188 |
| F Statistic (df = 3; 127) | 21.455*** | 1.366 | 4.207*** |

*Note:*          *p<0.1; **p<0.05; ***p<0.01

**Table 11: Results of Time Series Instrumental Variables Regression Factor Estimates**

|  | Dependent variable: | | | | | |
|  | i ~RM + SMB + HML \|USPROD + USPRODL1 + VRET + VRETL1 | | | | | |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| RM | 0.017*** | 0.011*** | 0.018*** | 0.018*** | -0.006 | 0.009*** |
|  | (0.002) | (0.001) | (0.002) | (0.002) | (0.004) | (0.003) |
| SMB | 0.004 | -0.002 | 0.010*** | 0.005 | 0.010* | -0.001 |
|  | (0.004) | (0.002) | (0.003) | (0.003) | (0.005) | (0.004) |
| HML | -0.006** | 0.0004 | -0.005** | -0.005** | 0.009** | 0.003 |
|  | (0.002) | (0.001) | (0.002) | (0.002) | (0.004) | (0.003) |
| Constant | 0.020* | 0.003 | -0.007 | -0.012 | -0.025 | -0.004 |
|  | (0.011) | (0.006) | (0.010) | (0.011) | (0.017) | (0.013) |
| Observations | 131 | 131 | 131 | 131 | 131 | 131 |
| $R^2$ | 0.338 | 0.437 | 0.482 | 0.384 | 0.064 | 0.113 |
| Adjusted $R^2$ | 0.322 | 0.424 | 0.470 | 0.370 | 0.042 | 0.092 |
| Residual Std. Error (df = 127) | 0.126 | 0.062 | 0.108 | 0.124 | 0.187 | 0.141 |

*Note:*          *p<0.1; **p<0.05; ***p<0.01

**Table 12: Results of Time Series Instrumental Variables Regression Factor Estimates Contd.**

|  | *Dependent variable:* | | | | | |
|---|---|---|---|---|---|---|
|  | i ˜RM + SMB + HML \|USPROD + USPRODL1 + VRET + VRETL1 | | | | | |
|  | (7) | (8) | (9) | (10) | (11) | (12) |
| RM | 0.010** | -0.001 | 0.006*** | 0.020* | 0.00004 | 0.014*** |
|  | (0.005) | (0.002) | (0.002) | (0.010) | (0.009) | (0.001) |
| SMB | 0.006 | 0.0003 | 0.005* | 0.029* | 0.032** | 0.002 |
|  | (0.007) | (0.003) | (0.002) | (0.015) | (0.013) | (0.002) |
| HML | -0.003 | -0.002 | 0.005** | 0.002 | 0.002 | 0.005*** |
|  | (0.005) | (0.002) | (0.002) | (0.010) | (0.009) | (0.001) |
| Constant | -0.007 | 0.012 | 0.008 | -0.013 | -0.033 | -0.004 |
|  | (0.024) | (0.008) | (0.008) | (0.048) | (0.041) | (0.006) |
| Observations | 131 | 131 | 131 | 131 | 131 | 131 |
| $R^2$ | 0.054 | 0.012 | 0.192 | 0.089 | 0.056 | 0.615 |
| Adjusted $R^2$ | 0.031 | -0.011 | 0.173 | 0.067 | 0.034 | 0.606 |
| Residual Std. Error (df = 127) | 0.264 | 0.091 | 0.090 | 0.529 | 0.452 | 0.064 |

*Note:*        *p<0.1; **p<0.05; ***p<0.01

**Table 13: Results of Time Series Instrumental Variables Regression Factor Estimates Contd.**

|  | Dependent variable: | | | | | |
|---|---|---|---|---|---|---|
|  | i ~RM + SMB + HML \|USPROD + USPRODL1 + VRET + VRETL1 | | | | | |
|  | (13) | (14) | (15) | (16) | (17) | (18) |
| RM | 0.012*** | 0.006*** | 0.018*** | 0.009*** | 0.008*** | 0.019*** |
|  | (0.002) | (0.002) | (0.005) | (0.002) | (0.002) | (0.005) |
| SMB | 0.002 | 0.009*** | 0.005 | 0.0004 | 0.002 | 0.034*** |
|  | (0.003) | (0.003) | (0.007) | (0.003) | (0.003) | (0.007) |
| HML | -0.003 | 0.010*** | -0.008* | 0.004* | 0.010*** | -0.014*** |
|  | (0.002) | (0.002) | (0.005) | (0.002) | (0.002) | (0.005) |
| Constant | 0.009 | -0.007 | -0.017 | 0.019** | -0.0004 | -0.060*** |
|  | (0.011) | (0.008) | (0.023) | (0.009) | (0.010) | (0.023) |
| Observations | 131 | 131 | 131 | 131 | 131 | 131 |
| $R^2$ | 0.195 | 0.354 | 0.139 | 0.201 | 0.281 | 0.354 |
| Adjusted $R^2$ | 0.176 | 0.338 | 0.118 | 0.183 | 0.264 | 0.339 |
| Residual Std. Error (df = 127) | 0.124 | 0.092 | 0.252 | 0.096 | 0.108 | 0.251 |

Note:    *p<0.1; **p<0.05; ***p<0.01

**Table 14: Results of Time Series Instrumental Variables Regression Factor Estimates Contd.**

| | *Dependent variable:* | | |
|---|---|---|---|
| | i ~RM + SMB + HML \|USPROD + USPRODL1 + VRET + VRETL1 | | |
| | (19) | (20) | (21) |
| RM | 0.008*** | 0.003** | 0.008** |
| | (0.001) | (0.001) | (0.004) |
| SMB | 0.006*** | -0.005*** | 0.016*** |
| | (0.002) | (0.002) | (0.005) |
| HML | 0.002 | -0.0001 | 0.005 |
| | (0.001) | (0.001) | (0.004) |
| Constant | 0.004 | 0.008 | -0.032* |
| | (0.007) | (0.005) | (0.016) |
| Observations | 131 | 131 | 131 |
| $R^2$ | 0.306 | 0.084 | 0.169 |
| Adjusted $R^2$ | 0.290 | 0.062 | 0.150 |
| Residual Std. Error (df = 127) | 0.077 | 0.059 | 0.180 |

*Note:*          *p<0.1; **p<0.05; ***p<0.01

**Table 15:   Results of Hausman test, Sargan over-identification test and Cragg-Donald weak instrument test**

Panel A

|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Hausman test: null $H_0$ OLS estimates consistent probability | 0.77 | 0.21 | 0.84 | 0.60 | 0.26 | 0.51 | 0.35 | 0.58 | 0.13 | 0.27 |
| Sargan over-identification test: null $H_0$ all instruments are valid probability | 0.83 | 0.036** | 0.28 | 0.71 | 0.28 | 0.061 | 0.53 | 0.23 | 0.13 | 0.75 |
| Cragg-Donald Minimum Eigenvalue | 0.687 | 0.687 | 0.687 | 0.687 | 0.687 | 0.687 | 0.687 | 0.687 | 0.687 | 0.687 |
| TSLS bias v OLS | <5% | <5% | <5% | <5% | <5% | <5% | <5% | <5% | <5% | <5% |

Panel B

|  | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Hausman test: null $H_0$ OLS estimates consistent probability | 0.37 | 0.57 | 0.032** | 0.002*** | 0.12 | 0.39 | 0.80 | 0.54 | 0.33 | 0.04** | 0.30 |
| Sargan over-identification test: null $H_0$ all instruments are valid probability | 0.004*** | 0.28 | 0.42 | 0.82 | 0.99 | 0.49 | 0.51 | 0.92 | 0.77 | 0.36 | 0.58 |
| Cragg-Donald Minimum Eigenvalue | 0.687 | 0.687 | 0.687 | 0.687 | 0.687 | 0.687 | 0.687 | 0.687 | 0.687 | 0.687 | 0.687 |
| TSLS bias v OLS | <5% | <5% | <5% | <5% | <5% | <5% | <5% | <5% | <5% | <5% | <5% |

Note: ***significant at 1% , **significant at 5%

**Table 16: Non-parametric sign tests of the differences between estimates of factor loadings by OLS and TSLS using instruments**

| | SIGN Test | | |
|---|---|---|---|
| | Number of Differences | Number of cases $B_i > IB_i$ | Under Null Hypothesis of equal probability |
| Beta $R_M$ OLS versus TSLS with instrument | n=19 | B1>IB1: w =16 (84.21%) | Prob(W => 16) = 0.0022125** |
| Beta $SMB$ OLS versus TSLS with instrument | n=20 | B2 > IB2: w = 5 (25.00%) | Prob(W <= 5) = 0.0206948* |
| Beta $HML$ OLS versus TSLS with instrument | n=21 | B3 > Ib3: w = 12 (57.14%) | Prob(W <= 12) = 0.808345 |

Note: ** indicates significance at the 1% level, * indicates significance at the 5% level.

We have now established that potential problems associated with endogeneity are a significant issue when estimating betas, in preparation for subsequent cross-sectional asset pricing tests. We have also demonstrated that the relationship between the factors is not necessarily linear. This is not referred to in typical empirical finance studies as linearity is assumed. We will, therefore, leave this issue aside. In the next section, we re-visit the issues raised by Petersen (2009) concerning problems likely to be encountered in the estimation of standard errors in finance panel data sets. These are a feature of typical empirical asset pricing tests.

## 4. Clustering in Finance Panel Data Sets

Petersen (2009) explored the difficulties associated with two general forms of dependence that are commonly encountered in empirical finance applications: the first centres on the fact that the residuals of a given firm may be correlated across years for a given firm (time series dependence). He termed this an 'unobserved firm effect'. On the other hand, the residuals may be correlated across different firms (cross-sectional dependence). He termed this a 'time effect'. Petersen (2009) used simulations to demonstrate that estimates that are not robust to the form of dependence produce biased standard errors and confidence intervals that are often too small, that is, biased downwards.

Petersen (2009) provides the following explanation of the difficulties caused by these issues. He notes that the standard regression for a panel data set can be written as:

$$Y_{it} = X_{it}\beta + \varepsilon_{it}, \tag{11}$$

where equation (9) includes observations on firms $i$ across years $t$. $X$ and $\varepsilon$ are assumed to be independent of each other, and $\varepsilon$ to possess a zero mean and finite variance. The beta coefficient estimated by OLS is:

$$\hat{\beta}_{OLS} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}Y_{it}}{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}^2} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}(X_{it}\beta + \varepsilon_{it})}{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}^2}$$

$$= \beta + \frac{\sum_{i=1}^{N}\sum_{t=1}^{N} X_{it}\varepsilon_{it}}{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}^2}. \tag{12}$$

The asymptotic variance of the estimated coefficient is given by:

$$AVAR[\hat{\beta}_{OLS} - \beta] \stackrel{[}{=} T \ fixed]plimN \to \infty \left[ \frac{1}{N^2} \left( \sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}\varepsilon_{it} \right)^2 \left( \frac{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}^2}{N} \right)^{-2} \right]$$

$$\stackrel{[}{=} T \ fixed]^{plim} N \to \infty \left[ \frac{1}{N^2} \left( \sum\sum_{t=1}^{T} X_{it}^2\varepsilon_{it}^2 \right) \left( \frac{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}^2}{N} \right)^{-2} \right]$$

$$= \frac{1}{N} \left( T\sigma_X^2\sigma_\varepsilon^2 \right) \left( T\sigma_X^2 \right)^{-2}$$

$$= \frac{\sigma_\varepsilon^2}{\sigma_X^2 NT.} \tag{13}$$

The above expression is the OLS formula which is correct when the errors are i.i.d..

Petersen (2009) then assumes that the errors are no longer independent. First, he assumes that the data have an unobserved firm effect that is fixed. This suggests that the residuals contain a firm-specific component $\gamma_i$, and an idiosyncratic component that is unique to each observation, $\eta_{it}$. It follows that the residuals can be specified as:

$$\varepsilon_{it} = \gamma_i + \eta_{it}. \tag{14}$$

Petersen (2009) also assumes that the independent variable $X$ has a firm-specific component:

$$X_{it} = \mu_i + \nu_{it}. \tag{15}$$

The components of $X$ ($\mu$ and $\nu$) and $\varepsilon$ ($\gamma$ and $\eta$) have zero mean, finite variance, and are independent of one another. This ensures that the estimated coefficients are consistent. The independent variable and the errors are correlated

across obsevations of the same firm, but are independent across firms. This can be shown as:

$$corr(X_{it}, X_{js}) = 1 \ for \ i = j \ and \ t = s$$

$$= \rho_X = \sigma_\mu^2/\sigma_X^2 \ \ for i = j \ and \ all \ t \neq s$$

$$= 0 \ for \ all \ i \neq j,$$

$$corr(\varepsilon_{it}, \varepsilon_{js}) = 1 \ for \ i = j \ and \ t = s$$

$$= \rho_\varepsilon = \sigma_\gamma^2/\sigma_\varepsilon^2 \ for \ i = j \ and \ all \ t \neq s$$

$$= 0 \ for \ for \ all \ i \neq j. \tag{16}$$

It follows that the square of the summed errors is not equal to the sum of the squared errors. The same observation can be made about the independent variable. This means that the covariances between the errors must be included. The asymptotic variance of the OLS coefficient estimate can then be written as:

$$AVar[\hat{\beta}_{OLS} - \beta] = \overset{[}{T} \ fixed] plim N \to \infty \left[ \frac{1}{N^2} \left( \sum_{i=1}^{N} \sum_{t=1}^{T} X_{it}\varepsilon_{it} \right)^2 \left( \frac{\sum_{i=1}^{N} \sum_{t=1}^{T} X_{it}^2}{N} \right)^{-2} \right]$$

$$= \overset{[}{T} \ fixed] plim N \to \infty \left[ \frac{1}{N^2} \sum_{i=1}^{N} \left( \sum_{t=1}^{T} X_{it}\varepsilon_{it} \right) \left( \frac{\sum_{i=1}^{N} \sum_{t=1}^{T} X_{it}^2}{N} \right)^{-2} \right]$$

$$= \overset{[}{T} \ fixed] plim N \to \infty \left[ \frac{1}{N^2} \sum_{i=1}^{N} \left( \sum_{i=1}^{T} X_{it}^2\varepsilon_{it}^2 \right) + 2 \sum_{t=1}^{T-1} \sum_{s=t+1}^{T} X_{it}X_{is}\varepsilon_{it}\varepsilon_{is} \right]$$

$$\tag{17}$$

$$\left( \frac{\sum_{i=1}^{N} \sum_{t=1}^{T} X_{it}^2}{N} \right)^{-2}$$

$$= \frac{1}{N}(T\sigma_X^2\sigma_\varepsilon^2 + T(T-1)\rho_X\sigma_X^2\rho_\varepsilon\sigma_\varepsilon^2)(T\sigma_X^2)^{-2}$$

$$= \frac{\sigma_\varepsilon^2}{\sigma_X^2 NT}(1 + (T-1)\rho_X\rho_\varepsilon.$$

Petersen (2009) explores the manner in which empirical finance researchers have estimated standard errors when using finance panel data sets. He observes that the methods adopted vary considerably, and that their relative accuracy depends on the structure of the data. He suggests that estimates that are robust to the form of dependence in the data produce unbiased standard errors and correct confidence intervals; while estimates that are not robust to the form of dependence in the data produce biased standard errors and confidence intervals that are often too small.[1]

We explore this issue in relation to the two sets of estimates of the factor loadings undertaken in the context of the 3-factor Fama-French model: the estimates which use time series regressions based on OLS, versus those which applied two stage least squares (TSLS) and instrumental variables to adjust for endogeneity. We undertake a limited example of the cross-sectional panel regression analysis typical of asset pricing tests using the companies downloaded from 'tiingo'. We use a total of 20 companies because one company had a data set which ended in 2014, as opposed to continuing to the end of 2017. The cross-sectional monthly returns sample is from February 2011 to the end of December 2017.

The regressions feature a basic asset pricing test in which the dependent variable is the actual return on the sample companies. The predicted return is constructed by applying the estimated company market beta to the actual return on the market in month $t$ to produce a series of predicted returns. We decided to switch to a one-factor model, rather than a 3-factor model to concentrate on the impact on the beta estimates, given that the instrument used in the TSLS time series regressions in the first stage of the analysis was related to the market factor.

The results for the stage one time series regressions of beta for the excess return on the market factor estimated by OLS, with the second stage asset pricing tests in a panel context using OLS, Robust OLS, clustered by date, by firm, and by both date and firm, are shown in Table 17.

Table 18 provides estimates in which the first stage estimates of the 3-factor loadings were by two stage least squares (TSLS), plus an instrument. We repeat that, in the cross-sectional tests reported in Tables 17 and 18, we have concentrated on a one-factor model using beta on the excess market return, as this estimate has been one of the main focusses of the adoption of the instrument.

The key issue is the variation in the estimated standard errors. Fama and French (2018) suggest ranking models by the intercept estimate, but Tables 17 and 18 show that the standard errors of the intercept are likely to vary, according to whether we use vanilla OLS, Robust OLS, or tests which allow for potential clustering of errors in the panel regressions used in the asset pricing tests.

---

[1]We are grateful to Mitchell Petersen and Robert McDonald for supplying copies of the R code to replicate Petersen (2009).

**Table 17: Vanilla, Robust and Clustered Standard Errors for OLS using Returns adjusted by Standard Betas**

| Variables | OLS | Robust OLS | Cluster: date | Cluster: firm | Cluster:both |
|---|---|---|---|---|---|
| X | 0.450603 | 0.4216 | 0.4506026 | 0.450602552 | 0.45060261 |
| SE | 0.120462 | 0.0458 | 0.1060870 | 0.095742356 | 0.1170480 |
| t Statistic | 3.741*** | 9.2124*** | 4.2475*** | 4.70641*** | 3.7687*** |
| Constant | -0.007239 | -0.0002 | -0.0072391 | -0.00723912 | -0.0072391 |
| SE | 0.005895 | 0.0022 | 0.0064021 | 0.006699706 | 0.0075205 |
| t Statistic | -1.228 | -0.1076 | -0.9996 | -0.00723912 | -0.8815 |
| Adj. RSquare | 0.007506 | n.a. | 0.007506 | 0.007506 | 0.007506 |
| F statistic | 13.99*** | 83.297*** | 13.99** | 13.99** | 13.99** |
| Res SE | 0.2317 | 0.07618 | 0.2317 | 0.2317 | 0.2317 |
| Observations | 1721 | 1721 | 1721 | 1721 | 1721 |

Table 18 repeats the analysis but, in this case, uses estimates of the beta on the market factors which have used two stage least squares plus two instruments to correct for endogeneity.

**Table 18: Vanilla, Robust and Clustered Standard Errors for OLS using Returns adjusted by Betas estimated by TSLS**

| Variables | OLS | Robust OLS | Cluster: date | Cluster: firm | Cluster:both |
|---|---|---|---|---|---|
| X | 0.6651151 | 0.7038 | 0.66511515 | 0.665115150 | 0.66511515 |
| SE | 0.1475692 | 0.0557 | 0.14894697 | 0.114488052 | 0.15380479 |
| t Statistic | 4.507*** | 12.6381*** | 4.4654*** | 5.80947*** | 4.3244*** |
| Constant | -0.0004105 | 0.0041 | -0.00041052 | -0.000410517 | -0.00041052 |
| SE | 0.0055778 | 0.0021 | 0.00742999 | 0.006121148 | 0.00781964 |
| t Statistic | -0.074 | 1.9344 | -0.0553 | -0.06707 | -0.0525 |
| Adj. RSquare | 0.01112 | n.a. | 0.01112 | 0.01112 | 0.01112 |
| F statistic | 20.31*** | 152.48*** | 20.31*** | 20.31*** | 20.31*** |
| Res SE | 0.2313 | 0.07516 | 0.2313 | 0.2313 | 0.2313 |
| Observations | 1721 | 1721 | 1721 | 1721 | 1721 |

It is apparent in Table 18 that the standard error (SE) of the coefficient of the market factor, $X$, is different from those in Table 17. Table 18 reflects the outcome of estimating beta on the market factor, using TSLS and two instruments, in this case, the OECD monthly "Business Tendency Surveys for Manufacturing: Confidence Indicators: Composite Indicators: European Commission and National Indicators for the United States" plus the return on the VXO, and one lag in both cases. The rest of the results reflect Petersen's original (2009) finding that clustering by date, firm, or both firm and date, will impact on the estimated standard errors in the typical panel data tests of asset pricing models.

### 5. Conclusion

In this paper we have used data that are acessible from French's website to explore the relationship between the data for three monthly market factors relating to US markets, representing the excess return on the market portfolio $R_M - R_F$, $SMB$, and $HML$. We first downloaded the entire monthly and weekly series of the three market factors, which commenced in July 1926 and terminated in July 1918, and estimated rolling bivariate regressions between the three series to explore their relationship through time. The rolling regressions revealed that there are prolonged periods during which the factors are related, and also intervals when they are not. Their relationship is not constant and changes sign in some periods. These results suggest that endogeneity between the factors needs to be considered in certain sub-periods drawn from this 92-year sample. This was confirmed by non-parametric tests of the independence of the series against the predictions obtained from pairwise OLS regressions.

We then used monthly data from January 2000 to December 2010 to further examine these relationships. An exploration of the relationships between these three factors in this sub-period, using OLS, revealed a significant relationship between $R_M - R_F$ and $SMB$. Ramsey's RESET test also revealed a non-linear relationship between $HML$ and $R_M - R_F$. This was further explored via the application of Naradaya (1964) and Watson (1964) kernel regressions, which suggested the existence of non-linearities.

Given that asset pricing tests assume linearity between factors and return series, we set aside the issue of non-linearity and concentrated on the issue of endogeneity, which empirical evidence suggests as being a complication in linear time series estimates of factor loadings in a multiple regression context. We used "Business Tendency Surveys for Manufacturing: Confidence Indicators: Composite Indicators: European Commission and National Indicators for the United States", (BSCICP02USM460S), a monthly OECD indicator series as an instrument in the estimation of factor loadings using two stage least squares (TSLS), plus the return on VXO, and one lag of each of these variables.

Non-parametric sign tests on the beta estimates for the loadings on $R_M - R_F$ and $SMB$ suggested that there are significant difference in the loadings on these factors estimated by OLS, as opposed to TSLS using instrumental variables. Given this finding, we then used a small sample of company returns to undertake cross-sectional tests of sensitivity to the market factor $R_M - R_F$, allowing for clustering of standard errors in this panel of 20 firms, as suggested by Petersen (2009). The results suggested that the estimated standard errors in the panel tests are different when the beta estimates were estimated by TSLS, which adjusted for endogeneity, than by OLS. They also varied if clustering was present by date, or firm, or both, as originally suggested by Petersen (2009).

These 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 problematic in that the standard errors are sensitive to the correct model specification, in both the initial estimation of the factor loadings, and in the subsequent panel data tests, in which error clustering may be a

serious issue.

## References

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

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

[3] Cragg, J.G., and S.G. Donald (1993) Testing identifiability and specification in instrumental variable models, *Econometric Theory*, 9(2) 222-240.

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

[5] 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.

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

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

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

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

[10] Hayfield, T. and J.S. Racine (2008) Nonparametric econometrics: the np package, *Journal of Statistical Software,* 27(5). URL http://www.jstatsoft.org/v27/i05/.

[11] Maasoumi, E. and J.S. Racine (2002) Entropy and predictability of stock market returns, *Journal of Econometrics*, 107(2) 291–312.

[12] Nadaraya, E.A. (1964) On estimating regression, *Theory of Probability and its Applications,* 9(1), 141–2.

[13] 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.

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

[15] Sargan, J.D. (1958) The estimation of economic relationships using instrumental variables, *Econometrica*, 26(3) 393–415.

[16] Sargan, J.D. (1975) Testing for misspecification after estimating using instrumental variables, Mimeo, London School of Economics.

[17] Watson, G.S. (1964) Smooth regression analysis, *Sankhyā: The Indian Journal of Statistics, Series A*, 26(4), 359–372.

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