

The Fama and French model in financial crises: Evidence from Turkey

Hoa Duong⁺

Abstract

The paper recognises the association between the 2008 financial crisis and performance of asset pricing models. The testing of their link has been hindered by a number of reasons, namely the insufficiency of post-shock observations, the use of a single break date and the controversial choice of the crisis time. The paper proposes an approach to avoid these without the problems specific to the test. Cointegration tests are performed to find Turkey as a candidate market to act as a cross-reference to the U.S data as well as one to enable us to perform unbiased examination on. The result reveals that the financial shock can make a substantial impact on a model.

Keywords: asset pricing model, financial crisis, independence tests

JEL classification: G11, G12, G14

⁺ Department of Finance and Management Studies (DeFiMS), SOAS, University of London, Thornhaugh Street, Russell Square, London, WC1H 0XG, the United Kingdom. Tel.: +44 796 829 7662; E-mail address: Hoa_Duong@soas.ac.uk.

I thank Pasquale Scaramozzino, Hong Bo, Christine Oughton, Ciaran Driver, seminar participants at DeFiMS and Sevkett Pamuk- Centre of Contemporary Turkish Studies, London School of Economics and Political Science (LSE), and Hoang Duong for helpful discussions and comments. Financial support from UK International Students Trust and VIED is gratefully acknowledged. All remaining errors are my own.

1. Introduction

Value stocks are documented to earn higher average returns than growth stocks¹. This return differential is called the value premium. In a series of papers in 1992, 1993, 1996, 2006, Fama and French argue that any risk missed by the Capital Asset Pricing Model (CAPM) is the reason for this anomaly. They show that, in a risk-based multifactor model, firm fundamental variables, such as size and book-to-market equity, can proxy for the risk patterns in value premium. Even though the explanatory success of the two factors is robust in many markets, some other research found the argument unsatisfactory². There exist other theories behind the success of the size and book-to-market factors. Sample-specific (Lo and MacKinlay 1990, and Kothari, Shanken and Sloan 1995), overreaction (Lakonishok, Shleifer and Vishny 1994), risks (Fama and French 1993), and firm characteristics (Daniel and Titman 1997) are the common explanations. With the exception of the risk-based reasoning, other theories do not seem to surpass criticism³.

This paper argues that similar to other types of risk, negative economic shocks do play a crucial part in explaining stock returns. Indeed, there are grounds to believe that the performance of asset pricing model is affected by shocks, events and especially financial crises. As pointed out by Campbell (1996), one of the main criteria for successful state variables in a model is forecasting the market return. To demonstrate its predictive ability, a strong model is expected to capture and go through intervention effects, such as the most

¹Value stocks refer to firms with high ratios of book-to-market value (B/M), earnings-to-price (E/P), cash flow-to-price, or dividend yield (DY). Stocks, of which the above fundamental values are small, are characteristically called as growth stocks.

²To capture the risk factor, Bhandari (1988) use the beta, size and leverage factors, Carhart (1997) used a four factor model while Avramov and Chordia (2006) favour liquidity risk.

³Chan, Jegadeesh and Lakonishok (1995) and Barber and Lyon (1997) point out that sample selection procedure does not account for the patterns of size and B/M in stock returns. Against the overreaction explanation are Rozeff and Zaman (1998) and Doukas, Kim and Pantzalis (2002) to name but a few, and Clare and Thomas (1995) who further argue that the psychological biases are in fact size effects. Daniel and Titman (1997) believes that distress or growth characteristics have been mistaken with risks, since they both relate to the covariance of stock returns and are difficult to be differentiated. The results are however questioned due to their rather short sample period (Davis, Fama and French 2000).

severe 2008 credit crunch that has diverse effects on the whole market's overall functionality. In support of this view, Buelens (2012) examined three groups of inflation models, and noticed that the financial crisis introduces substantial forecasting errors. Moreover, market return often signals the market conditions and should be one of the first victims of financial crises. With regards to the value premium strategies, Fama and French (1993 and 1996) document this possibility in the sense that value stocks do very badly due to typically being in distress and that their risk aspect rises notably (when a crisis comes along, firms are more likely to be in distress). As a consequence, Cooper (2006) uses an alternative approach (through systematic risk) to support that distressed firms are riskier particularly in "bad time". Supporting this papers' view, Lettau and Ludvigson (2001) use a different methodology to document the importance of good and bad states, and the level of the price of risk involved. Furthermore, Zhang (2005) simply finds value stocks are riskier than growth stocks in bad times, linked through consumption and production toward the risk. For a risk-based model, it is therefore relevant to examine the importance of states and distress level elaborated by the financial crisis.

An existing problem is how to know if a model starts failing due to a shock. One method is using event studies techniques summarised in MacKinlay (1997), which observe the effects before and after the shock, and note down the difference. The problem with this method is that with a financial crisis, it is not a short-term event and its starting date is also controversial. Thus, one cannot choose a precise break date in order to apply event studies techniques or to control for its effects with dummies. Additionally, the 2008 financial crisis, by nature, is recent and hence does not have a sufficient number of post-crisis observations to avoid misleading the results.

Moreover, studies in extreme states have a tendency to distort the value premium and the associated investment strategies. Concern is pointed out by Lakonishok, Shleifer and Vishny

(1994) who defend traditional measures of standard deviation and beta, saying dwelling on the examination of extremely bad times provides refuge for those looking for proofs that high return strategies are riskier. This prompts us to find ways to test the whole time frame without having to concern about the distorting effect of the states as well as overcome the problems with previous methods that hinder the hypothesis testing.

With that in mind, the paper experiments this method for the U.S. market in Section 5.3, Table 3 and the outcomes do give the first impression of the shock affecting assessing asset pricing models. Next, the paper starts applying its own approach which overcomes the common overlooks of traditional methods in testing this particular hypothesis. The results are found to support a main argument that the Fama and French model survives against criticism when shocks are taken into account.

More specifically, this paper proposes an alternative method which is further testing the model on a market less affected by the shock. This can isolate the shock-affecting possibility enables this study to compare the effectiveness of the model on two (or more) testing samples. The study pays attention to the U.S. market cross-section data as the main testing sample for models' efficiency. In addition, it looks at the cointegration tests between four global leading indices and the Istanbul Stock Exchange (ISE) to look for proofs of their linkage and dynamic interaction. Transiting from a closed economy to a more liberal system in 1980s, Turkey has its very own economic and legal characteristics that result a neutral response of Turkish market to the financial crisis. For examples, their relatively limited borrowing, low currency risk, a strong banking sector, and tight fiscal regulations that saved it from both domestic and external shocks. Indeed, the degree of interdependence and international co-movements are found relatively low using various state-of-the-art examinations.

One advantage of this method is that the tests have the same time period of long enough observations. This is significant because, taking the 2008 credit crunch as an example, there are insufficient observations for the post-event testing to carry unbiased examination. The second key advantage is that unlike firms' announcements, crisis is not a point break, but rather a lasting period without an agreed single date of representation or commencement. Using the proposed method, future studies can avoid splitting the sample and arrive to a more accurate approach than the splitting sample method. The results presented later in section 5.3 show that even splitting the sample does not provide a clear conclusion.

The Fama and French (1993) three-factor model and some of the most discussed risk-based models are selected as sample-cases to test the paper's hypotheses. Through mathematical reasoning, Petkova (2006) offers a model with different factor loadings and documents that the Fama and French (1993) three-factor model – the so-called FF model – fails to explain time-varying patterns in returns. Furthermore, the FF success is believed to be merely because their factors are parts of her innovations in predictive variables which are necessary according to Chen (2003). Hahn and Lee (2003) presented a model using once again a different set of factors to replace completely the previous works. However, Petkova points out that the Hahn and Lee's study lacks of sufficient proofs. The driving motivation behind their works is to come up with an arguably better model for a market(s) in mind. What if the FF model still works and widely used across markets. In fact, it is noticed that the two studies both covered periods including a great market downturn while Fama and French (1993) did not, which this study hypothesises to be the explanation. The question this study approaches to answer is whether the previous model itself has any problem. It uses a method that can tell if it decays or externally faulted by a specific shock.

In terms of methodology, the approach proposed in this paper allows an unbiased test for model strength to be done. The chosen models are observed on the US market where the

impact from the crisis was high. We scan through many markets, including the emerging market groups, using cointegration tests in order to find Turkey as the best candidate to cross-reference its results with the U.S. data results. The data are ideally suited to the comparison of the effects of a financial shock because the Turkish stock returns are little affected by it. Furthermore, testing on the Turkish market provides some valuable insights as to why the FF model behaves like it does on the U.S. market, and into the nature of the value premium without the obstacle mentioned above by Lakonishok et al. (1994). This study examines the whole time frame of markets without extremely bad states, due to one of the tested markets has such states' effect being minimal in the 2007/2008 credit crunch, and will act as a reference for the others. So, the test of the models' performance will be more accurate and comprehensive. Secondly, with this method the study would be able to indicate models through which anomalies were well explained, but fail doing so as a result of a crisis.

Furthermore, it is imperative to also take into consideration a number of other studies, as they claim to have formed a better explanatory variable set than the three factors and that the FF factors lost their powers in explaining returns. In a pioneer study Campbell (1996) points out that a desirable cross-sectional asset pricing model is one whose variables must proxy changes in investment opportunities. So, in this paper, four more related factors are included for experimental purpose: the short-term Treasury bill rate, aggregate dividend yield, term spread, and default spread. In addition, Avramov and Chordia (2006) suggest the use of momentum factor to explain value premium. Moreover, themselves and Pastor and Stambaugh (2003) document liquidity has a strong effect on stock returns. They observe a lower return as a result of higher turnover. Thereby, this paper shall follow this anomaly and also investigate past returns and liquidity.

Final results show that the FF is a long-standing model; however it is found that when taking into account the 2008 financial crisis, there is evidence to support its weakening,

mentioned by Petkova, and Hahn and Lee. Comparing with other models on a different market of small crisis effects has further assured on the primary cause of this finding. Efforts to augment the model using a number of factors do not seem to be significant beneficial. The improvement, if any, is immaterial.

The rest of the paper is organised as follows: Section 2 documents the selected models. Section 3 discusses the data selection criteria. Section 4 elaborates the Cointegration tests and explains a need to do them, which are necessary in order for section 5 to proceed with the tests of models using the paper's method. Section 6 summarises and concludes.

2. Selected models

This study collectively experiments 13 models, listed in Table 4. Among these models, the paper bases its concerns around the Fama and French (1993) model, while the Petkova's (2006) and the Hahn and Lee's (2006) models are also considered for the counterargument. The latter two are chosen as examples, because they are two which are able to clearly document their better models, and the failure of the FF. Additionally, out of many risk factors discussed in literature, the paper selects and later argues for the most frequently used, which results in two extra variables, liquidity and prior returns. These models do not aim to be exhaustive but to provide comprehensive and ideal case studies of how commonly-used measures and method are insufficient.

3. Data

For cointegration tests, the four leading stock indices are DAX 30, Nikkei 225, FTSE 100 and S&P 500, testing in relation to ISE 100. Monthly closing data for all five indices are obtained from DataStream Thomson Reuters database and over the period beginning in January 1988 and ending in December 2010. When the stock exchanges were closed due to

holidays or unexpected events, the index level was assumed to stay the same as that for the previous trading day.

For subsequent tests, data cover the period from July 1994 when sufficient data are available to December 2010, 198 months, for non-financial firms. Bond and bill rates in treasury auctions are obtained from the Central Bank's record. Variables are monthly data for most parts, however, default spread is based on annual data since the accounting reports are available at the yearly frequency. The sample excludes financial firms since the meaning of these firms' variables differs from that of non-financial firms. For example, a high leverage in financial firms is common, while it normally refers to a distress situation in non-financial firms. Checking a sub-sample shows that the exclusion of financial firms does not materially distort the results⁴.

In this study's estimation, survivorship bias is eliminated, since it also includes delisted firms in year(s) they were listed⁵. DataStream reports full data of only live trading firms, although there are firms that used to be listed and satisfied selection criteria. This study takes into account distress risk, bankruptcy risk and hence the need to include dead firms in the past. This study cross-references different sources to confirm the delisted firms' identity and data and forms its own portfolios based on collected data. Further advantages of this method are discussed in Kothari, Shanken and Sloan (1995).

Fama and French provide estimations for their three factors, namely the HML, SMB and the Market, on Professor Kenneth French's website for many countries⁶, the figures for the U.S. market are obtained in such a way while for Turkey the authors form these factors using the Fama and French (1993) approach with breakpoints set at the 40th and the 60th percentiles.

⁴For the sake of brevity, the paper do not report these results, they are however available upon request.

⁵ Delisted firms are defined as firms which went bankrupt, engage in merger and acquisitions or are disqualified. A check if exclusion of dead firms changes the performance of each pair of portfolios was also performed and found it statistically and economically impacts the market anomalies test.

⁶ We thank Kenneth French for making the portfolio and factor data available on his Website: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Accordingly, six size/BM portfolios were constructed as the intersection between three B/M and two size portfolios.

In order to be included in year t of the sample, a stock must have sufficient data on return index from July of year $t-1$ to June of year t , and on market value of equity as at 30th, June of year $t-1$. To mirror the real investment environment, negative B/M, negative E/P and zero DY stocks, which are rare during the testing period, are excluded from B/M, E/P and DY portfolios, respectively.

4. The independence proofs

It is considerably difficult to find out whether a time-series is affected by a specific shock and to what extent. The analysis resorted to use the cointegration tests to discover the linkage between highly affected markets and Turkey. There are a few reasons why this had to be done:

Firstly, the financial crisis originates from the U.S. market, and markets known to be heavily linked to this market, such as the UK, are strongly affected by it. Secondly, if the Turkey is strongly affected by the crisis, it should have had some level of dependence with the U.S. market directly, or with some other markets experiencing the crisis's effect, especially after the crisis⁷. Or where else could it have been affected from.

Even in this case a few short-comings and argument will look at the issue and the alternative methods at the end of this section.

⁷Arshanapalli and Doukas (1993) document that the co-movements between the US, UK, Germany, French, and Wu and Su (1998) show the co-movements between the US, UK, Japan and Hong Kong are stronger after the October 1987 market crash

4.1. Constructing the test

Although the concepts of Engle-Granger and Johansen cointegration tests can be found in many sources, it is worth to briefly summarise their main arguments and show how to augment the traditional tests to account for the 2008 financial shock.

In order to determine the cointegration between two market indices, it is necessary to first carry out the unit root test of whether each series is integrated of order one, I(1).

4.1.1. Unit Root Tests

Since the pioneer study on unit root by Dickey-Fuller (1981), there has been a mounting concern about the impact of structural breaks on the validity of the test. A brief review of the main studies on this subject will follow.

4.1.1.1. The ADF and PP

The conventional Augmented Dickey-Fuller (1981) and Phillips and Perron (1988) regressions are used respectively when one thinks of testing for the stationarity or the presence of a unit in a series.

The ADF test shall examine whether y_t is I(0) can be written as:

$$\Delta y_t = a_0 + a_1 y_{t-1} + \sum_{i=1}^p (\alpha_i \Delta y_{t-i}) + u_t \quad (1)$$

where p is large enough to make u_t white noise. The series can also subjects to PP test:

$$y_t = b_0 + b_1 y_{t-1} + v_t \quad (2)$$

where v_t is serially correlated.

4.1.1.2. Allowing multiple breaks

Although our interest remains to uncover the impact of crises on stock indices, it is important to rule out the likelihood that any found unit-root may have been the result of

structural change in data (Perron 1989). The use of dummy variables in the specification allows this. Further literature by Zivot and Andrews (1992) and Perron (1997) have included an exogenous breakpoint in the equation. By allowing one breakpoint they have found more evidence against the unit root hypothesis than the Dickey and Fuller (1981) method. To avoid favouritism toward the hypothesis, the paper aims to incorporate multiple unknown breaks in the testing process, as mentioned by Lumsdaine and Papell (1997). This paper employs this idea but in the more recently developed procedure by Lee and Strazicich (2003), whose advantages allow for breaks under not only the null, but also the alternative hypothesis. The test is arbitrary enough to allow for more than one crisis, which may not necessarily be the same across markets, to ensure unbiased evidence that all series are I(1). In addition, as break points are documented to be less inference sensitive than the assumptions about the number of breaks, this paper shall fix neither the number of breaks nor their dates using the Minimum Lagrange multiplier (LM) test proposed by Lee and Strazicich (2003)⁸ on series y over the period of T .

$$\Delta y_t = \mu + \beta t + \theta DU_{1t} + \gamma DT_{1t} + \omega DU_{2t} + \psi DT_{2t} + \alpha y_{t-1} + \sum_{s=1}^p \Delta y_{t-s} + \varepsilon_t \quad (3)$$

$t = 1, \dots, T$

where, DU_{it} are indicator dummy variables for a mean shift occurring at times T_i , that equals to 1 for $t > T_i$, DT_{it} are the corresponding trend shift variables, and equals to $(t - T_i)$ when $t > T_i$.

4.1.2. Cointegration Tests

Once the time-series are I(1), it is also important to ensure cointegration tests are not sensitive to the presence of structural break(s) occurring during the testing period. Thus,

⁸ We thank Junsoo Lee for generously providing the code for the test in RATS

beside the conventional cointegration tests such as Engle-Granger and Johansen tests, an augmented Engle-Granger test which allows for one break will be also carried out.

4.1.2.1. Engle and Granger

Now the test of cointegration can be taken once each series is found to have one unit root, when test statistics are derived from residuals of the following regression:

$$Y_t = \alpha_0 + \alpha_1 X_t + \alpha_2 t + e_t \quad (4)$$

where Y_t and X_t are the regressand and regressor, respectively and t is a trend.

If the series are cointegrated, the ADF test shall examine whether e_t is $I(0)$:

$$\Delta e_t = \alpha_0 e_{t-1} + \sum_{q=1}^p \phi_q \Delta e_{t-q} + v_t, \quad (5)$$

where q is large enough to make v_t white noise. The estimated residuals are also subject to PP test:

$$e_t = \beta_0 + \beta_1 e_{t-1} + \gamma_t \quad (6)$$

where γ_t is serially correlated.

Two series are said to be cointegrated when there exists a stationary linear combination of the two. In terms of inter-market efficiency, cointegration implies that the markets are linked even if they are non-stationary. Denote $z_t = y_t - \theta x_t$, where $x_t, y_t \sim I(1)$ and z_t is stationary and invertible ARMA then the two series x_t and y_t are cointegrated if and only if $E(z_t) = \delta_z$ and $\text{Var}(z_t) = \sigma^2 < \infty$.

4.1.2.2. Johansen

Johansen's maximum likelihood method suggest testing cointegration using dynamic VAR(k) specification of vector X_t , of the size $n \times 1$, consisting of $I(1)$ variables, to construct common stochastic trends:

$$X_t = \mu + A_1 X_{t-1} + \dots + A_k X_{t-k} + \varepsilon_t \quad (7)$$

where ε_t is assumed to be an *i.i.d.* Gaussian process

Next, denote $\Delta \equiv I - L$ where L is the lag operator.

The model above is rewritten as follows

$$\Delta X_t = \mu + \Gamma_1 \Delta X_{t-1} + \dots + \Gamma_{k-1} \Delta X_{t-k+1} + \Pi \Delta X_{t-k} + \varepsilon_t \quad (8)$$

where

$$\Gamma_i = -(I - A_1 - \dots - A_i), \quad i = 1, \dots, k-1$$

$$\Pi = -(I - A_1 - \dots - A_k)$$

The benefit of this is that it allows all long-run information of X_t to be recapitulated by ‘long-run impact matrix’, Π , whose rank shall determine the cointegrating vectors number.

If the coefficient matrix, Π , has rank of $r < n$, then there exist $n \times r$ matrices φ and ψ each with rank such that $\Pi = \varphi \psi'$ and $\psi' X_t$ is $I(0)$.

The analysis of these matrices can be referred to in Johansen (1991, 1995) for details.

In summary, the strength of Johansen approach lies in his assumptions which allow the Maximum Likelihood estimation to be incorporated with the cointegration issue and its testing framework.

4.1.2.3. *The cointegration with a structural break*

One approach is the LM test proposed by Westerlund (2006 p.101), which allows “for unknown number of breaks to be located at different dates and for different individual, endogenous regressors as well as serial correlation”. The test is however designed for a panel cointegration check between a number of series across markets, which is beyond the interest of this paper and will be useful for other research.

The paper contributes a method to test whether, even with a break, the cointegration might still exist. From the cointegrating regression (4), with an exogenous structural break in the level occurs at time $1 < T_B < T$ is added, the equation is followed.

$$Y_t = \alpha_0 + \alpha_1 X_t + \alpha_2 D(T_B)_t + \alpha_3 D(T_B)_t X_t + \alpha_4 t + e_t \quad (9)$$

where $D(T_B)_t = 1$ if $T_B + 1 \leq t \leq T$ and equals to 0 otherwise.

A rejection of the null hypothesis implies that the series are cointegrated with an exogenous change in the level at time T_B .

For the cointegration test with a break at the recent financial crisis, commonly chosen at September 2008, the above test does not include the break in trend function, which appears to be more appropriate for recent effects caused by the current financial crisis.

4.2. Empirical Features of the Co-integration Tests

Table 1 reports the results of unit root and cointegration tests with and without breaks.

In order to test the cointegration between markets, the series have to be determined integrated of order 1. First, the ADF and Phillips-Perron tests were used allowing no breaks and confirm that indeed they are (as reported in panel A). Then, accounting for a number of structural breaks the minimum LM test by Lee and Strazicich (2003) reassured the results at the 5% significant level.

Choosing Turkey as the base index, the null hypothesis of no cointegration between the ISE and each of the rest of the indices cannot be rejected at 5% significant level by Engle – Granger and Phillips-Ouliaris methods. At this stage, the paper continues to double-check the analysis of their interdependence by Johansen’s (1991, 1995) dynamic method with a maximum rank order, r , of one. For all cointegrating vectors, panel B of Table 1 confirms the previous results through the common stochastic trends, and that both Trace statistics (λ_{trace}) and Max-Eigen statistics (λ_{max}) cannot reject the null hypothesis under ranks $r=0$ and $r \leq 1$.

In addition, panel C is constructed with a structural break at the recent financial crisis, T_B . The tests of whether T_B is in fact a structural break in the time series indices show significant coefficients associated with the dummy variable, $D(T_B)$. This implies that the break has an important role in the cointegration tests. Additionally, both τ -statistics and z -statistics are not significant using MacKinnon (1996) critical values confirming the null hypothesis of no cointegration.

4.3. Implication and Short-comings

From an empirical point of view, we can see that the U.S. stock market has no significant linkage to the Turkish market. Over the 2008 crisis while the four major markets are strongly linked, the Turkish is reported to be substantially independent with all of the four.

Of the four major markets, if there is a strong degree of inter-dependence amongst them, an exogenous factor which affects one, would affect all others. The financial crisis has affected them severely. In this case, if the financial crisis had impact on Turkish market, it is likely to be little.

In terms of institutional characteristics, there are a number of economic and legal reasons for the neutral response of Turkish market to the financial crisis. As pointed out by Macovei (2009) and Turhan and Kilinc (2011) to name but a few, Turkey has a relatively limited borrowing, low currency risk, a strong banking sector, and tight fiscal regulations that saved it from both domestic and external shocks. Aside from the country norm of preferring low private and public debts, Turkey has strict regulations when it comes to balancing their foreign exchange position and government fiscal policy. Another reason for their stable performance throughout the 2007/2008 credit crunch is that Turkish banking system was comparatively strong as a result of major structural reforms after its 2001 crisis. With such

significant improvement in regulations and supervision over the banking sector, banks are not allowed to carry “bad” assets while still relatively independent to the Central bank.

Having mentioned above of the crisis affect the market returns of U.S., UK, Germany, and Japan, however there have not been any empirical efforts to document a method of measuring the effects. In order to examine an impact of an event to a time-series, one often requires employing either of the following techniques.

First is the event study, which this study documents earlier to be inappropriate. This technique aims to take in to account the effect of an event, such as firm announcement, which occurs generally within a certain day. Or one could use structural break tests to detect the financial crisis effect. It is a custom to believe that a global shock such as the financial crisis will cause a structural break in a time-series. This method could be put to use, however the problem is that a failure to reject the null of no structural break does not necessarily imply that financial crisis has no effect. Otherwise a proxy to capture its effect would be a conventional and defined method. Had this been possible, its effect could have been well studied simply using either ordinary least squares (OLS) or generalized method of moments. However, to proxy for the effect of the financial crisis is an original idea. We hope ourselves and future research will look into this possibility, but for now this is the limitation of this paper.

Another point worth noticing is the possibility of endogeneity occurring in the tests. Negative shocks can lead to a poor proxy for expected asset returns in testing asset pricing models. The paper, however, bases its analysis on a sample where the effects of shocks appear to be limited. Therefore, the likelihood of endogeneity within a model, if any, would not significantly affect the outcomes. There are also many financial shocks in international markets which could be incorporated to better enhance the results. However, for simplicity,

this paper emphasizes on a methodology and looks at the case of the most recent and severe crisis; it encourages future research to look into larger data set.

5. Asset pricing models

The FF model is tested here using two markets; however, it is necessary to furthermore consider the alternatives. Petkova (2006) and Hahn and Lee (2006) document failures and complete replacements of the model, this paper will supply a few points regarding their approaches. Upon correction of these that had been previously overlooked, this paper re-runs these two models and also experiment some others to see if they really work as well as proposed.

5.1. Composite Variables

Table 2 summarised descriptive statistics of explanatory variables. The High-minus-Low, HML, and Small-minus-Big, SMB, portfolios meant to mimic the risk factor in returns associated with B/M and size, respectively, using the lag of 1-month. The market portfolio, R_m , consists of all stocks in the value, growth, small and big portfolios plus negative B/M firms.

This study differs from previous studies in the way that it does not use correlated regressors within its analysis. Specifically, Petkova (2006) defines default spread, DEF, as differential in returns between Long-term Corporate Baa Bond and Long-term Government Bond; TERM as 10-year government bond minus 1-year government bond, where using the FRED[®] database long-term government used is the 10-year bond. It is, therefore, undoubted that the two variables are correlated.

Similarly, Hahn and Lee (2006) propose the usage of ΔDEF and $\Delta TERM$ as an alternative set of variables to capture the value effect. They define DEF as the yield spread between Baa corporate bond index (Bond Index) and 10-year Treasury constant maturity (10yTbill), and

TERM as the spread between 10yTbill and one-year Treasury bill (1yTbill) rates. They are defined and rearranged as follows:

$$\begin{aligned}\Delta DEF_t &= - [DEF_t - DEF_{t-1}] = - [(Bond\ Index_t - 10yTbill_t) - (Bond\ Index_{t-1} - 10yTbill_{t-1})] \\ &= - (Bond\ Index_t - Bond\ Index_{t-1}) + (10yTbill_t - 10yTbill_{t-1})\end{aligned}$$

$$\begin{aligned}\Delta TERM_t &= TERM_t - TERM_{t-1} = (10yTbill_t - 1yTbill_t) - (10yTbill_{t-1} - 1yTbill_{t-1}) \\ &= (10yTbill_t - 10yTbill_{t-1}) - (1yTbill_t - 1yTbill_{t-1})\end{aligned}$$

The term $(10yTbill_t - 10yTbill_{t-1})$ appears in the construction of both variables. While the dependence between them resulting a multicollinearity problem does not invalidate the model as a whole, a high correlation between regressors, especially in OLS estimation, will call off the predictability power of each correlated individual predictor, and whether one predictor is redundant with respect to others.

Two problems arise as a result. The first is that a high correlation between independent variables can damage the accuracy with which each of the variables' slopes is measured (Pastor and Stambaugh 2003). Also, the second issue is that no conclusion on the significance of each regressor can be made. Even Petkova (2006) herself documents the need of a variable to be significant to be important. Unable to determine the significance implies this would mislead many interpretations, such as each factor's role in the regression. Indeed, to examine the relationship between *SMB* and ΔDEF and *HML* and $\Delta TERM$ in a view to counteract Fama-French factors, Hahn and Lee (2006) document the following 2 regressions:

$$SMB_t = a_1 + b_1 R_{m,t} + c_1 \Delta DEF_t + d_1 \Delta TERM_t + e_{1,t} \quad (10)$$

$$HML_t = a_2 + b_2 R_{m,t} + c_2 \Delta DEF_t + d_2 \Delta TERM_t + e_{2,t} \quad (11)$$

Next, from regression (10) they interpreted the c_1 coefficient as significant while d_1 is not in order to conclude that the ΔDEF regressor can replace *SMB*, but it has been given that its

respective regressors ΔDEF and $\Delta TERM$ are correlated, hence it is not possible to obtain such interpretations. In the regression (11), similar bias conclusion appeared.

This paper's method can largely avoid such problems. The use of the O-score as a good proxy for default risk is also in the context of Dimson, Nagel and Quigley (2003) and Griffin and Lemmon (2002). Default spread, DEF , is the difference in returns between firms with the highest probability of bankruptcy, measured by Ohlson's (1980) O-score⁹, and firms with the lowest O-score. The DEF factors formed in this study using portfolio excess return has the advantage of being able to capture the corresponding risk premia, which are compensations for investment in firms with high default risk. The default spread is the universal proxy for the risk of change in credit market conditions. Increases in DEF signal the market's anticipation for bad credit market state. Followers of risk-based models always aim to capture abnormal returns through the kind of risk that concerns their assets, and hence models capturing more risk or explain the most risk compensation are more successful in bringing explanatory power. Different methods would be needed to go around this problem, if otherwise. Supporting this view, Fama and French (1993) is also able to perform such formation to construct their DEF factor in a portfolio basis but not to the TERM factor. This is natural as the TERM factor captures the change in market interest rates, thus it is unreflective to build a portfolio around it.

DEF, TERM together with prior return and liquidity variables are going to be used to proxy risk premia under changing investment opportunities. To proxy for these characteristics, past return and liquidity level are among the most prominent factors in explaining common variation in stock returns. This opens a possibility that adding them could enable augmented

⁹O-score = $-1.32 - 0.407 \log(\text{total assets}/\text{GNP price-level index})$
 $+ 6.03 (\text{total liabilities}/\text{total assets}) - 1.43 (\text{working capital}/\text{total assets}) + 0.076 (\text{current liabilities}/\text{current assets}) - 1.72(1 \text{ if total liabilities} > \text{total assets, else } 0) - 2.37(\text{net income}/\text{total assets}) - 1.83(\text{funds from operations}/\text{total liabilities}) + 0.285(1 \text{ if net loss for last two years, else } 0) - 0.521(\text{net income}_t - \text{net income}_{t-1})/(\text{net income}_t + \text{net income}_{t-1})$.

models to capture time-varying and cross-sectional patterns in stock price movement that has not been covered by the three factors. The purpose of this experiment is not to test all potential risk-based models, it is rather to robustness check the validity of FF model (model 1 in **bold**, Table 4) and its counterparts.

There are a number of proxies for liquidity factor, the usage is dependent on the goals. Two recent measurements are Pastor and Stambaugh (2003)'s pioneer non-traded liquidity factor measuring the impact of liquidity risk in association with daily price changes; and Avramov and Chordia's (2006) measurement focusing on the stock sensitivity to the liquidity factor. However, Avramov and Chordia document that neither of the two measurements can capture the impact of turnover on expected returns. This study employs monthly turnover ratio to act as an indicator for the purpose of capturing liquidity effect on monthly portfolio returns. Turnover, *TURN*, proxy for liquidity factor, is a mimicking portfolio that is long in low turnover stocks and short in high turnover ratio stocks. Turnover ratio is defined as trading volume divided by number of shares outstanding.

Additionally, one other common discussed factor, leverage, suggested by Bhandari (1988) is *not* included for collinearity reasons, since liabilities and O-score are interdependent (derived from footnote 9) and the meaning behind leverage such as probability of financial distress and default should be captured through O-score. The short-term Treasury bill variable, *STBill*, is a monthly rate of the 3-month T-Bill. DIV is the one-year aggregate dividend yield and term spread, *TERM*, is the difference between monthly long-term government bond rate and the short-term T-Bill rate¹⁰. The momentum factor is Winner-minus-Loser, *WML*, based on the Jegadeesh and Titman (1993) momentum strategy of

¹⁰The classification of long and short maturity of these government securities in each year varies upon the availability of the bonds and bills in the year.

buying high 11-month past returns and selling low 11-month past returns, lagged 1 month. Finally, a full list of explanatory proxies is also found in the first column of Table 4.

5.2. Methodology

First, the ordinary least squares time-series and cross-section analyses in this study assume that returns are generated by a linear regression of a k-factor model.

$$R_{it} = \alpha_i + \gamma_{M,i} (R_{mt} - R_{ft}) + \sum_{k=1}^n \gamma_{ki} U_{kt} + \varepsilon_{it} \quad (12)$$

where R_i is the excess return on portfolio i , R_m is the market portfolio, R_f is the risk-free asset, and U_k is factor loadings of factor k^{th} , and n is the number of factors added in the CAPM.

In Fama-MacBeth estimation, to judge of the goodness of fit of the model over different periods, the paper uses three measurements: the cross-sectional R-square measure employed by Jagannathan and Wang (1996) and Lettau and Ludvigson (2001), the Mean absolute pricing error measure suggested by Phalippou (2007), and a visual assessment by plotting fitted versus realised returns.

The R-square proposed by Jagannathan and Wang (so called *JW-R-square*) is defined as:

$$JW-R-square = (\sigma_r^2 - \sigma_e^2) / \sigma_r^2 \quad (13)$$

where σ_r^2 is the variance of the average returns and σ_e^2 is the variance of the average pricing errors across N portfolios.

Mean absolute pricing error for asset n generated based on pricing errors, $u_{t,n}$, from the Fama-MacBeth's second step is

$$u_n = 1/T \sum_{t=1}^T u_{t,n} \quad (14)$$

As a further visual assessment of the performance of the model over time, Figure 1 plots the fitted expected return of each of the 25 size-B/M sorted portfolios against its realised

average returns. A perfect model is one where all plot lie on the 45 degree line through the origin.

The OLS regressions is, on the other hand, utilizing the R^2 and standard errors adjusted for the degree of freedom, and the standard-pricing error is the square root of time-series average of pricing errors squared, which will avoid cancelling out.

$$u_n = 1/(T-1) \sqrt{\sum_{t=1}^T \hat{u}_{t,n}^2} \quad (15)$$

Although there are many statistics measurements used to assess a model performance, this paper employs these three main statistics but also aware that more test statistics can improve the power of its arguments. This is, perhaps, an interesting point for our future works.

In terms of methodology, this paper proposes a new approach of testing the model on a market less affected to the shock. This will isolate the shock-affecting possibility, leaving us able to compare the effectiveness of the model on two (or more) testing samples. The advantage of this method is that the tests have the same time period of long enough observations for both tests. This is significant because say for the 2008 shock, there is not enough observations following the event up to now to carry unbiased examination. A second advantage is that unlike firms' announcements, crisis is not a point break, but rather a period; thus this method can avoid splitting the sample and arrive to a more accurate approach than the splitting sample. The interpretation of results will further explain this part.

5.3. Empirical Results

This section presents the main findings on performance of the FF model, its counterparts and other augmented models.

5.3.1. The Fama and French (1993) three-factor model

The first part of this section analyses the results of the model in the U.S. market together with some robustness measures, the second is for the Turkish market, then some discussions follow.

For the U.S. sample, this section first assesses the goodness of fit of the FF model broken down to three different periods, each one involving the Fama-MacBeth two-step procedure. Accordingly, the average of estimated coefficients stands for risk premium for each corresponding factor. The Fama-MacBeth cross-sectional approach first estimates a multiple time-series regression and then cross-sectionally relates the average excess returns of all assets to the risk patterns in factor loadings. For example, the pre-crisis period test runs 25 time-series regressions on the three factors from July 1963 to August 2008 (542 months), and then run $T=542$ regressions cross-sectional on the exposures to the risk factors, which are the average of monthly regression coefficients from the previous step. The two-step cross-section regression results are presented in Table 3.

As Petkova (2006) updates results of the FF model up until 2001, this paper extends this examination to 2011 data. The FF model is examined in the U.S. market as a benchmark comparison for other works' coming together, as well as for this paper's own hypothesis of whether the crisis has made an impact on the model.

Let us first looking at the first period in the U.S. market 1963 – 2008. The paper starts its check from 1963 to coincide with the starting date of Fama and French, Petkova, and Hahn and Lee models to better compare the model performance. The reported *JW-R-square* is: 77.03% (Table 3). Next, the post 2008 period spanning up to December 2011 has the corresponding value of only 24.46%. This in itself can provide an initial proposal that the

financial crisis has noticeable impact on this model effectiveness. It could even suggest that the model works worse after the shock event.

Before looking at the whole period from 1963 up and including the financial crisis to December 2011, it is worth noticing the *JW-R-square* is at 71% for the period 1963- 2001, according to Petkova; then it is found at 77.03% up until 2008. However, now the whole period, 1963-2011, has a stumbling 72.27% *R-square*. So, there seems to be more evidence to support the model's failure, especially after crisis. Later sections will further interpret the validity of these.

In the meantime, the section returns to the commonly-used checks in order to re-investigate this result. Even when the paper undertakes these disadvantaged traditional approaches, the results are inconclusive. The goodness of fit of the model is showed through the average pricing error for the 1963-2011 being a minimal 0.082%. The pricing error drops after the crisis happens by 0.002% from 0.084%. In fact, this is an improvement from the 0.10% and 0.15% reported in Phalippou (2007) for periods 1963-2001 and 1980-2001, respectively. So according to this measure, the model gets better fitted over the timeframe. This lays further foundation to investigate the model performance. From Figure 1, the graph of the three periods shows that the post-crisis period the fitted values are far from the true values. The graph the entire period and the pre-September 2008 period are indistinguishable. In this example, it is not possible to draw any conclusions on whether the model has improved or worsen after the crisis, even when a choice of a single break date had to be done.

However, as mentioned earlier, one of two disadvantages of the method involving splitting sample is that estimations do not always provide unbiased comparison between the two samples of significant difference in observation number. More specifically, the post-2008 sample, by nature, is recent and therefore has a small number of observations. Also, the

difference in the sample length between the 1963-2008 and 1963-2011 is insignificant, being 542 month and 582 month respectively.

The second problem is that no one can satisfactorily argue for an exact date of the 2008 crisis to be a breakpoint. In fact, a single date cannot represent a crisis. Hence, testing its effect on a model has been virtually always biased. However, this paper's method is able to test the crisis effect on a certain model without having to choose any break dates. Further investigations will now follow.

In the Turkish market, let us look at the cross-section sample where insignificant 2007/2008 impact was found. The sample is selected from 1994 when the ISE data were sufficiently established. The start year of the sample does not coincide with that of the U.S. sample. However the nature of the market is, at this point in time, mature enough to carry out the test. The model performance effectiveness on the market will not be affected depending on how long the market has been actively trading for.

From Table 4, the adjusted R^2 value is as high as 83.11%. The average pricing error is 0.21%— a small amount in comparison to what is commonly found. Cross-referencing these results to table 3, it is noticed that FF model performs well from 1963-2008, but after the crisis took place, the R^2 drops from the respective 77.03% to 72.27% implying that the shock may be the reason for the decrease in the explanatory ability of the model. A check in Turkish market verifies that, when the crisis does not significantly damage the market, the model retains a much higher explanatory power, of 83.11%.

That could give an insight as to why Petkova (2006) and Hahn and Lee (2006) document that for the U.S. market the FF model does not work as well as theirs, with a lower adjusted R^2 . The reason could be that their tests do not consider the influence of shocks. Indeed, their testing periods cover up to December 2001 which coincides with a dramatic market downturn

from which, until now, most of the main markets have not reached their heights before 2001. This downturn would have affected the FF as discussed so far in this paper and as a result it performed worse than their two alternatives. Hence it makes sense to devote the next section to further evaluate this hypothesis.

To summarise, a crisis has made an effect on the model's efficiency not just on one market, but in general. It can furthermore tell, when it fails, the model reduces in explanatory power over time, or while the model itself was good it was a victim of the crisis. In general, this methodology can be used in a variety of contexts where a cointegration test can be done. Subsequently the observed isolated sample can be ideally suited for examination and comparison of any model's effectiveness.

Though the FF model, in this paper, shows a reduced explanatory ability because a crisis appears, it is overall a good model as tested here on two markets and widely studied elsewhere. It is, however, imperative to take into account the arguments against. Petkova (2006) and Hahn and Lee (2006) are two studies that are able to document the failures and complete replacements of this model. This paper has provided a few arguments regarding the composition of their models, now upon correction, the paper re-examines these two models to see if they really work as well as created to be. As a general convention, a good model should work for many markets, so the choice of one in this paper should remain an arbitrary one. In the next sections, the models will be tested on the Turkish data where no linkage to a strongly crisis-affected market is detected.

5.3.2. The Petkova (2006) and Hahn and Lee (2006) models

This section aims to evaluate Petkova's and Hahn and Lee's models and their results against those of the FF model.

Table 4 reports the models' effectiveness for the Turkish stocks market. Overall, they also seem to be working well. The adjusted R^2 measure shows the overall statistical goodness of fit of each cross-sectional model. The Petkova's model has the adjusted R^2 of 79.94% implying that a significant 79.94% of the variation in average returns can be explained by this approach. The presence of their alternative variables seems to result in low estimation standard errors and pricing errors. The estimated standard errors for the 1994-2010 period are 0.0710 and the average pricing errors are as low as 0.25 per month¹¹. Similarly, the Hahn and Lee's model captures 79.68% of the movement of the cross sectional average returns. The model standard errors and pricing errors are of the same magnitudes as the Petkova's, 0.0715 and 0.26 respectively.

Although both Petkova and Hahn and Lee documented the two models to be better than the FF, the FF model is performing better in Turkey. Its adjusted R^2 is 3.17% and 3.43% higher than the Petkova's and the Hahn and Lee's, respectively. The estimation errors are also lower for the FF. The FF's estimation is as small as 0.0652 standard errors from zero, and its pricing error is 0.04% less than those of the two alternative models. Moreover, for a model to be a good model the intercept should be indistinguishable from zero, which is precisely the case in the Fama and French's model while the other two's intercepts are statistically significant at 5%.

The reason for this inconsistency in results for these two models in the Turkish and in the U.S. markets, is perhaps due to the structural breaks occurring in the U.S. market were not accounted for. Negative shocks such as financial crises can affect the factor loadings, which aim to capture changes in investment opportunities set, of the two alternative models. According to these studies, there are two aspects of investment opportunity changes that asset

¹¹Vector Auto-regression method (VAR) is also employed to have the results comparable with Petkova's argument and for robustness check. The results are qualitatively similar to the ones presented for the OLS estimation.

pricing models aim to capture: yield curve and conditional distribution of asset returns. In which, short-term T-Bill and term spread are proxies for the former while default spread and dividend yield mean to proxy for the latter aspect. It is sensible to expect that market downturns may overstate the default spread and understate the changes in expectation on interest rates, such as short-term T-bill. The crises see firms facing higher risk of going default, each to difference degree in which distress firms tend to suffer higher barrier than growth firms. The conditional distribution hence tends to be overstated. During bad times, the investment yields also must be high enough to induce people to invest instead of consume. It is therefore vital to account for the downturn shocks when pricing risks of changes in investment opportunity.

In line with the previous hypothesis of the crisis having impact on the models, the paper notices that the factors included by both Petkova and Hahn and Lee are strongly affected by the crisis. During the period, the default likelihood experienced by firm is changed drastically. The Default spread factor – either measured using bankruptcy probability or corporate bond rate – is only a firm specific variable, while the financial crisis would affect the whole market hence manipulate this factor accordingly. This factor proxy misses out the effect of the crisis.

In summary, the paper does not find evidence to support that the two models are better than the FF model. The next section seeks for additional explanatory variables to see if they could improve the models.

5.3.3. Other model augmentations

Augmented Petkova models (5b, 5c, 5d in Table 4) are no better solutions, offering less than 1% increase in R^2 and the pricing error reduced very slightly. Adding TURN and/or WML into Hahn & Lee (6b, 6c, 6d) offers similar effects with the intercepts remain

significant from zero. Model 7: when all factors combined, many variables lose their significance in the regression with the exception of Fama and French's three factors. The augmented Fama and French models increase in all goodness of fit measures, especially when both TURN and WML are included; however the gains in adjusted R^2 are no more than half a percentage point and like other measures, the differences are immaterial.

The FF is well-performing, as adding variables common in testing value premium only helps the model very slightly. One may argue that having these extra factors could expectedly help the model but are simply examples of more redundant variables can be better performing than omitting necessary variables. Perhaps, it was not such a good idea. The results remain consistent when using E/P and DY as classifiers of value¹².

The TURN and WML factor seem to play a significant role at 1% level, when added to all the three: the FF, Petkova or Hahn and Lee. The models, however, see no significant improvement (0.25% at most) in explanatory ability, when appearing with these two factors.

6. Conclusion

As documented earlier, Lakonishok, Shleifer and Vishny (1994) defend traditional measures of standard deviation and beta, saying dwelling into the examination of extremely bad times provides refuge for those looking for proofs that high return strategies are riskier. As extreme states have a tendency to distort the value premium, this study was able to examine the whole time frame of a market without having to worry of the obstacle. The Fama and French model tested on markets such as the Turkish's ISE can prove to have this advantage.

The paper bypasses a number of obstacles when concerning the recent financial shock. It brings together the use cointegration tests to find a market little affected by a financial crisis

¹²The results are available upon request.

in order to test the strength of any models. The results confirm that financial shocks concern the model, and the finding can be applied to all markets. Without the need to consider all individuals, the model's explanatory power against shocks can be tested.

The paper also recognises potential areas for future research, such as incorporating more financial crises in international markets and more statistics testing. They would perhaps provide a stronger evidence of the proposed method validity which is beneficial for any new methods.

Overall, the model considered in this paper – the Fama and French – is one that does not perform better after the 2008 crisis occurred but one that performs better without the financial crisis. However when tested in a market where the crisis impact was relatively low, the model performs well and better than other alternatives. The augmentations do not significantly improve the model.

References

- Arshanapalli, B., Doukas, J., 1993. International stock market linkages: Evidence from the pre- and post-October 1987 period. *Journal of Banking and Finance* 17, 193-208.
- Avramov, D., Chordia, T., 2006. Asset pricing models and financial market anomalies. *Review of Financial Studies* 19(3), 1001-1040.
- Barber, B.M., Lyon, J.D., 1997. Firm size, book-to-market ratio, and security returns: A holdout sample of financial firms. *Journal of Finance* 52(2), 875-883.
- Bhandari, L.C., 1988. Debt/Equity ratio and expected common stock returns: Empirical evidence. *Journal of Finance* 43 (2), 507-528.
- Buelens, C., 2012. Inflation forecasting and the crisis: assessing the impact on the performance of different forecasting models and methods. *European Economy, Economic Papers* 451, European Commission.
- Campbell, J., 1996. Understanding risk and return. *Journal of Political Economy* 104, 298-345.
- Carhart, M. M., 1997. On Persistence in Mutual Fund Performance. *Journal of Finance* 52, 57-82.
- Chan, L.K.C., Jegadeesh, N., Lakonishok, J., 1995. Evaluating the performance of value versus glamour stocks: The impact of selection bias. *Journal of Financial Economics* 38, 269-296.
- Chen, J., 2003. Intertemporal CAPM and the cross-section of stock returns. Working Paper. University of Southern California.
- Clare, A., Thomas, S., 1995. The overreaction hypothesis and the UK stock market. *Economics Letters* 40, 105-112.
- Cooper, I., 2006. Asset pricing implications of nonconvex adjustment costs and irreversibility of investment. *Journal of Finance* 61(1), 139-170.
- Daniel, K., Titman, S., 1997. Evidence on the characteristics of cross sectional variation of stock returns. *Journal of Finance* 52 (1), 1-33.
- Davis, J.L, Fama, E.F., French, K.R., 2000. Characteristics, covariances, and average returns: 1929-1997. *Journal of Finance* 55(1), 389-406.
- Dickey, D.A., Fuller, W.A., 1981. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica* 49, 1057-1072.
- Dimson, E., Nagel, S., Quigley, G., 2003. Capturing the value premium in the United Kingdom. *Financial Analysts Journal* 59(6), 35-45.
- Doukas, J.A., Kim, C., Pantzalis, C., 2002. A test of the errors-in-expectations explanation of the value/glamour stock returns performance: Evidence from analysts' forecasts. *Journal of Finance* 57, 2143-2165.
- Fama, E. F., French, K. R., 1992. The cross-section of expected stock returns. *Journal of Finance* 47 (2), 427-465.
- Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33, 3-56.
- Fama, E. F., French, K. R., 1996. Multifactor explanations of asset pricing anomalies. *Journal of Finance* 50 (1), 55-84.

- Fama, E.F., French, K.R., 2006. The value premium and the CAPM. *Journal of Finance* 61(5), 2163-2185.
- Fama, E.F., MacBeth, J., 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81, 607-636.
- Griffin, J.M., Lemmon, M.L., 2002. Book-to-market equity, distress risk, and stock returns. *Journal of Finance* 57(5), 2317-2336.
- Hahn, J., Lee, H., 2003. Yield spreads as alternative risk factors for size and book-to-market. Working paper, University of Washington.
- Hahn, J., Lee, H., 2006. Yield spreads as alternative risk factors for size and book-to-market. *Journals of Financial and Quantitative Analysis* 41(2), 245-269.
- Jagannathan, R., Wang, Z., 1996. The conditional CAPM and the cross-section of expected returns. *Journal of Finance* 51, 3-53.
- Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: Implications for stocks market efficiency. *Journal of Finance* 48(1), 65-91.
- Johansen, S., 1991. Estimation and hypothesis testing of cointegration vectors in Gaussian Vector Autoregressive models. *Econometrica* 59, 1551-1580.
- Johansen, S., 1995. Likelihood-based inference in Cointegration Vector Autoregressive models. Oxford: Oxford University Press.
- Kothari, S.P., Shanken, J., Sloan, R., 1995. Another look at the cross-section of expected returns. *Journal of Finance* 50, 185-224.
- Lakonishok, J., Shleifer, A., Vishny, R., 1994. Contrarian investment, extrapolation, and risk. *Journal of Finance* 49, 1541-1578.
- Lee, J., Strazicich, M.C., 2003. Minimum Lagrange Multiplier unit root test with two structural breaks. *Review of Economics and Statistics* 85(4), 1082-1089.
- Lettau, M., Ludvigson, S., 2001. Resurrecting the C(CAPM): a cross-sectional test when risk premia are time-varying. *Journal of Political Economy* 109, 1238-1287.
- Lo, A.W., MacKinlay, A.C., 1990. Data-snooping biases in tests of financial asset pricing models. *Review of Financial Studies* 3(3), 431-467.
- Lumsdaine, R.L., Papell, D.H., 1997. Multiple trend breaks and the unit root hypothesis. *Review of Econometrics and Statistics* 79(2), 212-218.
- MacKinlay, A.C., 1997. Event studies in economics and finance. *Journal of Economic Literature* 35(1), 13-39.
- MacKinnon, J.G., 1996. Numerical distribution functions for unit root and cointegration tests. *Journal of Applied Econometrics* 11(6), 601-618.
- MacKinnon, J.G., Haug, A.A., Michelis, L., 1999. Numerical distribution functions of likelihood ratio tests for cointegration. *Journal of Applied Econometrics* 14, 563-577.
- Macovei, M., 2009. Growth and economic crises in Turkey: Leaving behind a turbulent past?. *European Economy, Economic Papers* 386, European Commission.
- Ohlson, J.A., 1980. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research* 18(1), 109-131.

- Pastor, L., Stambaugh, R.F., 2003. Liquidity risk and expected stock returns. *Journal of Political Economy* 111(3), 642-685.
- Perron, P., 1989. The Great Crash, the Oil-price shock and the unit root hypothesis. *Econometrica* 57, 1361-1401.
- Perron, P., 1997. Further evidence on breaking trend function in macroeconomic variables. *Journal of Econometrics* 80, 355-385.
- Petkova, R., 2006. Do the Fama-French factors proxy for innovations in predictive variables?. *Journal of Finance* 61(2), 581-612.
- Phalippou, L., 2007. Can risk-based theories explain the value premium?. *Review of Finance* 11, 143-166.
- Phillips, P.C.B., Ouliaris, S., 1990. Asymptotic properties of residual based tests for cointegration. *Econometrica* 58, 165-193.
- Phillips, P.C.B., Perron, P., 1988. Testing for a unit root in time series regression. *Biometrika*, 75(2), 335-346.
- Rozeff, M. S., Zaman, M. A., 1998. Overreaction and insider trading: evidence from growth and value portfolios. *Journal of Finance* 53(2), 701-716.
- Turhan, I.M. 2011. Turkey's response to the global economic crisis. MPRA Paper 31214, University Library of Munich.
- Westerlund, J., 2006. Testing for panel cointegration with Multiple structural breaks. *Oxford Bulletin of Economics and Statistics* 68(1), 101-132.
- Wo, C., Su, Y-C., 1998. Dynamic relations among international stock markets. *International Review of Economics and Finance* 7(1), 63-84.
- Zivot, E., Andrews, D.W.K, 1992. Further evidence on the Great crash, the Oil-price shock, and the unit root hypothesis. *Journal of Business and Economic Statistics* 10(3), 251-270.
- Zhang, L., 2005. The value premium. *Journal of Finance* 60, 67-103.

Figure 1: Fitted expected returns versus Realized returns for the 25 FF portfolios

This figure shows performance of the Fama and French three factor model in three periods, i.e. 1963-2011, pre-September 2008 and post-September 2008. The average realized returns are shown on the horizontal axis, fitted returns on the vertical axis for 25 size/BM portfolios and the straight line is the 45-degree line. For each portfolio, the realized returns are the time-series average portfolio returns and the fitted expected returns are the fitted value for expected returns from the model.

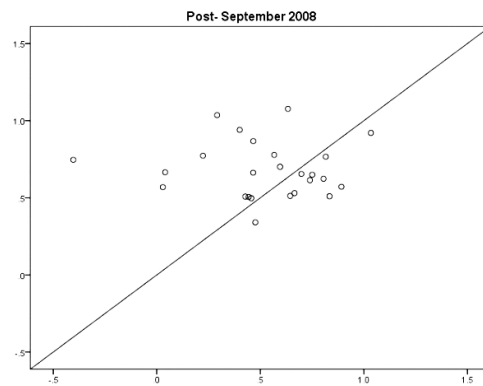
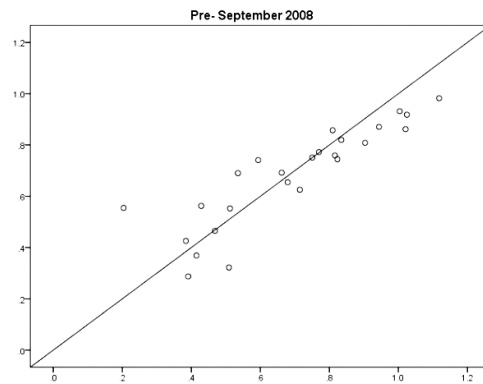
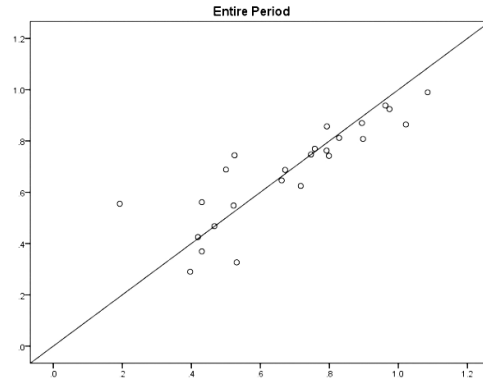


Table 1: Independence Proofs

Unit root (Panel A) and Cointegration test (Panel B and C) results with and without break(s) from January 1988 to December 2010.

Panel A: Unit-root tests

Country	Test on Index level		Test on First differences		Lee and Strazicich (2003)			
	ADF	PP	ADF	PP	Lags	TB ₁	TB ₂	t-LM τ (n)
Turkey	-1.3 (0)	-1.51 (7)	-17.52* (0)	-17.50* (6)	4	2004:05	2008:06	-5.39
Germany	-2.17 (1)	-2.34 (7)	-14.34* (0)	-14.48* (6)	7	2001:02	2007:12	-4.92
Japan	-2.52 (0)	-2.75 (8)	-16.17* (0)	-16.17* (7)	8	1991:03	1993:10	-5.29
UK	-1.8 (0)	-1.93 (8)	-16.12* (0)	-16.13* (7)	3	2003:03	2008:04	-4.34
U.S.	-1.50 (0)	-1.83 (9)	-15.31* (0)	-15.48* (9)	5	2002:10	2007:10	-4.42

Note: The panel presents the Augmented Dickey-Fuller (ADF), Phillips-Perron (PP) tests and multiple structural breaks tests proposed by Lee and Strazicich (2003) for unit roots.

The first two conventional test in the autoregressive representations of Turkish and four major stock indices and on the first differences. For ADF, numbers in parentheses are optimal lag length estimated by Schwarz Information Criterion (SIC). For PP, these are automatic bandwidth based on Newey-West bandwidth selection method and Bartlett kernel. Lee and Strazicich's (2003) procedure allows for two unknown breaks, TB₁ and TB₂, in constant and trend, it employs the Lagrange Multiplier (LM τ) test, in which the lag lengths are selected using Akaike information criteria.

The asterisk indicates statistical significant at the 1% level. Critical values for ADF and PP tests are from MacKinnon (1996) and these for LM τ test are reported in Lee and Strazicich (2003).

Table 1 - continued

Panel B: Cointegration tests without structural breaks

Country	Engle-Granger		Phillips-Ouliaris		Johansen						
	τ - statistic	z- statistic	τ - statistic	z- statistic	Ho: r (rank)	Lag order ^a	Trace ^b (λ_{trace})	Max Eigen ^c (λ_{max})	Ho: r (rank)	Trace ^b (λ_{trace})	Max Eigen ^c (λ_{max})
Germany	-1.59	-4.46	-1.63	-4.74	$r \leq 1$	2	1.72	1.72	$r=0$	5.29	3.57
Japan	-1.97	-7.58	-2.11	-8.77		1	0.91	0.91		4.99	4.08
UK	-1.6	-3.92	-1.57	-3.74		2	1.42	1.42		5.69	4.27
U.S.	-1.16	-2.55	-1.15	-2.49		2	0.74	0.74		6.03	5.29

Note: The Panel displays the static Engle-Granger, Phillips-Ouliaris, the dynamic Johansen tests and most importantly an augmented test from the Engle-Granger allowing for an unknown number of breaks in both level and trend. The asterisk implies the rejection of the null hypothesis of no cointegration.

^a:Lag order determined by Akaike Information Criteria (AIC)

^b:Trace critical value in MacKinnon-Haug-Michelis (1999) at 5% level are 3.8415 and 15.4947 for $r \leq 1$ and $r=0$, respectively

^c:Max-Eigen critical value in MacKinnon-Haug-Michelis (1999) at 5% level are 3.8415 and 14.2646 for $r \leq 1$ and $r=0$, respectively.

Panel C: Cointegration tests with a structural break

Country	Dummy	Augmented Engle-Granger	
		τ -statistic	z-statistic
Germany	Yes	-2.13	-9.03
Japan	Yes	-1.20	-8.14
UK	Yes	-2.08	-8.22
US	Yes	-1.85	-5.66

Note: The test control for a dummy variable, $D(T_B)$, which takes the value of 1 after September 2008 and the value of zero elsewhere. The second column shows if coefficients associated with $D(T_B)$ is significant. See notes to Panel B for Engle-Granger test clarification.

Table 2: Explanatory variables

Descriptive statistics of explanatory variables from July 1994 to December 2010, 198 months

Explanatory returns	Mean	Std	<i>t</i> -statistics
$R_m - R_f$	0.0382	0.1432	3.75***
HML	0.0042	0.0645	0.91
SMB	0.0036	0.0714	0.71
DIV	0.0329	0.0128	36.1***
STBill	0.0317	0.0195	22.9***
TERM	0.0065	0.0109	8.3***
DEF	0.0142	0.0927	2.15**
Δ TERM	0.0001	0.0065	0.01
Δ DEF	0.0029	0.1304	0.31
TURN	-0.0056	0.0770	-1.02
WML	-0.0124	0.1377	-1.45

*: significant at 10% level ** : significant at 5% level ***: significant at 1% level

Note: The mean is presented in percentages. Std is the standard deviations, *t*-statistics is the mean divided by its time-series standard error.

The excess market returns, $R_m - R_f$, is the return differential between market portfolio and 1-month LIBOR which stands for the risk-free asset. The High-minus-Low, HML, and Small-minus-Big, SMB, portfolios meant to mimic the risk factor in returns associated with B/M and with size, respectively, with 1-month lag. The short-term Treasury bill variable, STBill, is a monthly rate of the 3-month Turkish T-Bill and DIV is the one-year aggregate dividend yield. Term spread, TERM, is the difference between monthly long-term government bond rate and the short-term T-Bill rate. Default spread, DEF, is the difference in returns between firms with the highest probability of bankruptcy, measured by Ohlson's (1980) O-score and firms with the lowest O-score. For each month *n*, Δ TERM_{*n*} (Δ DEF_{*n*}) is constructed in accordance to Hahn and Lee (2006) as the difference between the term spread (minus default spread) of month *n* and that of the prior month. The momentum factor is Winner-minus-Loser, WML, based on the Jegadeesh and Titman (1993) momentum strategy of buying high 11-month past returns and selling low 11-month past returns, lagged 1 month. Turnover, TURN, proxy for liquidity factor, is a mimicking portfolio that long in low turnover stocks and short in high turnover ratio stocks. Turnover ratio is defined as trading volume divided by number of shares outstanding.

Table 3: Cross-sectional three-factor regressions for U.S. market

The Fama- MacBeth two-stage estimation of the below Fama-French cross-sectional regressions on the excess returns on 25 Size/BM portfolios over the entire period, pre- and post-crisis:

$$R_{it} - R_{ft} = \gamma_{0,i} + \gamma_{M,i} (R_{mt} - R_{ft}) + \gamma_{HML,i} HML + \gamma_{SMB,i} SMB + \varepsilon_{it}$$

	γ_0	γ_M	γ_{HML}	γ_{SMB}	$t(\gamma_0)$	$t(\gamma_M)$	$t(\gamma_{HML})$	$t(\gamma_{SMB})$	<i>JW-R-square</i>	<i>Pricing Error</i>
7/63-6/11	1.12	-0.65	0.42	0.20	1.37	-0.82	3.47***	1.95*	72.27	0.082
7/63-8/08	1.18	-0.71	0.46	0.20	1.48	-0.93	3.95***	1.98**	77.03	0.084
9/08-12/11	1.05	-0.65	-0.31	0.26	0.26	-0.16	-0.41	0.43	24.46	0.191

*: significant at 10% level **: significant at 5% level ***: significant at 1% level

Note: 25 size/BM stock portfolios are formed as the intersections of five size and five B/M groups in July each year and held to June of the following year. The sample is from July 1963 to December 2011. Monthly returns are from Kenneth French website. The excess market returns, $R_m - R_f$, is the return differential between market portfolio and 1-month Treasury bills which stands for the risk-free asset. The High-minus-Low, HML, portfolios meant to mimic the risk factor in returns associated with B/M and is the difference each month between the average returns of the two high B/M portfolios (S/H and B/H) and those of the two low B/M portfolios (S/L and B/L). Small-minus-Big, SMB, portfolios meant to mimic the risk factor in returns associated with size, is the difference each month between the average returns of the three small portfolios (S/L, S/M and S/H) and those of the three big stock portfolios (B/L, B/M and B/H). *JW-R-squares* and Pricing errors are reported in percentage form.

Table 4: Cross-section regressions for Turkey

Cross-section regression of excess returns on the 10 regressors (incl. intercept) listed in the first column; subsequent columns report estimation results for each model.

	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Intercept	0.0013 (0.64)	0.0016 (0.79)	0.0009 (0.46)	0.0012 (0.61)	0.0166 (2.55**)	0.0050 (2.3**)	0.0045 (0.74)
$R_m - R_f$	1.0122 (75.8***)	0.9946 (70.7***)	1.0131 (75.9***)	0.9954 (70.9***)	0.9773 (64.4***)	0.9866 (67.8***)	0.9948 (68.7***)
HML	0.2054 (6.2***)	0.1721 (5.0***)	0.2012 (6.0***)	0.1677 (4.9***)			0.1595 (4.5***)
SMB	0.4820 (15.9***)	0.4665 (15.4***)	0.4638 (14.9***)	0.4479 (14.3***)			0.4404 (13.5***)
DIV					-0.4251 (-2.6**)		-0.1928 (-1.27)
TERM					-0.0004 (-0.00)	-0.0567 (-3.5***)	0.1784 (0.97)
DEF					0.1080 (4.6***)	0.2356 (0.73)	0.0111 (0.47)
STBill					0.0333 (0.30)		0.0587 (0.58)
TURN		-0.1017 (-3.8***)		-0.1023 (-3.8***)			-0.0952 (-3.3***)
WML			-0.0541 (-2.3**)	-0.0549 (-2.4**)			-0.0575 (-2.7**)
$Adj R^2$	83.11	83.29	83.17	83.36	79.94	79.68	83.35
<i>s.e.</i>	0.0652	0.0648	0.0651	0.0647	0.0710	0.0715	0.0647
<i>Pricing error</i>	0.21	0.21	0.21	0.21	0.25	0.26	0.21
<i>F-statistics</i>	1947	1481	1467	1190	947	1552	661
<i>p-value</i>	0.00	0.00	0.00	0.00	0.00	0.00	0.00

*: significant at 10% level

**: significant at 5% level

***: significant at 1% level

Note: The models numbered in the table are respectively: [1] Fama-French (1993)'s (in **bold**), [2,3,4] Fama-French model augmented by TURN, by WML factor, and by both factors, [5] model proposed by Petkova (2006), [6] model suggested by Hahn and Lee (2006), and [7] model combining all factors so far. For the sake of brevity, in the Hahn and Lee's (model 6), TERM and DEF in the first column represent for Δ TERM and Δ DEF. See notes to Table 2 for explanations of variables. The table presents estimated coefficients and t-statistics (in parentheses). R^2 and standard errors (s.e.) are adjusted for degrees of freedom. F-statistics and their p-value testing the joint significance of the corresponding loadings. Adjusted R^2 's and Pricing errors are reported in percentage form.

Table 4 - continued

Other augmented models

	[5b]	[5c]	[5d]	[6b]	[6c]	[6d]
Intercept	0.0164 (2.5**)	0.0099 (1.52)	0.0098 (1.51)	0.0051 (2.4**)	0.0037 (1.7*)	0.0040 (1.8*)
$R_m - R_f$	0.9639 (61.5***)	0.9807 (65.5***)	0.9680 (62.7***)	0.9650 (63.2***)	0.9918 (69.0***)	0.9706 (64.4***)
HML						
SMB						
DIV	-0.3972 (-2.4**)	-0.2909 (-1.77*)	-0.2668 (-1.63)			
TERM	0.0243 (0.12)	0.2169 (1.11)	0.2367 (1.21)	-0.0435 (-2.7***)	-0.0622 (-3.9***)	-0.0492 (-3.1***)
DEF	0.0781 (3.1***)	0.1182 (5.1***)	0.0899 (3.7***)	0.2604 (0.82)	0.3573 (1.13)	0.3800 (1.21)
STBill	0.0187 (0.17)	0.0170 (0.16)	0.0035 (0.03)			
TURN	-0.1003 (-3.3***)		-0.0945 (-3.1***)	-0.1273 (-4.4***)		-0.1243 (-4.4***)
WML		-0.1545 (-6.2***)	-0.1520 (-6.1***)		-0.1498 (-6.1***)	-0.1480 (-6.1***)
<i>Adj R²</i>	80.10	80.55	80.69	79.99	80.29	80.58
<i>s.e.</i>	0.0707	0.0700	0.0697	0.0710	0.0704	0.0699
<i>Pricing Error</i>	0.25	0.24	0.24	0.25	0.25	0.24
<i>F-statistics</i>	797	820	710	1187	1209	986
<i>p-value</i>	0.00	0.00	0.00	0.00	0.00	0.00

*: significant at 10% level

**: significant at 5% level

***: significant at 1% level

Note: The models numbered in the table are respectively: [5b, 5c, 5d] model 5 augmented by TURN, by WML factor and by both factors, [6b, 6c, 6d] model 6 augmented by TURN, by WML factor and by the both factors. For the sake of brevity, in the augmented Hahn and Lee's models (models 6b, 6c, 6d), TERM and DEF in the first column represent for Δ TERM and Δ DEF, respectively. See notes to Table 2 for explanations of variables. The table presents estimated coefficients and t-statistics (in parentheses). R^2 and standard errors (s.e.) are adjusted for degrees of freedom, and F-tests report the joint significance. Adjusted R^2 's and Pricing errors are reported in percentage form.