Review of Finance (2007) 11: 143–166
doi: 10.1093/rof/rfm014
Advance Access publication: 1 June 2007

Can Risk-Based Theories Explain the Value Premium?*

LUDOVIC PHALIPPOU
University of Amsterdam

Abstract. This paper shows that some of the most prominent risk-based theories offered as explanation for the value premium are at odds with data. The models proposed by Fama and French (1993), Lettau and Ludvingson (2001), Campbell and Vuolteenaho (2004), and Yogo (2006) can capture the cross-section of returns of portfolios sorted on book-to-market ratio and size, but *not* of portfolios sorted on book-to-market ratio and institutional ownership. These models generate economically large pricing errors in all the institutional ownership quintiles and each statistical test indicates that these pricing errors are significant. More generally, these results show that a minor alteration of the test assets can lead to a dramatically different answer regarding the validity of a given asset pricing model.

JEL Classification: G12, G14, G20

1. Introduction

Companies with a high book-to-market ratio (BE/ME), referred to as "value firms", have a higher return than companies with a low book-to-market ratio, or "growth firms". Some authors have warned that this "value premium" may

* I am grateful to Nicolas Barberis, Alexandra Bonczoszek, Jerry Coakley (EFA discussant), Kent Daniel, John Doukas, Joost Driessen, Bernard Dumas, Francesco Franzoni (LSE discussant), Jose-Miguel Gaspar, Harald Hau, Pierre Hillion, Patrick Kelly, Herwig Langohr, Anders Loflund (EFMA discussant), Martin Lettau, Andrew Metrick, David Musto (WFA discussant), Stefan Nagel, Arzu Ozoguz, Marco Pagano (the editor), Lubos Pastor, Joel Peress, Jeff Pontiff, Bruno Solnik, Alex Taylor (EFA discussant), Ning Zhu (BF discussant), and an anonymous referee for their comments. I am also grateful to seminar participants at the University of Amsterdam, UT Austin, UC Berkeley, HEC Paris, INSEAD, Singapore MU, Stockholm school of Economics, Yale University and London School of Economics (joint LBS-LSE-Oxford asset pricing workshop), and the participants at the meetings of the AFFI 2002 in Lyon, EFMA 2002 in Helsinki, Inquire 2003 in Barcelona, WFA 2004 in Vancouver, Behavioral Finance 2004 conference in Notre Dame University, and EFA 2006 in Zurich. Part of this paper has been previously circulated under the titles "What drives the value premium?" and "Institutional Ownership and the value premium". The first draft was written and circulated in November 2001 as the first chapter of the author's doctoral thesis at INSEAD.

result from sample selection biases or data-snooping.[1] Its apparent persistence, however, both out of sample and after correction for selection biases, has lead to a near consensus on its authenticity. As a result, debate has focused on two central lines of argument. The first posits that a high BE/ME implies a higher discount rate. Advocates of this "rational" explanation propose various adaptations of the Capital Asset Pricing Model (CAPM) to capture the premium. The second approach views BE/ME as a proxy for mispricing. A combination of certain systematic errors made by investors with limited arbitrage constitutes the argument.

Recently, a host of asset pricing models have been proposed in support of the first argument. These models are deemed successful because the pricing errors that they generate with 25 size BE/ME sorted portfolios cannot be statistically distinguished from zero.[2] This paper assesses the robustness of some of these results by re-testing the models on a different time period and different set of test assets. This exercise can be seen as an out-of-sample test or a model/data snooping check which is necessary given the many unrelated and deemed successful models that have been offered. The purpose of this article is not to determine which model is correct but to assess the sensitivity of results to a departure from the standard 25 BE/ME size test assets and typical time period frame. The selected models are chosen because the factors are available online and because they are prominent models.

The main change I employ to the conventional asset pricing testing approach is to modify the test assets, namely the 25 size-BE/ME sorted portfolios. As I want to assess explanations for the value premium, the BE/ME sorting dimension should be kept, and I thus need to replace the size dimension. To obtain a wide dispersion of returns and thus have enough power for the test, the selected characteristic needs to be related to stock returns either directly or via a cross-effect with BE/ME. There are several candidate characteristics for the latter. Indeed, several characteristics have been shown to be highly related to the magnitude of the value premium. Namely, the value premium has been found to be primarily concentrated in stocks with low analyst coverage (Griffin and Lemmon, 2002), stocks with high idiosyncratic volatility (Ali et al., 2003), and stocks with low Institutional Ownership (IO) (Nagel, 2005). I find that IO

[1] Kothari et al. (1995) argue that significant biases arise when analysis is conditioned to assets appearing in both the CRSP and COMPUSTAT databases. This claim is, however, disputed by Chan et al. (1995). Ball et al. (1995) stress microstructure/liquidity problems when measuring returns of small value stocks. They suggest forming portfolios at June-end instead of December-end. Finally, Lo and MacKinlay (1990) and Conrad et al. (2002) warn against data-snooping.

[2] A list of models is provided by Lewellen et al. (2006), in Table 1 of Daniel and Titman (2005), and in Appendix A.

is the variable most closely related empirically to the value premium out of all these characteristics. I thus form BE/ME-IO sorted portfolios as an alternative to the standard BE/ME-size sorted portfolios.

Selected models are those of Fama and French (FF, 1993), Lettau and Ludvingson (LL, 2001), Campbell and Vuolteenaho (CV, 2004), and Yogo (Y, 2006). As IO is available only after 1980 and most factors are available until 2001, models are tested (i) with 25 size-BE/ME portfolios on the entire time period (1952–2001 for LL and Y or 1963–2001 for CV and FF), which is similar to the time period on which these models have been originally tested, (ii) with 25 size-BE/ME portfolios on the period 1980–2001, in order to isolate the effect due to the change in time period, and (iii) with 25 IO-BE/ME portfolios on the period 1980–2001.

The first test, which basically replicates the original studies, uses the 25 size-BE/ME portfolios as test assets on the entire time period (1952–2001 for LL and Y or 1963–2001 for CV and FF). I find that each model indeed generates monthly absolute pricing errors that are very low and of the same order of magnitude: 0.10% for FF, 0.11% for CV and 0.12% for LL and Y.

The second test changes the time period to 1980–2001 and leaves test assets unchanged. Pricing errors increase for each model with different magnitudes: from 0.10% to 0.15% for FF, from 0.11% to 0.17% for CV, from 0.12% to 0.23% for Y, and from 0.12% to 0.32% for LL.

Finally, when test assets are changed, the pricing errors increase significantly. They reach 0.27% for FF, 0.31% for CV, 0.30% for Y, and 0.47% for LL. Other statistics for goodness-of-fit such as the R-square proposed by Campbell and Vuolteenaho (2004) indicate a poor fit for each model when presented the 25 IO-BE/ME portfolios (e.g., the maximum CV-R-square is 22% and is generated by the FF model). Furthermore, nontrivial errors are generated for portfolios that contain very large stocks. For example, among the 10 portfolios in the top two IO quintiles, the FF model generates alphas above 10 basis points per month (in absolute value) for 8 portfolios (out of 10), and alphas above 24 basis points per month (in absolute value) for 3 portfolios (out of 10). The spread between the alpha of value stocks and the alpha of growth stocks is a substantial −0.41% for each of the top two IO quintiles, with t-statistics of 2.02 and 2.12, respectively. That is, the three-factor model indicates that there exists a *growth* premium for high-IO stocks that is both statistically and economically significant. Similar results are found with the other models, which show that the failure of these models is not due to economically insignificant stocks.

These results are robust to the estimation methodology and to the choice of goodness-of-fit criterion. I report results for both a Fama-McBeth estimation approach and several Stochastic Discount Factor-Generalized Method of

Moment specifications. Results are also robust to the way portfolios are formed (e.g., value-weighting versus equally weighing portfolios, using NYSE cut-offs or not).

This paper thus shows that just a tiny alteration in test asset construction can have a big effect on how we judge asset pricing models. As a result, the main implication is that authors should always experiment with different representations of the value effect when testing models that seek to explain this effect. An additional implication is that using recent asset pricing models for calculations of the cost of capital might be premature given that some of the most prominent of these models generate large pricing errors. A related point is that it is not clear that the value premium can be traced to risk, at least not to the sources of risk that are proposed by the models tested in this paper.

Recent complementary studies by Daniel and Titman (2005) and Lewellen et al. (2006) argue that the statistical tests used for these asset pricing models lack power. Both studies demonstrate that the finding that returns of BE/ME sorted portfolios are aligned with the loadings on a given factor is *not sufficient* to infer that the model explains the value premium. Daniel and Titman (2005) argue that this is because BE/ME is a 'catch-all' variable. Lewellen et al. (2006) argue that it is the result of a strong common factor structure of stock returns. In this paper, I show that one *can* conclude on the validity of recent risk-based explanations offered for the value premium. If test assets are changed to IO-BE/ME then each model that I test fails to explain *the value premium*.[3]

The rest of the paper reads as follows: Section 2 discusses the selected models. Section 3 lays out the data selection scheme. Section 4 shows the relation between stock characteristics and the value premium. Section 5 presents the methodology. Section 6 discusses the empirical findings. Section 7 gives a brief conclusion.

2. Selected Models

Several models have recently been proposed to explain the existence of a value premium on "rational grounds". A list of 13 such models is given in Table 1 of Daniel and Titman (2005), 8 other models are listed in the introduction of Lewellen et al. (2006), 3 other models of this type that are not included in either study are listed in Appendix A. Out of these 24 models, I have selected the

[3] Daniel and Titman (2005) also show that both the CAPM and the Campbell–Vuolteenaho model are rejected, not because they cannot explain the value premium but because they cannot explain the dispersion of returns after controlling for BE/ME and size. Lewellen et al. (2006) also show that five risk-based models (e.g., Lettau and Ludvingson, 2001, and Yogo, 2005) are rejected, again not because they cannot explain the value premium but because they fail to explain the cross-section of industry portfolios (when added to the 25 size-BE/ME portfolios).

models that are most frequently used and for which data about the factors are available online. This results in the selection of four models (Fama and French, 1993; Lettau and Ludvingson, 2001; Campbell and Vuolteenaho, 2004, and Yogo, 2006). These models are described in the Appendix B. The purpose of this paper is not to test all potential risk-based explanations or to re-evaluate a series of published articles. The purpose is rather to assess how much a simple change in test assets could affect the pricing errors generated by asset pricing models and their conclusion regarding the source of the value premium. The selected models are both among the most prominent asset pricing models and focus on the value premium. As such, they provide ideal case studies for the purpose of this paper.

3. Data and Summary Statistics

Data on accounting and market performance are collected through the Center for Research in Security Preside-COMPUSTAT Merged Database. I also take the complementing accounting data based on *Moody's Industrial Manuals*.[4] Only common and non-financial stocks are included. Delisting returns are taken into account as proposed by Shumway (1997). The book value of equity (BE) is computed as the book value of equity (COMPUSTAT Data Item, C.D.I 60), plus deferred taxes (C.D.I. 74) and investment tax credits (C.D.I 208), minus the book value of preferred stock (C.D.I 56, 10, or 130, in that order). From July of year t to June of year $t + 1$, the book-to-market ratio is: $BE/ME_t = BE_{t-1}/ME_{t-1}$, where ME_{t-1} (BE_{t-1}) is the market (book) value of equity at the end of year $t - 1$.

Negative values of BE/ME are discarded and, in order to reduce selection biases, information about BE_{t-2} is required. I also construct the time-series of stock illiquidity ratios proposed by Amihud (2002).[5] Idiosyncratic volatility is constructed as in Ali et al. (2003). It is denoted IVOL and is obtained by regressing daily returns on the CRSP value-weighted index over a maximum of 250 days ending on June of year t, and then computing the variance of the residuals.

[4] These data can be downloaded on: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. I am very grateful to Jim Davis for generously providing me with many details about the data construction. I am also thankful to Martin Lettau for his comments and for posting the factors on http://pages.stern.nyu.edu/~mlettau, to Motohiro Yogo for his comments and for posting the factors on http://finance.wharton.upenn.edu/~yogo/, and to Tuomo Vuolteenaho for posting the data on his website: post.economics.harvard.edu/faculty/vuolteenaho/papers.html.

[5] Illiquidity ratio is the average, over a year, of the (daily) ratios of the daily absolute return to the dollar trading volume (in million) on that day.

Data about IO from 1980 to 2001 are constructed from the CDA/Spectrum 13F database. These data come from the SEC 13F form in which large institutional investment managers (banks, insurance companies, mutual funds, large brokerage firms, pension funds, and endowments) report their *quarterly* common-stock positions. Gompers and Metrick (2001) provide a detailed analysis of this database and I follow their methodology in constructing IO. IO is defined as the fraction of a company's shares that are owned by institutions filing the 13F form. These data are noisy as a result of the way CDA/Spectrum handles late filings and the fact that small positions in a stock do not have to be disclosed. Late filing is rare (4% of the observations) and concentrated in small stocks. I include such holdings and adjust the reported number of shares if a stock split has occurred. IO data are available from the first quarter of 1980. This study's time period then covers 1980 to 2001, as most factors' data ends in 2001.

4. Value Premium and Stock Characteristics

The main motivation for using 25 size-BE/ME sorted portfolios in testing asset pricing models is that they provide the widest dispersion of returns with a small number of test assets. This is so because both BE/ME and size are directly related to returns. Moreover, the value premium varies greatly as a function of stock size. Small stocks typically exhibit a larger value premium than big stocks.

In this paper, I assess the sensitivity of the pricing errors generated by certain recent models to a change in test assets. As I want to evaluate the validity of these selected models as an explanation for the value premium, the BE/ME sorting dimension should be kept, and I thus need to replace the size dimension. To obtain a wide dispersion of returns and thus have enough power for the test, the selected characteristic needs to be related to stock returns either directly or via a cross-effect with BE/ME. There are several candidate characteristics for the latter (and only a few for the former). Indeed, several characteristics have been shown to be highly related to the magnitude of the value premium.

Griffin and Lemmon (2002) find that firms with highest distress risk exhibit a large value effect, which cannot be explained by the three-factor model of Fama and French (1993). They also find that the value premium is stronger for firms with low analyst coverage. Ali et al. (2005) find that the value premium is more concentrated in stocks with higher idiosyncratic volatility, which they defend as a proxy for cost of arbitrage (see Pontiff, 1996). Nagel (2005) finds that the value premium is stronger for stocks with tighter short sale constraints.

Table I. Correlation between stock characteristics

This table reports the average monthly cross-sectional correlation between the following stock characteristics: institutional ownership (IO), the illiquidity ratio (ILLIQ) of Amihud (2002), the number of analyst following the stocks (Analyst), (minus) idiosyncratic volatility (IVOL), and size. Time period is July 1980 to December 2001.

	IO	Liquidity	Analyst	−1*IVOL	Size (log)
IO	100%	44%	48%	37%	64%
Liquidity	44%	100%	31%	54%	58%
Analyst	48%	31%	100%	25%	65%
−1*IVOL	37%	54%	25%	100%	44%
Size (log)	64%	58%	65%	44%	100%

He uses residual institutional ownership as one of his proxies. Finally, liquidity is also sometimes seen as a potential explanatory variable for the magnitude of the value premium and is therefore included.

I compare the marginal explanatory power of IO, size, liquidity, idiosyncratic volatility, and analyst coverage as a summary variable for the likelihood to observe a value effect.[6] These characteristics have all been argued to be highly related to the value premium but they have not been tested against one another. Table I shows that they are highly correlated, but not perfectly so.

I compare the marginal explanatory power of these characteristics by running independent regressions every quarter with individual stock returns as a dependent variable. The time-series averages of the coefficient estimates are displayed in Table II. The t-statistics are computed in the usual Fama–MacBeth fashion and are adjusted for autocorrelation using the Newey–West method. As certain explanatory variables display extreme skewness, I take a natural logarithm transformation and compute z-scores for the explanatory variables in order to obtain economically meaningful cross-effect variables (i.e., scale free cross-effects).

I find that the book-to-market effect is stronger for stocks with high idiosyncratic volatility, small stocks, stocks that have low analyst coverage, and stocks that are held mainly by individual investors as reported in the literature (Griffin and Lemmon, 2002, Ali et al., 2003 and Nagel, 2005). When comparing these effects separately, IO appears to have the largest marginal explanatory power for the value effect. When all the characteristics are included in the regression, the cross-effect of IO and BE/ME is the only one that is

[6] Nagel (2005) shows that IO is more related to the value premium than size but he restricts his sample to mid-cap and large-cap stocks (50% of the universe in number of stocks) and does not compare the explanatory power of IO to the other characteristics.

Table II. Marginal effects of IO—regression evidence

This table reports the average slope of independent cross-sectional regressions of individual stock returns on various stock characteristics, from Q3 1980 to Q4 2001. The set of characteristics include the book-to-market ratio (BE/ME), size, and cross-effects of BE/ME with size, the inverse of idiosyncratic volatility ($IVOL^{-1}$), the number of analyst following the stock (Analyst), the inverse of the illiquidity ratio ($ILLIQ^{-1}$), and institutional ownership (IO). Size, BE/ME, IVOL, Analyst, and ILLIQ are expressed as ln(X). The z-score of each dependant variable is computed by subtracting its cross-sectional mean and then dividing by its standard deviation. Fama–McBeth t-statistics (Newey–West adjustment for autocorrelation) are reported below each coefficient, in italics. A constant is included but not reported. There are on average 2181 observations per quarter. *, **, *** are attributed to coefficients that are statistically different from zero at the 10%, 5%, and 1% significance level, respectively.

Dep. variable is	Specification						
stock return	0	1	2	3	4	5	6
BE/ME	3.81***	3.67***	3.32**	3.78***	3.81***	3.31**	3.32**
	2.91	*2.76*	*2.55*	*2.78*	*2.91*	*2.55*	*2.44*
Size	1.44	1.45	1.48	1.33	1.44	1.28	1.34
	0.89	*0.89*	*0.91*	*0.83*	*0.89*	*0.79*	*0.83*
BE/ME*Size		−0.99*					1.34
		−1.83					*1.59*
BE/ME*$IVOL^{-1}$			−1.45***				−1.07*
			−2.81				*−1.66*
BE/ME*Analyst				−1.12**			*−0.25*
				−2.26			*−0.47*
BE/ME*$ILLIIQ^{-1}$					−0.77		−0.76
					−1.07		*−1.03*
BE/ME*IO						−2.11***	−2.02***
						−3.73	*−2.94*
Average R^2	3.48	3.67	3.77	3.63	3.54	3.66	4.30

significant at a 5% level test. From these results, we can deduct that forming portfolios based on BE/ME and IO is likely to lead to a wide dispersion in stock returns and thus offer an appropriate out-of-sample test for the selected explanations for the value premium.

It is both important and interesting to assess the difference between sorting on IO and sorting on size in the context of the value premium. The correlation between IO and size is not perfect (see Table I) but the out-of-sample test is more relevant if IO has a distinct effect than size. To assess how forming portfolios on BE/ME-IO is different to forming portfolios on the traditional BE/ME-size dimensions, I report the results from a simple exercise that consists in creating terciles based on size (IO), and within each of these, terciles based on IO (size). The value premium in each tercile is then computed.

Table III. Marginal effect of size and institutional ownership

This table reports the marginal effect of size and institutional ownership (IO) on the value premium. Stocks in panel A are first separated in terciles according to their size. Then, each of these terciles is divided into three IO-terciles. Panel B is formed the same way but stocks are first assigned to IO-terciles and then to size-terciles. Finally, the pooled average of IO (in percentage) and size (in million) are reported for each group.

Panel A

	Size-terciles								
	Small IO-terciles			Mid IO-terciles			Big IO-terciles		
	Low	Blend	High	Low	Blend	High	Low	Blend	High
Ret value	2.64	2.13	1.74	1.55	1.54	1.45	1.54	1.53	1.34
Ret growth	0.81	1.02	1.22	−0.42	0.61	1.01	0.70	1.31	1.13
Value premium	**1.73**	**1.11**	**0.52**	**1.93**	**0.93**	**0.44**	**0.84**	**0.22**	**0.21**
	3.94	*3.32*	*1.61*	*6.02*	*3.33*	*1.61*	*2.33*	*0.73*	*0.94*
Mean IO	1%	6%	22%	9%	25%	47%	26%	48%	66%
Mean Size	$12	$16	$20	$82	$99	$123	$2965	$3872	$2557

Panel B

	IO-terciles								
	Low Size-terciles			Blend Size-terciles			High Size-terciles		
	Small	Mid	Big	Small	Mid	Big	Small	Mid	Big
Ret value	2.93	1.63	1.52	1.62	1.63	1.41	1.42	1.44	1.42
Ret growth	2.52	−0.01	−0.60	0.92	0.70	0.81	1.10	1.14	1.21
Value premium	**0.41**	**1.62**	**2.12**	**0.70**	**0.93**	**0.60**	**0.32**	**0.30**	**0.21**
	0.61	*4.73*	*6.41*	*2.30*	*3.16*	*1.72*	*1.32*	*1.03*	*0.61*
Mean IO	3%	5%	7%	21%	24%	26%	50%	55%	58%
Mean Size	$6	$19	$154	$37	$92	$2124	$549	$569	$5461

Panel A of Table III shows that, within each size tercile, the value premium decreases dramatically with IO, which indicates a strong marginal effect of IO. Small stocks with a high IO do not exhibit a significant value premium even though they are very small ($20 million average market capitalization). Intermediate size stocks—that are substantially bigger but have a low IO—have, in contrast, a value premium as high as 1.9% per month and a market capitalization of $82 million on average. The domination of IO over size is even clearer when I perform the reverse operation, i.e., rank first by IO and then by size (Panel B). The value premium in the low-IO stocks is actually concentrated in the largest stocks ($154 million on average) for which it reaches a staggering 2.1% per month. Among the other IO terciles, the value

premium is smaller and similar across size sub-samples, despite wide variations in market capitalization. Size has therefore no marginal explanatory power over IO in explaining the value premium. This result also explains why the subsequent tests are unaffected by using value-weighted portfolios instead of equally weighted portfolios.

5. Methodology

5.1 TEST ASSETS

The "benchmark" test assets are the 25 size-BE/ME portfolios posted on French's website. The proposed alternative test assets consist of 25 quarterly rebalanced portfolios sorted on IO and BE/ME ratio. The returns are value-weighted, computed at a monthly frequency from July 1980 to December 2001, and NYSE-based IO cut-off points are used to form IO quintiles. For example, on December 31,1981, stocks are classified according to their IO level at that date. They are assigned to one of the five IO groups and the value-weighted return of each group in January 1982 is computed. If a stock is delisted in January 1982, then its return is the "delisting return" for that month and the portfolio is rebalanced across surviving stocks. The return for February 1982 is then computed and soon. For the models with quarterly frequency factors, these portfolio returns are compounded to offer quarterly returns. Finally, as in Fama and French (1993), the sorts on IO and BE/ME are made independently.

Table IV reports the average return for the two main sets of test assets from 1980 to 2001: 25 size-BE/ME (Panel A) and 25 IO-BE/ME (Panel B). In Panel A, it can be seen that in contrast to the period 1963–1990, from 1980 to 2001, small stocks do not outperform. Among growth stocks, it is actually big stocks that outperform. This pattern change has been documented at length (e.g., Gompers and Metrick, 2001). It is also worth noting that the value premium varies from 1.1% per month among small stocks to 0.13% per month among big stocks. Panel B shows that there is no relationship between IO and returns except among growth stocks (two bottom quintiles), where it is significantly negative. The most striking observation in Panel B is the relationship between the value premium and IO. The value premium decreases from an extremely high 1.8% per month (t-statistic = 5.15) for stocks in the lowest IO quintile to a negligible 0.2% per month (t-statistic = 0.85) for stocks belonging to the highest IO quintile.

Table IV. Average returns of test portfolios

This table reports the average monthly return of various portfolios and the implied value premium between July 1980 and December 2001. The portfolio returns are all value-weighted by market capitalization. Panel A reports the average return of the 25 size-BE/ME sorted portfolios (posted on French's website). Panel B reports the average return of the 25 IO-BE/ME portfolios computed as in Fama and French (1993).

Panel A: 25 Size-BE/ME portfolios, 1980–2001

	Small	2	3	4	Big	Small-Big	*t*-stat
Value	1.59	1.53	1.67	1.47	1.44	0.15	*0.51*
2	1.64	1.58	1.39	1.44	1.25	0.39	*1.05*
3	1.5	1.52	1.33	1.32	1.23	0.27	*0.84*
4	1.39	1.28	1.39	1.26	1.29	0.10	*0.34*
Growth	0.49	0.9	1.06	1.32	1.31	−0.82	*−2.31*
Value premium	1.10	0.63	0.61	0.15	0.13	0.97	*2.55*
t-stat	*3.66*	*2.20*	*1.82*	*0.50*	*0.53*		

Panel B: 25 IO-BE/ME portfolios, 1980–2001

	Low	2	3	4	High	Low-High	*t*-stat
Value	1.69	1.46	1.44	1.19	1.46	0.23	*0.82*
2	1.29	1.32	1.15	1.45	1.19	0.10	*0.32*
3	1.17	0.85	1.02	1.3	1.3	−0.13	*−0.34*
4	0.36	1.1	1.21	1.24	1.26	−0.90	*−2.81*
Growth	−0.09	0.15	0.79	1.24	1.24	−1.33	*−3.12*
Value premium	1.79	1.31	0.65	−0.05	0.23	1.56	*3.31*
t-stat	*5.15*	*2.87*	*1.65*	*−0.15*	*0.85*		

5.2 ECONOMETRIC APPROACHES

There exist two main approaches for testing an asset pricing model. The first approach is the so-called "covariance estimation" (or SDF estimation) and the second approach is the so-called "two-pass regression" (or Fama–McBeth estimation).[7] In this section, I detail these two approaches and detail the different specifications that I use to test the asset pricing models.

The covariance estimation of a K-factor asset pricing model when presented to N test assets for a time period going from $t = 1$ to $t = T$ can be summarized as follows: Let us define an SDF m (1 by T), a matrix of factors f (K by T), a vector of ones O_T (T by 1), and a set of factor loadings b (K by 1). The SDF is defined as $m = O_T - b'(f - E(f))$ and the assets to be priced is the time-series of the excess returns of N portfolios, denoted $R_t - Rf_t$. The pricing error for the model under consideration for asset n is:

[7] Details are provided by Cochrane (2005).

$$g_{T,n}(b) = \frac{1}{T}\sum_{t=1}^{T}(R_{t,n} - Rf_{t,n})m_t.$$

Denoting $g_T(b) = [g_{T,1}(b), \ldots, g_{T,N}(b)]$, the covariance estimation consists in finding the vector of factor loadings (b^*) that minimizes the weighted sum of pricing errors $g_T(b)Wg_T(b)'$. In this case, only N moments enter this optimization. In certain specifications, K moments (and K unknowns) are added by allowing the sample mean of the factors to deviate from the true mean. In this case the matrix W is $(N + K)$ by $(N + K)$ and $g_T(b)$ is $(N + K)$ by 1. The pricing error generated by the model under consideration for asset n is thus

$$g_{T,n}(b^*) = \frac{1}{T}\sum_{t=1}^{T}(R_{t,n} - Rf_{t,n})m_t(b^*).$$

The second approach–Fama-McBeth estimation–consists in first running N Ordinary Least Square (OLS) time-series regressions:

$$(R_{t,n} - Rf_{t,n}) = a_{t,n} + \beta_n' f_t + \varepsilon_{t,n} \qquad t = 1, \ldots, T \text{ for each asset } n,$$

and, as a second step, running T OLS cross-section regressions:

$$(R_{t,n} - Rf_{t,n}) = \beta_n'\lambda_t + u_{t,n} \qquad n = 1, \ldots, N \text{ for each time period } t.$$

The pricing error generated by the model under consideration for asset n is

$$u_n = \frac{1}{T}\sum_{t=1}^{T} u_{t,n}.$$

Note that parameters b and λ are related as follows:

$$\lambda = E(ff')b.$$

Note also that when the "covariance estimation" is chosen, there is myriad of specifications that the researcher can choose from and each of the six models under investigation uses a different methodology. In this study, I report the results for eight different methods:

Estimation Method 1. One-step GMM with the second-moment matrix as weight. This is advocated by Hansen and Jagannathan (1997). I use $N + K$ moments corresponding to N equations related to the pricing errors of each portfolio and K equations related to the difference between the mean of the factor in sample and its theoretical mean (see above).

Estimation Method 2. One-step GMM with the identity matrix as weight and $N + K$ moments.

Estimation Method 3. Two-stage GMM, the weighting matrix is thus the so-called optimal one (i.e., the inverse of the spectral density matrix). To estimate the spectral density matrix, the Newey–West methodology (i.e., a Bartlett kernel) is used and the number of lag is set to one. $N + K$ moments are used.

Estimation Method 4. As method 1 but with N moments. That is, the mean of the factor is constrained to be equal to the in-sample mean.

Estimation Method 5. As method 2 but with N moments. That is, the mean of the factor is constrained to be equal to the in-sample mean.

Estimation Method 6. As method 3 but with N moments. That is, the mean of the factor is constrained to be equal to the in-sample mean.

Estimation Method 7. Fama–McBeth estimation (see above).

Estimation Method 8. Campbell–Vuolteenaho (2004) estimation: this approach will be applied to their model only. Their two betas are defined as:

$$\beta_{i,DR} = \frac{\text{cov}(R_{t,i} - N_{t,DR}) + \text{cov}(R_{t,i} - N_{t-1,DR})}{\text{Var}(N_{t,CF} - N_{t,DR})}$$

and

$$\beta_{i,CF} = \frac{\text{cov}(R_{t,i} - N_{t,CF}) + \text{cov}(R_{t,i} - N_{t-1,CF})}{\text{Var}(N_{t,CF} - N_{t,DR})}$$

where N_{DR} and N_{CF} are the estimated cash flow and expected return news from a VAR model. Note that their beta estimators deviate from the usual OLS estimator in two respects. First, they include one lag of the market news term in the numerator. Second, the covariance is normalized by the sample variance of the unexpected market return.

Cochrane (2005) argues at length in favor of prespecified weighting matrices of the type described above (methods 1, 2, 4 and 5). He points out that using such a methodology rather than the traditional two-step efficient GMM used in the literature (methods 3 and 6) has the advantage of avoiding the "trap of blowing up standard errors rather than improving pricing errors" and leads to estimates that are more robust to small model misspecifications.

For each model, I report the following statistics that are commonly used to evaluate the goodness of fit of a model:

1. Mean absolute pricing error for the N portfolios generated by each method.
2. CV-R-square: one minus the inverse of the variance of average portfolio returns. A negative CV-R-square says that the model has less explanatory power than a model that predicts constant returns across portfolios (see Campbell and Vuolteenaho, 2004).
3. Hansen and Jagannathan (1997) distance measure: HJ-distance. It is the least-square distance between the tested pricing kernel and the closest point in the set of pricing kernels that price the assets correctly. The distribution of this statistic is a weighted sum of $n-k$ iid random variables distributed as a chi-square with one degree of freedom.
4. J-test: provides a test of overidentifying restrictions, which in turn tests the null hypothesis that the pricing errors are jointly zero across the N test assets.
5. Chi-square test that all pricing errors are jointly equal to zero when using the Fama–McBeth representation, with standard errors correction as recommended by Shanken (1992).
6. JW-R-square. This is a measure proposed by Jagannathan and Wang (1996). It is defined as: JW-R-square $= 1 - V(e)/V(r)$, where, $V(r)$ is the variance of the time-series-average of returns across the N portfolios and $V(e)$ is the variance of the time-series-average of pricing errors across the N portfolios.

Some remarks are in order about the different methodologies. First, method 4, and to a lower extent, method 2, are very similar to the Fama–McBeth approach. Second, Lettau and Ludvigson (2001) point out that in small samples the SDF approach (except methods 2 and 4 as far as pricing errors are concerned) is not suited for small time-series like theirs (146 observations).[8] Such a remark applies to the present study as the number of time-series

[8] Using long time series as done in most studies is also problematic because both the Fama-McBeth approach and the SDF approach requires certain stationarity assumptions, which might be a strong assumption when considering a 50-year time interval.

observations is 86 for quarterly-frequency tests and 258 for monthly-frequency tests. Consequently, I will report only results from the Fama-McBeth approach when performing quarterly-frequency tests and measures such as HJ-distance and J-test will receive relatively low emphasis. As in the literature, the quality of a model will be mainly judged on the economic magnitudes of the pricing errors. This is the only common statistic across all models and probably the most economically relevant.

6. Empirical Findings

6.1 THE THREE-FACTOR MODEL OF FAMA AND FRENCH (1993)

I start with a test of the three-factor model of Fama and French (1993) using both the Fama–McBeth approach and the SDF approach (methods 1 to 4).[9] Results are reported in Table V.

To begin with, I replicate the well-known result that the model of Fama and French (1993) is successful at pricing the 25 size-BE/ME portfolios from 1963 to 2001. The R-squares are above 67% and the average absolute pricing error is as low as 0.10% per month on average (Panel A). When the time period is changed to more recent years, the model generates larger errors, that is the model is more successful over the time period used in Fama and French (1993). Panel B shows that pricing errors for the 1980–2001 period average 0.15% per month. In addition, the R-square is reduced by half and the HJ-distance more than doubles compared to the overall time period (Panel A).

Panel C displays the goodness of fit of the model when test assets are changed to 25 IO-BE/ME portfolios. The average pricing error (in absolute value) jumps to 0.27% per month, that is above 3% on an annual basis and about three times the magnitudes found in Panel A for the 25 size-BE/ME portfolios from 1963 to 2001. This average pricing error is the same across methodologies and the R-square of both Jagannathan and Wang (1996) and Campbell and Vuolteenaho (2004) are about 40%. The chi-square test for pricing errors being jointly zero (the null hypothesis) strongly rejects this hypothesis.

Panel D displays the pricing errors for each portfolio obtained with the Fama–McBeth approach. Eight portfolios out of 25 display alphas above 25 basis points per month (in absolute value). Even among high IO portfolios, which tend to contain very large stocks, the errors are substantial. Among the 10 portfolios in the top two IO quintiles, 8 have an alpha above 0.10 basis points per month (in absolute value) and 3 have an alpha above 25 basis points

[9] Methods 5 and 6 described above also generate very similar pricing errors and J-statistics for each model; results obtained with these methods are therefore omitted in the tables.

Table V. Three-factor model and portfolios formed on Institutional Ownership (IO) and book-to-market(BE/ME), 1980–2001

This table shows the performance of the three-factor model of Fama French (1993) on different time periods and for different test assets. Panels A, B, and C report statistics that measure the goodness of fit of the model when tested with 25 size-BE/ME portfolios from 1963 to 2001, with 25 size-BE/ME portfolios from 1980 to 2001, and with 25 IO-BE/ME portfolios from 1980 to 2001, respectively. Panels D and E show the monthly abnormal performance of the 25 IO-BE/ME and 25 size-BE/ME portfolios respectively derived with the Fama–McBeth methodology. All returns are value-weighted when aggregated at the portfolio level and are sampled at a monthly frequency. The different methods of estimation are detailed in the text. In Panels A to C, the first column reports results from the Fama–McBeth approach, and the other four columns report results from the various specifications of the SDF approach.

Panel A: Goodness of fit for the three-factor model, 1963–2001, 25 size-BE/ME portfolios

	FM	SDF-S1	SDF-S2	SDF-S3	SDF-S4
JW-*R*-square	67%				
CV-*R*-square	68%				
Chi-square or J stat	72	74	74	72	71
P-value chi-square	0.00	0.00	0.00	0.00	0.00
Average abs. pricing error	0.10%	0.11%	0.11%	0.10%	0.11%
HJ-distance					0.17

Panel B: Goodness of fit for the three-factor model, 1980–2001, 25 size-BE/ME portfolios

FM	SDF-S1	SDF-S2	SDF-S3	SDF-S4	
JW-*R*-square	30%				
CV-*R*-square	33%				
Chi-square or J stat	118	145	144	130	128
P-value chi-square	0.00				
Average abs. pricing error	0.15%	0.31%	0.16%	0.15%	0.15%
HJ-distance					0.49

Panel C: Goodness of fit for the three-factor model, 1980–2001, 25 IO-BE/ME portfolios

	FM	SDF-S1	SDF-S2	SDF-S3	SDF-S4
JW-*R*-square	20%				
CV-*R*-square	22%				
Chi-square or J-stat	67	64	60	65	63
P-value chi-square	0.00				
Average abs. pricing error	0.27%	0.26%	0.27%	0.33%	0.26%
HJ-distance					0.27

Table V. Three-factor model and portfolios formed on IO and BE/ME, 1980–2001 (continued)

Panel D: Pricing errors generated by the FF model—25 IO-BE/ME portfolios, 1980–2001

	Alphas					
	Growth	2	3	4	Value	Value premium
Low	−1.28	−0.87	−0.07	0.43	0.53	***1.81***
2	−0.61	0.02	−0.37	0.06	0.17	***0.78***
3	−0.12	−0.09	−0.19	−0.10	0.10	***0.22***
4	0.17	0.06	0.11	0.13	−0.24	***−0.41***
High	0.27	−0.02	−0.14	−0.40	−0.14	***−0.41***
			t-statistics			
Low	−4.53	−3.22	−0.24	1.74	1.89	*5.51*
2	−1.77	0.10	−1.73	0.39	0.90	*2.31*
3	−0.47	−0.46	−1.03	−0.69	0.52	*1.34*
4	1.07	0.41	0.67	0.93	−1.31	*−2.02*
High	2.13	−0.17	−0.98	−3.01	−0.88	*−2.12*
Average abs. pricing error: 0.27%						
CV-*R*-square: 22%						

Panel E: Pricing errors generated by the FF model–25 size-BE/ME portfolios, 1963–2001

	Alphas					
	Growth	2	3	4	Value	Value premium
Low	−0.39	0.05	0.04	0.19	0.13	***0.52***
2	−0.16	−0.07	0.09	0.09	0.04	***0.20***
3	−0.06	0.02	−0.08	0.01	0.03	***0.08***
4	0.16	−0.20	−0.05	0.06	−0.07	***−0.23***
High	0.22	−0.03	−0.03	−0.10	−0.18	***−0.40***
			t-statistics			
Low	−3.64	0.60	0.54	2.91	1.97	***2.25***
2	−2.04	−1.02	1.33	1.46	0.61	***1.31***
3	−0.75	0.21	−1.01	0.12	0.31	***0.51***
4	2.11	−2.32	−0.62	0.84	−0.77	***1.46***
High	3.56	−0.37	−0.32	−1.44	−1.85	***1.99***
Average abs. pricing error: 0.10%						
CV-*R*-square: 68%						

per month (in absolute value). The spread between the alpha of value stocks and the alpha of growth stocks is a staggering −0.41% for each of the top two IO quintiles, with *t*-stats of 2.02 and 2.12 respectively. That is, the three-factor model indicates that there exists a growth premium for high-IO stocks that is both statistically and economically significant. This shows that the failure of this model is not due to economically insignificant stocks. Overall, both the large growth premium for high-IO stocks and the large value premium

for low-IO stocks are inconsistent with the claim that the three-factor model captures well the cross-section of returns.[10]

To conclude, the three-factor model generates *slightly* larger pricing errors than initially reported (Fama and French, 1993) when the time period is changed and generates *substantially* larger pricing errors when the test assets are changed.

6.2 THE CAMPBELL AND VUOLTEENAHO (2004) MODEL

I now turn to the model of Campbell and Vuolteenaho (CV, 2004). They propose an economically motivated two-beta model. These two betas are the result of breaking the CAPM beta into two components: one reflecting news about future market's cash flows and one reflecting news about the market's discount rates.

For comparability purposes, I follow their unique approach to test their model, I present the results separately and use as goodness of fit their criteria the CV-R-square and average absolute pricing error (with the intercept restricted to be equal to the risk-free rate). In addition, the two betas of Campbell and Vuolteenaho (CV, 2004) are computed from the data posted by Vuolteenaho on his website. Consequently, unlike for the other tested models, it is not possible to make direct inference on the statistical significance of each individual alpha that would be comparable to other models. I then focus on the pricing errors generated by this model on the two different time periods and two different test assets.

Results are reported in Table VI. First, I replicate CV's findings (Panel A) and find that the model has indeed an impressive fit for the 25 size-BE/ME portfolios for the period 1963–2001. The R-square is higher than 60% and the average pricing error is as small as 0.11%. That is, the pricing errors are of the same magnitude as Fama–French model while their model uses only two factors that are theoretically motivated and economically appealing. When the same test is done for the period 1980–2001, however, the fit deteriorates (Panel B). The R-square decreases to 24% and the average pricing error increases to 0.17%. Finally, when presented to IO-BE/ME portfolios, the model has an R-square that decreases to 14% and the average pricing error rises to 0.31%,that is nearly doubles (Panel C). The CV-R-square becomes slightly

[10] The magnitude of the failure of the three-factor model is striking here and no such magnitude has been found in the literature. However, the model's failure in itself is not a new result. For example, Griffin and Lemmon (2002) report that large pricing errors are also generated for large stocks with high probability of financial distress. They find that the value premium left unexplained is 0.47% per month for these stocks.

Table VI. Campbell-Vuolteenaho model

This table shows the monthly abnormal performance of portfolios calculated from the two-factor model of Campbell and Vuolteenaho (2004). The two betas are computed as in Campbell and Vuolteenaho (2004). In Panels A, B and C, the performance of the model is reported for 25 size-BE/ME portfolios from 1963 to 2001, 25 size-BE/ME portfolios from 1980 to 2001, and 25 IO-BE/ME portfolios from 1980 to 2001, respectively.

Panel A: Pricing errors generated by the CV model—25 size-BE/ME portfolios, 1963–2001

	Growth	2	3	4	Value	Value Premium
Small	−0.39	0.06	0.03	0.20	0.10	***0.49***
2	−0.02	−0.05	0.10	0.16	0.13	***0.15***
3	0.05	0.06	−0.05	0.01	0.11	***0.06***
4	0.22	−0.24	−0.06	0.07	−0.01	***−0.23***
Big	0.13	−0.11	−0.10	−0.19	−0.19	***−0.32***
Average abs. pricing error: 0.11%						
CV-*R*-square: 62%						

Panel B: Pricing errors generated by the CV model—25 size-BE/ME portfolios, 1980–2001

	Growth	2	3	4	Value	Value Premium
Small	−0.66	0.14	0.21	0.28	0.07	***0.73***
2	−0.13	−0.03	0.17	0.25	0.10	***0.23***
3	0.16	0.08	−0.04	−0.01	0.21	***0.05***
4	0.31	−0.19	−0.24	0.02	0.07	***−0.24***
Big	0.21	−0.15	−0.16	−0.34	−0.05	***−0.26***
Average abs. pricing error: 0.17%						
CV-*R*-square: 24%						

Panel C: Pricing errors generated by the CV model—25 IO-BE/ME portfolios, 1980–2001

	Growth	2	3	4	Value	Value Premium
Low	−1.33	−0.63	0.32	0.62	0.58	***1.92***
2	−0.35	−0.05	−0.12	0.24	0.13	***0.49***
3	−0.27	0.07	−0.10	0.12	0.55	***0.82***
4	0.49	0.05	0.31	0.12	0.28	***−0.22***
High	0.43	0.13	0.03	−0.34	0.08	***−0.34***
Average abs. pricing error: 0.31%						
CV-*R*-square: −1%						

negative (−1%) which means that a constant expected return model would perform slightly better at capturing the cross-section of returns.

Analyzing the alphas (Panel C) reveals that, as for the Fama-French model, even high-IO portfolios, and thus large stocks, exhibit large pricing errors.

It appears also that the errors are concentrated on the growth portfolios: the low-IO growth stocks are expected to have a *higher* return than what is observed (alpha is −1.33% per month) but the high-IO growth stocks are expected to have a *lower* return than what is observed (alpha is 0.43% per month). Compared to the three-factor model of Fama and French (1993), it performs better for value stocks but worse for growth stocks. The conclusion is the same for the CV model as it is for the three-factor model. The CV model generates *slightly* larger pricing errors than initially reported (Campbell and Vuolteenaho, 2004) when the time period is changed and generates *substantially* larger pricing errors when the test assets are changed.

6.3 OTHER MODELS

Two other important models recently proposed are the consumption-based asset pricing model of Lettau and Ludvingson (2001) and Yogo (2006). These models were originally tested at a quarterly frequency starting in 1952. I then report the pricing errors and other goodness-of-fit measures of these two models for the 1952–2001 period and original test assets (25 size-BE/ME portfolios), for the 1980–2001 period and original test assets, and for the 1980–2001 period and different test assets (25 IO-BE/ME portfolios). Results are reported in Table VII.

Lettau and Ludvingson (2001) argue that value stocks earn higher average returns than growth stocks, as they are more highly correlated with consumption growth in bad times. They show that their conditional consumption CAPM—denoted (C)CAPM—captures the value effect far better than unconditional specifications, and about as well as the three-factor model of Fama and French (1993). I confirm these results for both the period 1952–2001 and for the period 1963–2001.[11] LL's model generates pricing errors as low as 0.12% on a monthly basis (Panel A). However, when the time period is changed, the average absolute pricing error increases dramatically to 0.32% monthly. When, in addition, test assets are modified, the average absolute pricing error reaches 0.47% per month. Most of the increase in the pricing errors is thus due to the change in time period but the change in test asset has a nontrivial impact.

Results are very similar for Yogo's model, which is also a consumption-based asset pricing model (Panel B). For the period 1952–2001, pricing errors are indeed small and comparable to those of the ad hoc three-factor model of FF as they average 0.12% per month. When the time period is changed, the

[11] The results for 1963–2001 are omitted in the tables as they are very similar to the 1952–2001 results.

Table VII. Quarterly frequency models

This table shows the performance of the models of Yogo (2005) and Lettau and Ludvigson (2001) when presented either 25 IO-BE/ME portfolios or 25 size-BE/ME portfolios to price, either from 1980 to 2001 or from 1952 to 2001. All returns are value-weighted when aggregated at the portfolio level and are sampled at a quarterly frequency. Goodness of fit is assessed with the Fama–McBeth methodology (see text.)

Panel A: Lettau and Ludvigson (2001), goodness of fit with different test assets and time periods

1952–2001, 25 size-BE/ME Portfolios				
JW-R-square	CV-R-square	Chi-square	P-value	Average abs. pricing error
39%	41%	26	0.26	0.12% (monthly)
1980–2001, 25 size-BE/ME Portfolios				
JW-R-square	CV-R-square	Chi-square	P-value	Average abs. pricing error
−2%	−2%	23	0.42	0.32% (monthly)
1980–2001, 25 IO-BE/ME Portfolios				
JW-R-square	CV-R-square	Chi-square	P-value	Average abs. pricing error
−69%	−87%	42	0.00	0.47% (monthly)

Panel B: Yogo (2005), goodness of fit with different test assets and time periods

1952–2001, 25 size-BE/ME Portfolios				
JW-R-square	CV-R-square	Chi-square	P-value	Average abs. pricing error
28%	31%	17	0.77	0.12% (monthly)
1980–2001, 25 size-BE/ME Portfolios				
JW-R-square	CV-R-square	Chi-square	P-value	Average abs. pricing error
−50%	−47%	29	0.15	0.23% (monthly)
1980–2001, 25 IO-BE/ME Portfolios				
JW-R-square	CV-R-square	Chi-square	P-value	Average abs. pricing error
18%	18%	19	0.65	0.30% (monthly)

average absolute pricing error increases to 0.23% monthly. When, in addition, test assets are modified, the average absolute pricing error reaches 0.30% per month. The increase in the pricing errors is thus equally due to the change in time period and the change in test asset.

Yogo's model seems, therefore, more robust than LL's model and has similar performance as both CV's model and FF's model with the new test assets. Note that this result on the inability of the conditional CAPM to explain the value premium complements the arguments of Lewellen and Nagel (2005), who show that the conditional CAPM performs nearly as poorly as the unconditional CAPM in terms of explaining the value premium. They argue that the covariance of the conditional expected return on the market and of the conditional market betas of value and growth stocks is not high enough to explain the value premium. The result for the LL's model is also consistent with the finding by Hodrick and Zhang (2001) of large specification errors for this model. It is also worth noting that I report the goodness of fit using only the Fama–McBeth results for the quarterly frequency models because the time series is relatively short (86 quarters). When using the SDF-GMM approach,

the models still fail to capture the cross-section of returns and the statistics vary a lot across the SDF-GMM methodologies described in Section 4.[12]

7. Conclusion

This paper shows that some prominent risk-based theories proposed to explain the value premium are not satisfactory empirically. Indeed, I find that these models capture the cross-section of the 25 size-BE/ME sorted portfolios but fail to capture the cross-section of returns of a different set of test assets. This paper can be seen as a data/model snooping check. The conclusion is that an alteration in test assets has a large effect on the magnitude of the pricing errors generated by these asset pricing models. Authors should, therefore, always experiment with different test assets when testing models that seek to explain the value premium or other empirical regularities.

Appendix A: Models not Mentioned in Daniel and Titman (2005) or in Lewellen et al. (2006)

Title: "Long-Run Stockholder Consumption Risk and Asset Returns"
Authors: Malloy, C., T. Moskowitz, and A. Vissing-Jörgensen, working paper
Claim: The covariance of returns with long-run consumption growth of stockholders captures the cross-sectional variation in stock returns, including the size and value premia.

Title: "Consumption, Dividends, and the Cross-Section of Equity Returns"
Authors: Bansal, R., R. F. Dittmar, and C. T. Lundblad
Published: *Journal of Finance*, August 2005
Claim: Aggregate consumption risks embodied in cash flows can account for the puzzling differences in risk premia across book-to-market, momentum, and size-sorted portfolios.

Title: "Idiosyncratic Consumption Risk and the Cross-Section of Asset Returns"
Authors: Jacobs, K. and K. Q. Wang
Published: *Journal of Finance*, October 2004
Claim: The presence of uninsurable idiosyncratic consumption risk is relevant for constructing factors that help explain the cross section of asset returns.

[12] Similar results are obtained with other models whose factors are available online (Vassalou and Xing, 2004, and Ferguson and Shockley, 2003). These results are available upon request.

Appendix B: Selected Models for Empirical Tests

Title: "A Consumption-Based Explanation of Expected Stock Returns"
Authors: Yogo, M.
Published: *Journal of Finance*, forthcoming
Claim: The model explains both the cross-sectional variation in expected stock returns and the time variation in the equity premium. Small stocks and value stocks deliver relatively low returns during recessions, when durable consumption falls, which explains their high average returns relative to big stocks and growth stocks.

Title: "Resurrecting the (C)CAPM: A cross-Sectional Test When Risk Premia Are Time Varying"
Authors: Lettau, M. and S. Ludvingson
Published: *Journal of Political Economics*, 2001
Claim: The (C)CAPM can account for the difference in returns between low-BE/ME and high-BE/ME portfolios and exhibits little evidence of residual BE/ME effect.

Title: "Bad Beta, Good Beta"
Authors: Campbell, J. and T. Vuolteenaho
Published: *American Economic Review*, December 2004
Claim: Explains the size and value anomalies in stock returns using an economically motivated two-beta model. The beta of a stock with the market portfolio is broken into two components: one reflecting news about the market's future cash flows and one reflecting news about the market's discount rates. Value stocks have considerably higher cash- flow betas than growth stocks and this can explain their higher average returns.

Title: "Common risk factors in the returns on stocks and Bonds"
Authors: Fama, E. and K. French
Published: *Journal of Financial Economics*, February 1993
Claim: Explains cross-section of returns with three factors (Rm-Rf, SMB and HML).

References

Amihud, Y. (2002) Illiquidity and stock returns: Cross-section and time-series effects, *Journal of Financial Markets* **5**, 31–56.

Ali, A., Hwang, L-S. and Trombley, M. (2003) Arbitrage risk and the book-to-market anomaly, *Journal of Financial Economics* **69**, 355–373.

Ball, R., Kothari, S. P. and Shanken, J. (1995) Problems in measuring portfolio performance: An application to contrarian investment strategies, *Journal of Financial Economics* **38**, 79–107.
Campbell, J. and Vuolteenaho, T. (2004) Bad beta, good beta, *American Economic Review* **94**, 1249–1275.
Chan, L., Jegadeesh, N. and Lakonishok, J. (1995) Evaluating the performance of value versus glamour stocks, *Journal of Financial Economics* **38**, 269–296.
Cochrane, J. (2005) *Asset Pricing*, Princeton University Press, revised edition.
Conrad, J., Cooper, M. and Kaul, G. (2003) Value versus glamour, *Journal of Finance* **59**, 1969–1995.
Daniel, K. and Titman, S. (2005) Testing factor-model explanations of market anomalies, unpublished working paper, Northwestern University and UT Austin.
Fama, E. and French, K. (1993) Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* **33**, 3–56.
Ferguson, M. and Shockley, R. (2003) Equilibrium "Anomalies", *Journal of Finance* **58**, 2549–2580.
Gibbons, M., Ross, S. and Shanken, J. (1989) A test of the efficiency of a given portfolio, *Econometrica* **57**, 1121–1152.
Gompers, P. and Metrick, A. (2001) Institutional investors and equity prices, *Quarterly Journal of Economics* **116**, 229–259.
Griffin, J. M. and Lemmon, M. (2002) Book-to-market equity, distress risk, and stock returns, *Journal of Finance* **57**, 2317–2336.
Hansen, L. P. and Jagannathan, R. (1997) Assessing specification errors in stochastic discount factor models, *Journal of Finance* **52**, 557–590.
Hodrick, R. J. and Zhang, X. (2001) Evaluating the specification errors of asset pricing models, *Journal of Financial Economics* **62**, 327–376.
Jagannathan, R. and Wang, Z. (1996) The CAPM is alive and well, *Journal of Finance* **51**, 3–53.
Kothari, S. P., Shanken, J., and Sloan, R. (1995) Another look at the cross-section of expected stock returns, *Journal of Finance* **50**, 185–224.
Lakonishok, J., Shleifer, A., and Vishny, R. (1994) Contrarian investment, extrapolation and risk, *Journal of Finance* **49**, 1541–1578.
Lettau, M. and Ludvingson, S. (2001) Resurrecting the (C)CAPM: A cross-sectional test when risk premia are time varying, *Journal of Political Economy* **109**, 1238–1287.
Lewellen, J. and Nagel, S. (2005) The conditional CAPM does not explain asset pricing anomalies, *Journal of Financial Economics* forthcoming.
Lewellen, J., Nagel, S. and Shanken, J. (2006) A skeptical appraisal of asset-pricing tests, NBER working paper 12360.
Lo, A. W. and MacKinlay, S. C. (1990) Data-snooping biases in tests of financial asset pricing models, *Review of Financial Studies* **3**, 431–468.
Nagel, S. (2005) Short sales, institutional investors, and the book-to-market effect, *Journal of Financial Economics* forthcoming.
Pontiff, J. (1996) Costly arbitrage: Evidence from closed-end funds, *Quarterly Journal of Economics* **111**, 1135–1151.
Shanken, J. (1992) On the estimation of Beta-pricing models, *Review of Financial Studies* **5**, 1–55.
Shumway, T. (1997) The delisting bias in CRSP data, *Journal of Finance* **52**, 327–340.
Vassalou, M. and Xing, Y. (2004) Default risk in equity returns, *Journal of Finance* **20**, 831–868.
Yogo, M. (2006) A consumption-based explanation of expected stock returns, *Journal of Finance* **61**, 539–580.