MANAGEMENT SCIENCE
Vol. 59, No. 1, January 2013, pp. 122–141
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

http://dx.doi.org/10.1287/mnsc.1120.1557

Intertemporal CAPM with Conditioning Variables

Paulo Maio
Hanken School of Economics, 00101 Helsinki, Finland, paulmaio@hanken.fi

This paper derives and tests an intertemporal capital asset pricing model (ICAPM) based on a conditional version of the Campbell–Vuolteenaho two-beta ICAPM (bad beta, good beta (BBGB)). The novel factor is a scaled cash-flow factor that results from the interaction between cash-flow news and a lagged state variable (market dividend yield or consumer price index inflation). The cross-sectional tests over 10 portfolios sorted on size, 10 portfolios sorted on book-to-market, and 10 portfolios sorted on momentum show that the scaled ICAPM explains relatively well the dispersion in excess returns on the 30 portfolios. The results for an alternative set of equity portfolios (25 portfolios sorted on size and momentum) show that the scaled ICAPM prices particularly well the momentum portfolios. Moreover, the scaled ICAPM compares favorably with alternative asset pricing models in pricing both sets of equity portfolios. The scaled factor is decisive to account for the dispersion in average excess returns between past winner and past loser stocks. More specifically, past winners are riskier than past losers in times of high price of risk. Therefore, a time-varying cash-flow beta/price of risk provides a rational explanation for momentum.

Key words: asset pricing models; conditional CAPM; ICAPM; linear multifactor models; predictability of returns; cross-section of stock returns; time-varying risk aversion; momentum; value premium
History: Received April 25, 2010; accepted February 28, 2012, by Wei Xiong, finance. Published online in *Articles in Advance* August 20, 2012.

1. Introduction

The Merton (1973) intertemporal capital asset pricing model (ICAPM) postulates that state variables that predict market returns should act as risk factors to help price the cross-section of average equity returns. Among the papers that implement empirically testable versions of the original ICAPM are Campbell (1993, 1996) and, more recently, Chen (2003), Brennan et al. (2004), Guo (2006), and Campbell and Vuolteenaho (2004). Common to these papers is the assumption that the factor betas/risk prices in the expected return-beta representation are constant through time. Nevertheless, following the conditional CAPM literature it seems reasonable to assume that (some of) the factor betas/risk prices are time varying.

This paper derives an ICAPM based on a conditional version of the two-beta ICAPM (bad beta, good beta (BBGB)) from Campbell and Vuolteenaho (2004). The beta/price of risk of aggregate cash-flow news is assumed to be time varying, similarly to the conditional CAPM literature regarding the market beta/price of risk (Harvey 1989, Ferson and Harvey 1999, Lettau and Ludvigson 2001, and Petkova and Zhang 2005, among others). Specifically, the conditional cash-flow beta is assumed to be linear in a state variable, leading to a scaled ICAPM that contains three factors: revisions in future aggregate cash flows (cash-flow news), revisions in future market discount rates (discount-rate news), and a scaled factor that corresponds to the interaction of cash-flow news and the lagged state variable. The novelty relative to the BBGB model is this scaled factor that accounts for a time-varying cash-flow beta. Alternatively, the scaled ICAPM can be derived by assuming that the cash-flow risk price is a linear function of the state variable, which leads to two distinct fundamental components of the price of risk—a constant or long-term component of the risk price and a time-varying component that is negatively correlated with the business cycle or, more generally, positively correlated with "bad times," and thus is interpreted as the cyclical or short-term component of the cash-flow risk price (alternatively, risk aversion).

In the empirical implementation of the scaled ICAPM, I use the market dividend yield, which is a popular predictor of the market equity premium, and is also a countercyclical variable. The second conditioning variable employed is the consumer price index (CPI) inflation, which is used by Brandt and Wang (2003) as a determinant of risk aversion. I conjecture that times of rising inflation are "bad times" in the sense that positive shocks in inflation lead to negative shocks in real financial and nonfinancial wealth and thus a higher price of risk.

The cross-sectional tests with 10 portfolios sorted on size, 10 portfolios sorted on book-to-market

(BM), and 10 portfolios sorted on momentum (S10 + BM10 + M10) show that the scaled ICAPM explains relatively well the dispersion in excess returns on the 30 portfolios, with explanatory ratios varying between 47% and 60%. Thus, the three-factor model accounts for the value and momentum anomalies. The results for an alternative set of equity portfolios (25 portfolios sorted on size and momentum) show that the scaled ICAPM prices particularly well the momentum portfolios, with explanatory ratios varying between 74% and 86%. Furthermore, the risk price estimates show that the risk price for the scaled factor is significantly positive, indicating that the conditional beta/risk price of cash-flow news increases with the state variables; that is, the beta/price of risk rises in "bad times." The results show that the time-varying or cyclical component of the cash-flow risk price strongly dominates the constant or average component. Moreover, the scaled ICAPM compares favorably with alternative asset pricing models in pricing both sets of equity portfolios—the BBGB model; the unconditional CAPM (Sharpe 1964 and Lintner 1965); the conditional CAPM (Jagannathan and Wang 1996, Ferson and Harvey 1999, Lettau and Ludvigson 2001, among others); the Fama and French (1993) three-factor model; and the Carhart (1997) four-factor model. Among the competing models, only the Carhart (1997) model beats the scaled ICAPM.

The explanatory power of the model is robust to a series of robustness checks: using the factors as additional test assets; adding bond premia to the set of test assets; using alternative vector autoregression (VAR) specifications to estimate cash-flow and discount-rate news; including an intercept in the pricing equation; allowing the beta (or, alternatively, the price of risk) of discount-rate news to also be time varying; testing the model with 25 size-BM portfolios; using alternative measurement of the lagged state variables; conducting a bootstrap-based inference; or estimating the model in expected return-beta form. I also test the scaled ICAPM with alternative portfolio groups that combine momentum with several other stock characteristics (stock volatility, firm age, idiosyncratic volatility, volume, and credit risk).

The fact that the scaled ICAPM is able to account for momentum profits represents an important innovation to most of the empirical tests of the cross-section of stock returns found in the literature.[1] The scaled factor is decisive to account for the dispersion in average excess returns between past winner and past loser stocks. More specifically, past losers are riskier than past winners are in periods of low realizations of the state variable, that is, in times with a low price of risk. Therefore, a time-varying cash-flow beta/price of risk provides a rational explanation for momentum.

This paper is organized as follows. Section 2 presents the theoretical derivation of the scaled ICAPM. Section 3 presents the main empirical cross-sectional tests of the scaled ICAPM. Section 4 presents several robustness checks to the main results in the previous section. Section 5 provides additional discussion on why the scaled ICAPM prices the momentum portfolios; time-varying risk aversion as an alternative source of the time-varying cash-flow risk price is also discussed. Section 6 concludes.

2. A Scaled ICAPM

Campbell (1993) uses an Epstein and Zin (1989, 1991) utility function and a decomposition for innovations on consumption growth based on the investor's intertemporal budget constraint (combined with the assumptions of joint conditional log-normality and homoskedasticity for both asset returns and consumption growth) to derive a discrete-time version of the ICAPM:

$$\mathrm{E}_t(r_{i,t+1}) - r_{f,t+1} + \frac{\sigma^2_{i,t}}{2} = \gamma \sigma_{im,t} + (\gamma - 1)\sigma_{iH,t}, \quad (1)$$

where $\mathrm{E}_t(\cdot)$ denotes the conditional expectation at time t; $r_{i,t+1}$ and $r_{f,t+1}$ denote the log return on stock i and log risk-free rate, respectively; γ is the coefficient of relative risk aversion (RRA); $\sigma^2_{i,t}/2$ is a Jensen's inequality adjustment arising from the log-normal model, with $\sigma^2_{i,t} \equiv \mathrm{Var}_t(r_{i,t+1})$ denoting the conditional variance of stock i's return; and $\sigma_{im,t} \equiv \mathrm{Cov}_t[r_{i,t+1}, r_{m,t+1} - \mathrm{E}_t(r_{m,t+1})]$ and $\sigma_{iH,t} \equiv \mathrm{Cov}_t(r_{i,t+1}, r^H_{t+1})$ represent the conditional covariances of stock i's return with the current market return and news about future market returns, respectively. News about future market returns (discount-rate news) is given by

$$r^H_{t+1} \equiv (\mathrm{E}_{t+1} - \mathrm{E}_t)\sum_{j=1}^{\infty} \rho^j r_{m,t+1+j}. \quad (2)$$

The intuition behind the model in Equation (1) is that for an investor with greater risk aversion than an investor with log utility (i.e, $\gamma > 1$), assets that are positively correlated with future market returns (changes in future investment opportunities) have a higher expected return than predicted by the CAPM (the first factor). The reason is that such assets do not enable an investor to hedge against a deterioration in future investment opportunities.

By using the same framework as in Campbell (1993, 1996), Campbell and Vuolteenaho (2004; CV04 hereafter) rely on the dynamic accounting identity developed by Campbell and Shiller (1988) and

[1] Bansal et al. (2005) and Liu and Zhang (2008) test asset pricing models that also account for momentum.

Campbell (1991), which decomposes current unexpected market returns into discount-rate news, and the residual, which they interpret as cash-flow news:

$$
\begin{aligned}
r_{m,t+1} - \mathrm{E}_t(r_{m,t+1}) &= (\mathrm{E}_{t+1} - \mathrm{E}_t)\sum_{j=0}^{\infty}\rho^j \Delta d_{t+1+j} \\
&\quad - (\mathrm{E}_{t+1} - \mathrm{E}_t)\sum_{j=1}^{\infty}\rho^j r_{m,t+1+j} \equiv r^{CF}_{t+1} - r^{H}_{t+1}, \qquad (3)
\end{aligned}
$$

where $r^{CF}_{t+1} \equiv (\mathrm{E}_{t+1} - \mathrm{E}_t)\sum_{j=0}^{\infty}\rho^j \Delta d_{t+1+j} = r_{m,t+1} - \mathrm{E}_t(r_{m,t+1}) + r^{H}_{t+1}$ represents "news" about future cash flows. CV04 derive a version of the ICAPM based also on only two factors: the covariance (beta) with discount-rate news (good beta) and the covariance with cash-flow news (bad beta):

$$
\mathrm{E}_t(r_{i,t+1}) - r_{f,t+1} + \frac{\sigma^2_{i,t}}{2} = \gamma\,\sigma_{iCF,t} - \sigma_{iH,t}, \qquad (4)
$$

where $\sigma_{iCF,t} \equiv \mathrm{Cov}_t(r_{i,t+1}, r^{CF}_{t+1})$ is the covariance of asset i's return with cash-flow news.[2] The covariance with cash-flow news receives a risk price of γ, whereas the covariance with discount-rate news has a risk price of -1. Thus, for a conservative investor (one with $\gamma > 1$), the covariance with upward revisions in future cash flows has a higher risk price than the covariance with downward revisions in future market returns. The intuition provided by CV04 relies on the argument that adverse changes in future cash flows have a permanent negative impact on wealth, whereas adverse shocks (unexpected rises) in future discount rates have only a transitory negative impact on wealth because although current wealth declines it will be reinvested at a higher rate of return.

Similarly to CV04, one can extend the model in Equation (4) to allow for the possibility that the log return on the reference portfolio is a weighted average of the log stock market return and the log risk-free rate:

$$
r_{p,t+1} = \omega r_{m,t+1} + (1-\omega) r_{f,t+1}, \qquad (5)
$$

where ω stands for the weight associated with the stock index in the global portfolio. Given that $(\mathrm{E}_{t+1} - \mathrm{E}_t) r_{f,t+1} = 0$ (i.e., the risk-free rate is known at the beginning of the period) and by assuming that $r_{f,t+1}$ is approximately constant through time, then the discount-rate and cash-flow news associated with the aggregate return in Equation (5) are equal to ωr^{H}_{t+1} and ωr^{CF}_{t+1}, respectively. This leads to an "unrestricted" version of the BBGB model:

$$
\mathrm{E}_t(r_{i,t+1}) - r_{f,t+1} + \frac{\sigma^2_{i,t}}{2} = \gamma\omega\,\sigma_{iCF,t} - \omega\,\sigma_{iH,t}. \qquad (6)
$$

The main innovation relative to Equation (4) is that the risk price associated with $\sigma_{i,H}$ is estimated in the cross-section instead of being fixed at -1. By imposing $\omega = 1$ (i.e., the reference portfolio only contains a stock index), one obtains the restricted version (4) as a special case of (6). According to Equation (6) the cash-flow risk price should be positive, whereas the risk price for discount-rate news should be estimated negatively.

Next, to derive the scaled ICAPM, I assume that the conditional covariance with cash-flow news has the following form:

$$
\begin{aligned}
\mathrm{Cov}_t(r_{i,t+1}, r^{CF}_{t+1}) &= \lambda_t \mathrm{Cov}(r_{i,t+1}, r^{CF}_{t+1}) \\
&\quad + \rho_t \mathrm{Cov}(r_{i,t+1}, r^{CF}_{t+1} z_t), \qquad (7)
\end{aligned}
$$

where z_t is a time t state variable and both λ_t and ρ_t denote time t variables that depend on z_t and the conditional variance, $\mathrm{Var}_t(r^{CF}_{t+1})$. This functional form for the conditional covariance is equivalent to the conditional beta from conditional market regressions: $\beta_t(r_{i,t+1}, r^{CF}_{t+1}) = \beta(r_{i,t+1}, r^{CF}_{t+1}) + \beta(r_{i,t+1}, r^{CF}_{t+1} z_t) z_t$, where $\beta_t(r_{i,t+1}, r^{CF}_{t+1}) \equiv \mathrm{Cov}_t(r_{i,t+1}, r^{CF}_{t+1})/\mathrm{Var}_t(r^{CF}_{t+1})$ denotes the conditional cash-flow beta; $\beta(r_{i,t+1}, r^{CF}_{t+1}) \equiv \mathrm{Cov}(r_{i,t+1}, r^{CF}_{t+1})/\mathrm{Var}(r^{CF}_{t+1})$ is the unconditional cash-flow beta; and $\beta(r_{i,t+1}, r^{CF}_{t+1} z_t) \equiv \mathrm{Cov}(r_{i,t+1}, r^{CF}_{t+1} z_t)/\mathrm{Var}(r^{CF}_{t+1} z_t)$ denotes the unconditional beta with the scaled factor.[3] By substituting (7) on (6) and using the law of iterated expectations, I obtain a three-factor model in unconditional form:

$$
\mathrm{E}(r_{i,t+1}) - r_{f,t+1} + \frac{\sigma^2_i}{2} = \gamma_{CF}\,\sigma_{iCF} + \gamma_{CFz}\,\sigma_{iCFz} + \gamma_H\,\sigma_{iH}, \qquad (8)
$$

where $\gamma_{CF} \equiv \gamma\omega \mathrm{E}(\lambda_t)$, $\gamma_{CFz} \equiv \gamma\omega \mathrm{E}(\rho_t)$, $\gamma_H \equiv -\omega$, and all the covariances are unconditional. In the restricted version of the scaled ICAPM ($\omega = 1$), the risk prices are equal to $\gamma_{CF} \equiv \gamma \mathrm{E}(\lambda_t)$, $\gamma_{CFz} \equiv \gamma \mathrm{E}(\rho_t)$, and $\gamma_H = -1$.

Alternatively, one can derive the three-factor scaled ICAPM by allowing the cash-flow risk price to be time varying, making it related with z_t, as in Lettau and Ludvigson (2001), Petkova and Zhang (2005), or Zhang (2005):

$$
\gamma_t = \gamma_0 + \gamma_1 z_t. \qquad (9)
$$

By imposing (9) in the pricing Equations (4) and (6) and using the linearity of conditional expectations and the law of iterated expectations, one obtains:

$$
\mathrm{E}(r_{i,t+1}) - r_{f,t+1} + \frac{\sigma^2_i}{2} = \gamma_0\,\sigma_{iCF} + \gamma_1\,\sigma_{iCFz} - \sigma_{iH}, \qquad (10)
$$

$$
\mathrm{E}(r_{i,t+1}) - r_{f,t+1} + \frac{\sigma^2_i}{2} = \gamma_0\omega\,\sigma_{iCF} + \gamma_1\omega\,\sigma_{iCFz} - \omega\,\sigma_{iH}. \qquad (11)
$$

[2] In (4) the difference from CV04 is that $\sigma_{iH,t}$ appear with a minus sign in the pricing equation because they define the covariance (beta) with respect to the negative (favorable change) of discount-rate news.

[3] This specification is widely used in the conditional CAPM literature (see Shanken 1990, Lewellen 1999, Petkova and Zhang 2005, among others).

By using excess simple returns $R_{i,t+1} - R_{f,t+1}$ on the left-hand side (as in CV04), one obtains the unconditional pricing equations of the scaled ICAPM:

$$E(R_{i,t+1} - R_{f,t+1}) = \gamma_{CF}\beta_{i,CF} + \gamma_{CFz}\beta_{i,CFz} - \beta_{i,H}, \quad (12)$$

$$E(R_{i,t+1} - R_{f,t+1}) = \gamma_{CF}\beta_{i,CF} + \gamma_{CFz}\beta_{i,CFz} + \gamma_{H}\beta_{i,H}. \quad (13)$$

Henceforth, Equation (13) is denoted as the benchmark scaled ICAPM, whereas Equation (12) corresponds to the restricted scaled ICAPM. Similarly, the two versions of the BBGB model in unconditional form are defined as

$$E(R_{i,t+1} - R_{f,t+1}) = \gamma_{CF}\beta_{i,CF} - \beta_{i,H}, \quad (14)$$

$$E(R_{i,t+1} - R_{f,t+1}) = \gamma_{CF}\beta_{i,CF} + \gamma_{H}\beta_{i,H}. \quad (15)$$

Hereafter, Equations (14) and (15) are denoted as the restricted and benchmark bad beta good beta ICAPM, respectively. These four pricing equations will be empirically tested in the next section.

3. Asset Pricing Tests

3.1. Data

I use two sets of equity portfolios in the cross-sectional tests of the different ICAPM specifications. The first group contains 30 portfolios (S10 + BM10 + M10), which include 10 portfolios sorted on size (denoted as S10); 10 portfolios sorted on book-to-market; and 10 portfolios sorted on momentum (prior one-year returns, denoted as M10). The second group stands for 25 portfolios sorted on both size and momentum (SM25). All portfolio return data and the one-month Treasury bill rate, used to calculate excess returns, are obtained from Kenneth French's website. Return data on the value-weighted stock market index are from the Center for Research in Security Prices (CRSP), whereas monthly data on the price index, earnings, and dividends associated with the Standard & Poor's (S&P) Composite Index are obtained from Robert Shiller's website. Macroeconomic and additional interest rate data, including the seasonally adjusted CPI, 10-year and 1-year Treasury bond yields, and the three-month Treasury bill rate are all obtained from the FRED database, available from the St. Louis FED's website.

3.2. Estimating the "Shifts in the Investment Opportunity Set": A VAR Approach

Following Campbell (1991, 1993), I rely on a first-order VAR to estimate both r^{H}_{t+1} and r^{CF}_{t+1}. The VAR equation assumed to govern the behavior of a state vector $\mathbf{x}_t$, which includes the excess market return and other variables that help to forecast changes in the expected market return, is given by[4]

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \boldsymbol{\epsilon}_{t+1}. \quad (16)$$

In this framework the news components are estimated in the following way:

$$r^{H}_{t+1} \equiv (\mathrm{E}_{t+1} - \mathrm{E}_t)\sum_{j=1}^{\infty}\rho^j r_{m,t+1+j} = \mathbf{e1}'\rho\mathbf{A}(\mathbf{I} - \rho\mathbf{A})^{-1}\boldsymbol{\epsilon}_{t+1} = \boldsymbol{\varphi}'\boldsymbol{\epsilon}_{t+1}, \quad (17)$$

$$r^{CF}_{t+1} \equiv (\mathrm{E}_{t+1} - \mathrm{E}_t)\sum_{j=0}^{\infty}\rho^j \Delta d_{t+1+j} = r_{m,t+1} - \mathrm{E}_t(r_{m,t+1}) + r^{H}_{t+1} = [\mathbf{e1}' + \mathbf{e1}'\rho\mathbf{A}(\mathbf{I} - \rho\mathbf{A})^{-1}]\boldsymbol{\epsilon}_{t+1} = (\mathbf{e1} + \boldsymbol{\varphi})'\boldsymbol{\epsilon}_{t+1}. \quad (18)$$

Here, ρ is a discount coefficient linked to the average log consumption-to-wealth ratio, $\rho \equiv 1 - \exp[E(c_t - w_t)]$, or average market dividend-to-price ratio; $\mathbf{e1}$ is an indicator vector that takes a value of 1 in the cell corresponding to the position of the excess market return in the VAR; $\mathbf{A}$ is the VAR coefficient matrix; $\mathbf{I}$ is an identity matrix; and $\boldsymbol{\varphi}' \equiv \mathbf{e1}'\rho\mathbf{A}(\mathbf{I} - \rho\mathbf{A})^{-1}$ is the function that relates the VAR shocks with discount-rate news. To be consistent with previous work (e.g., CV04), I assume $\rho = 0.95^{1/12}$, i.e., an annualized discount factor of 0.95, corresponding to an annual consumption-to-wealth ratio of approximately 5%. In Equation (18) cash-flow news is the residual component of unexpected stock market returns, which has the advantage that one does not have to model directly the dynamics of aggregate dividends.[5]

The state-vector associated with the first-order VAR is given by $\mathbf{x}_t \equiv [RREL_t, TERM_t, VS_t, EY_t, r_{mt}]'$, which represents a parsimonious representation for the variables that forecast the market return and follows the specification used in CV04. *RREL* is the relative or detrended Treasury bill rate, representing the difference between the three-month Treasury bill rate (r_{f3}) and a moving average of r_{f3} over the previous 12 months.[6] Its inclusion in the VAR is justified by previous evidence that short-term interest rates forecast expected market returns, at least for short-term forecasting horizons (Campbell 1991, Hodrick 1992, Ang and Bekaert 2007, among others). *TERM* refers to the term structure spread (measured here as the difference between the 10-year and 1-year Treasury bond yields), which represents a proxy for the yield

[4] Any P-order VAR with $P > 1$ can be restated as a first-order VAR if the state vector is expanded by including lagged state variables, with $\mathbf{A}$ denoting the VAR companion matrix.

[5] This avoids problems like the seasonality and nonstationarity of dividends.

[6] $RREL_t = r_{f3,t} - (1/12)\sum_{j=1}^{12} r_{f3,t-j}$.

Table 1 VAR Estimation Results

	$RREL_t$	$TERM_t$	VS_t	EY	$r_{m,t}$	$\bar{R}^2$
$RREL_{t+1}$	0.95	0.07	−0.01	−0.00	0.00	0.83
	(**30.37**)	(**3.32**)	(−**2.70**)	(−0.00)	(0.73)	
$TERM_{t+1}$	−0.04	0.94	0.01	−0.00	0.00	0.94
	(<u>−2.41</u>)	(**76.00**)	(*1.83*)	(−*1.51*)	(0.01)	
VS_{t+1}	0.43	0.23	0.22	−0.00	0.07	0.05
	(**2.99**)	(1.50)	(**3.18**)	(−0.92)	(<u>2.18</u>)	
EY_{t+1}	2.03	1.16	0.23	0.99	−0.16	0.99
	(**7.58**)	(**4.22**)	(<u>2.02</u>)	(**96.78**)	(−1.45)	
$r_{m,t+1}$	−0.20	0.39	−0.06	0.01	0.08	0.02
	(−0.96)	(*1.80*)	(−0.89)	(*1.68*)	(1.40)	

Notes. This table presents the estimated coefficients (first row of each variable) and associated heteroskedasticity-robust t-statistics (second row) for a first-order VAR. The VAR vector is given by $[RREL_t, TERM_t, VS_t, EY_t, r_{mt}]'$, where *RREL* is the relative bill rate, *TERM* is the term structure spread, VS is the value spread, EY is the log earnings yield, and r_{mt} is the excess log value-weighted market return. The original sample is 1953:05–2009:12. Italic, underlined, and bold t-statistics denote statistical significance at the 10%, 5%, and 1% levels, respectively. $\bar{R}^2$ is the adjusted R^2.

curve slope. This variable has been widely used in the predictability of returns literature since Campbell (1987) and Fama and French (1989) have found that *TERM* tracks the business cycle and predicts market returns. VS is the small-value spread, which is computed as the spread in log returns of the small-high BM and small-low BM portfolios.[7] EY denotes the log earnings yield, calculated as the log of the earnings-to-price ratio associated with the S&P 500 index. This variable is used instead of the market dividend yield (Fama and French 1988) in light of recent evidence that the forecasting power of the market dividend yield has decreased since the 1990s and also to be consistent with CV04.[8] The fifth variable used in the VAR is the excess log market return (r_m), which takes as proxy the value-weighted market return from CRSP. The sample used in estimating the VAR is 1953:05–2009:12, because the Treasury bond yield data are not available before 1953:04.

Untabulated results show that both *TERM* and EY are quite persistent, with autoregressive coefficients (”) of 0.97 and 0.99, respectively, whereas *RREL* is also relatively persistent (” = 0091). Furthermore, the VAR state variables are not highly correlated with the exception of *RREL* and *TERM* with a correlation coefficient of −0.58 and, to a lower degree, *TERM* and EY (−0.43) and VS and r_m (−0.41).

The VAR coefficient estimates and associated heteroskedasticity-robust t-statistics (White 1980) are presented in Table 1. The market return equation, presented in the bottom row, shows that all four state variables forecast the excess market return with the expected sign: *RREL* predicts negative excess market returns one month ahead, consistent with previous evidence (Campbell 1991, Hodrick 1992, Ang and Bekaert 2007), and both *TERM* and EY predict positive market returns, also in line with previous evidence (Fama and French 1988, 1989; CV04). On the other hand, VS predicts negative changes in the aggregate equity premium, which is in line with the evidence in CV04 for a longer sample. Both *RREL* and VS are not statistically significant, whereas *TERM* and EY are significant at the 10% level. The small individual significance of the state variables might be a result of some multicollinearity induced by the correlations among some of the state variables as referred above. The small degree of one-month momentum in market returns, as captured by the estimated slope for r_{m1t}, is not statistically significant. The adjusted R^2 of 2% is in line with the values for monthly predictive regressions found in the predictability literature. A Wald test of the joint significance of the five predictors for the excess stock return yields a p-value of 0.00; that is, the predictors are jointly significant at forecasting the equity premium.[9] The other equations in the VAR show that *TERM* is close to an AR(1) process, but it

[7] A similar spread is used by CV04. The small-high and small-low portfolios belong to the group of six portfolios sorted on size and BM, available from Kenneth French's library.

[8] Instead of the annual earnings yield, CV04 use the smoothed price-earnings ratio, which is based on a 10-year moving average of past earnings. I have opted to use the annual earnings yield because in unreported results the forecasting power of EY for the market return at the one-month horizon is greater than the corresponding predictive power associated with the smoothed earnings yield.

[9] Preliminary results show that there is a decline in the predictability of the equity premium in recent years. By estimating the VAR over the 1953.05–2003.12 period, the adjusted R^2 is slightly higher (3%), whereas the slopes associated with $RREL$ and EY are significant at the 10% and 5% levels, respectively. If one uses one-sided t-stats (justified because the signs of the slopes are more or less consistent and have an economic justification, e.g., EY), all four predictors are individually significant over this shorter sample.

is also (negatively) predicted by *RREL*, whereas *RREL* is mainly explained by its lagged value, although both *TERM* and VS have also some forecasting power. Regarding the value spread, it is (positively) predicted by its lagged value, *RREL*, and the market return. In the equation for EY, the autoregressive coefficient is highly significant, and in addition, *RREL*, *TERM*, and VS all forecast an increase in EY.

The results for the variance decomposition associated with the components of the excess market return are given by

$$\frac{\mathrm{Var}(r^{CF}_{t+1})}{\mathrm{Var}[r_{m\,t+1} - \mathrm{E}_t(r_{m\,t+1})]} + \frac{\mathrm{Var}(r^{H}_{t+1})}{\mathrm{Var}[r_{m\,t+1} - \mathrm{E}_t(r_{m\,t+1})]} - \frac{2\,\mathrm{Cov}(r^{CF}_{t+1}, r^{H}_{t+1})}{\mathrm{Var}[r_{m\,t+1} - \mathrm{E}_t(r_{m\,t+1})]} = 158.33\% + 67.99\% - 126.31\% = 100\%. \quad (19)$$

Thus, cash-flow news represents more than twice the weight of discount-rate news over the excess market return. These results are in contrast with previous findings showing that the major component of aggregate (unexpected) stock returns is discount-rate news followed by cash-flow news (Campbell 1991, Campbell and Ammer 1993, Campbell and Vuolteenaho 2004, among others). The difference in results should be attributable to a declining forecasting power for market returns in recent years, which leads to a decrease in the weight of discount-rate news.[10] The absolute weights attached to cash-flow news and the covariance term are above 100%, because both components of market returns are positively correlated (0.61).

3.3. Econometric Framework

The asset pricing models presented in the previous section are estimated by a first-stage generalized method of moments (GMM) procedure (Hansen 1982, Cochrane 2005) in which all the moment conditions (that is, the pricing equations for the different test assets) receive the same weight. This procedure is equivalent to an ordinary least-squares (OLS) cross-sectional regression of average excess returns on factor covariances (betas) and enables us to assess whether the scaled ICAPM can explain the returns of a set of economically interesting portfolios (e.g., value or momentum portfolios). The first N sample moments correspond to the pricing errors for each of the N test assets:

$$g_T(\mathbf{b}) \equiv \frac{1}{T}\sum_{t=0}^{T-1} \begin{cases} (R_{i\,t+1} - R_{f\,t+1}) - \gamma_{CF}\, r_{i\,t+1}(r^{CF}_{t+1} - \mu_{CF}) \\ \quad - \gamma_{CFz}\, r_{i\,t+1}(r^{CF}_{t+1} z_t - \mu_{CFz}) \\ \quad - \gamma_{H}\, r_{i\,t+1}(r^{H}_{t+1} - \mu_{H}) \\ r^{CF}_{t+1} - \mu_{CF} \\ r^{CF}_{t+1} z_t - \mu_{CFz} \\ r^{H}_{t+1} - \mu_{H} \end{cases} = \mathbf{0}, \quad i = 1, \ldots, N, \quad (20)$$

where $(\mu_{CF}, \mu_{CFz}, \mu_H)$ denote the means of $(r^{CF}_{t+1}, r^{CF}_{t+1} z_t, r^{H}_{t+1})$. The last three moment conditions in the system above enable estimating the factor means, which implies that the standard errors of the estimated covariance risk prices (γs) account for the error-in-variables associated with the factor means, as in Cochrane (2005, Chap. 13), and Yogo (2006).[11] In the case of the benchmark scaled ICAPM, there are three risk prices to estimate, γ_{CF}, γ_{CFz}, γ_H, and thus $N-3$ overidentifying conditions, whereas the benchmark BBGB model has $N-2$ overidentifying conditions in the respective GMM system.

The asymptotic test that the pricing errors are jointly equal to zero, with $\hat{\boldsymbol{\alpha}} \equiv g_T(\hat{\mathbf{b}})$, is given by

$$\hat{\boldsymbol{\alpha}}'\, \mathrm{Var}(\hat{\boldsymbol{\alpha}})^{+}\, \hat{\boldsymbol{\alpha}} \sim \chi^2(N-K), \quad (21)$$

where K is the number of factors and $\mathrm{Var}(\hat{\boldsymbol{\alpha}})^{+}$ denotes a generalized inverse of the variance–covariance matrix of the N pricing errors.[12] The standard errors for the parameter estimates and pricing errors are computed as in Cochrane (2005, Chap. 11).[13]

Two simpler and more robust measures for the global fit of a given model within the cross-section of returns are the average pricing error (mean absolute pricing error, MAE) and the cross-sectional OLS coefficient of determination. MAE is represented by

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N} |\hat{\alpha}_i|, \quad (22)$$

[10] In a VAR estimated over the 1953.05–2007.12 period, the weights associated with $\mathrm{Var}(r^{CF}_{t+1})$, $\mathrm{Var}(r^{H}_{t+1})$, and $-2\,\mathrm{Cov}(r^{CF}_{t+1}, r^{H}_{t+1})$ are 98%, 104%, and −103%, respectively, which shows that the weights associated with discount-rate and cash-flow news are approximate. The results showing a large weight for cash-flow news are also consistent with the evidence from Campbell et al. (2011) that the recent bear market (2007–2009) is mostly attributable to negative cash-flow news rather than positive shocks in market discount rates.

[11] This method enables estimating the risk prices in only one stage, instead of first estimating the factor covariances (betas) and the factor risk prices in a second stage as in the time-series/cross-sectional regressions approach presented in Cochrane (2005, Chap. 12).

[12] Following Cochrane (1996, 2005), I perform an eigenvalue decomposition of the moments' variance–covariance matrix, $\mathrm{Var}(\hat{\boldsymbol{\alpha}}) = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}'$, where $\mathbf{Q}$ is a matrix containing the eigenvectors of $\mathrm{Var}(\hat{\boldsymbol{\alpha}})$ on its columns and $\boldsymbol{\Lambda}$ is a diagonal matrix of eigenvalues; then only the nonzero eigenvalues of $\boldsymbol{\Lambda}$ are inverted.

[13] As discussed above, the signs of the risk prices are a priori restricted by the model; thus, I use t-statistics with one-sided p-values in evaluating the individual statistical significance.

and the cross-sectional OLS R^2 is equal to

$$R^2_{\mathrm{OLS}} = 1 - \frac{\sum_{i=1}^{N} \bar{\hat{\epsilon}}_i^2}{\sum_{i=1}^{N} \bar{R}_i^2}, \tag{23}$$

where

$$\bar{R}_i = \frac{1}{T}\sum_{t=0}^{T-1}(R_{i,t+1} - R_{f,t+1}) - \frac{1}{N}\sum_{i=1}^{N}\left\{\frac{1}{T}\sum_{t=0}^{T-1}(R_{i,t+1} - R_{f,t+1})\right\}$$

denotes the (cross-sectionally) demeaned (average) excess returns, $\hat{\epsilon}_i$ represents the pricing errors, and $\bar{\hat{\epsilon}}_i$ stands for the (cross-sectionally) demeaned pricing errors. R^2_{OLS} assigns equal weight to all pricing errors and represents a proxy for the proportion of the cross-sectional variance of average excess returns explained by the factors associated with a given model.

3.4. Benchmark ICAPM: Estimating the Factor Risk Premia

In the cross-sectional tests the first group of test portfolios (S10 + BM10 + M10) is linked to the most important CAPM anomalies—size premium, value premium, and momentum. Therefore, if a given asset pricing model is to explain the cross-section of equity returns, then it is reasonable to force that model to price all three portfolio sorts simultaneously. The second group is 25 portfolios sorted on size and momentum, which enables assessing in greater detail the explanatory power of the scaled ICAPM for the momentum anomaly (Jegadeesh and Titman 1993).

In the empirical implementation of the scaled ICAPM, I use two alternative state variables that drive the time-varying cash-flow beta or cash-flow risk price, f_t. The first state variable is the market dividend yield (DY), computed as the ratio of the annual sum of dividends to the monthly price level of the S&P 500 index. The market dividend yield was one of the first variables used to forecast market returns at several horizons (see Fama and French 1988, 1989), being interpreted as a variable related with the longer-term component of business conditions, which makes it eligible to explain the time-varying cash-flow beta/risk price.[14] However, the predictive power of DY has declined in recent years (see Cochrane 2005, among others). On the other hand, the market dividend yield can be interpreted as a proxy for "bad times" (in which the cash-flow price of risk is high) because it represents a countercyclical variable. By using a business cycle dummy (*CYCLE*)—which takes the value one in an economic expansion as defined by the National Bureau of Economic Research (NBER) and takes the value zero in recessions—and performing a monthly regression of DY on *CYCLE*, one gets the following results (OLS t-statistics in parenthesis):

$$DY_t = \underset{(38.32)}{0.040} - \underset{(-8.20)}{0.009}\, CYCLE_t, \quad \bar{R}^2 = 0.091,$$

which shows that DY represents an (imperfect) proxy for the business cycle. An alternative interpretation is that DY is negatively correlated with either stock prices or financial wealth (because this ratio tends to be high when stock prices and hence, financial wealth, are low) and consequently represents a proxy for "bad times" (times in which financial wealth is at lower levels).

The second state variable is the CPI inflation rate, computed as the log (one-year) difference on the CPI index, $\ln(CPI_t/CPI_{t-12})$, which is denoted as *CPI*. One justification for using *CPI* as a state variable that drives conditional cash-flow betas or the time-varying cash-flow risk price follows from evidence showing that inflation forecasts aggregate stock returns (Fama and Schwert 1977). Another possible motivation for using *CPI* as a scaling variable is that unexpected rises in inflation lead to negative shocks in both real financial and nonfinancial wealth, and thus *CPI* represents a proxy for "bad times." In a related cross-sectional study with both bonds and portfolios of stocks, Brandt and Wang (2003) specify risk aversion as a function of shocks to inflation.

Therefore, in the specification for the cash-flow price of risk, $f_t = f_0 + f_1 z_t$, with $z_t = DY_t, CPI_t$, one expects f_1 (and thus f_{CFz}) to be positive in both cases; i.e., worsening business conditions or deteriorating (real) financial and nonfinancial wealth lead to a rising aggregate price of risk. Alternatively, higher aggregate risk premia in the economy (as captured by high realizations of DY) lead to a higher cash-flow price of risk. If one considers time-varying cash-flow betas as the building block of the scaled ICAPM, then there is a priori no constraint on the sign of f_{CFz}. Both state variables are highly persistent with autoregressive coefficients of 0.99, and they are moderately correlated (correlation coefficient of 0.46). Thus, although there is some overlapping, these two state variables might explain different properties of the time-varying cash-flow beta/price of risk.[15]

The estimation and evaluation results associated with the benchmark ICAPM are displayed in Table 2, in which the test assets are the 30 portfolios (panel A) and the 25 size-momentum portfolios (panel B). In

[14] Harvey (1989), Ferson and Harvey (1999), and Petkova and Zhang (2005), among others, specify market betas as a function of the market dividend yield.

[15] In the estimation of the pricing equations, the lagged conditioning variables are demeaned, which is a common practice in the conditional CAPM literature.

Table 2 Factor Risk Premia for Scaled ICAPM

	γ_{CF}	γ_{CFz}	γ_H	χ^2	MAE (%)	R^2_{OLS} (%)
			Panel A (S10 + BM10 + M10)			
BBGB	2.96		−7.13	77.87	0.16	−23.45
	(**2.72**)		(−1.00)	(0.00)		
DY	−1.15	1,656.77	−24.00	27.44	0.11	59.85
	(−0.42)	(**2.70**)	(−1.87)	(0.44)		
CPI	−4.11	568.53	3.72	33.73	0.13	47.18
	(−1.19)	(2.03)	(0.29)	(0.17)		
			Panel B (SM25)			
BBGB	0.63		−58.41	28.54	0.24	34.44
	(0.19)		(−**2.68**)	(0.20)		
DY	−1.20	1,259.31	−50.07	33.49	0.14	73.68
	(−0.38)	(1.80)	(−**2.33**)	(0.06)		
CPI	−5.78	556.81	−30.02	35.59	0.11	86.34
	(−1.57)	(1.81)	(−1.90)	(0.03)		

Notes. This table reports the estimation and evaluation results for the benchmark scaled ICAPM. The models are the BBGB model and the ICAPM scaled by the dividend yield (*DY*) and *CPI* inflation (*CPI*). The testing portfolios are 10 size portfolios, 10 book-to-market portfolios, and 10 momentum portfolios (S10 + BM10 + M10, panel A) and 25 portfolios sorted on both size and momentum (SM25, panel B). The estimation procedure is first-stage GMM with equally weighted moments. γ_{CF} and γ_H denote the (covariance) risk prices estimates for cash-flow and discount-rate news, respectively, whereas γ_{CFz} denotes the risk price associated with the scaled factor. Below the parameter estimates are reported asymptotic heteroskedasticity-robust *t*-statistics. The column χ^2 presents the level (first line) and associated *p*-values (second line) for the asymptotic χ^2 test. MAE refers to the mean absolute pricing error (in %), and R^2_{OLS} denotes the cross-sectional OLS R^2 (in %). The sample is 1953:05–2009:12. Underlined and bold numbers denote statistical significance at the 5% and 1% levels, respectively.

the test of BBGB with the 30 portfolios, the point estimate for f_{CF} is around 3 and is significant at the 1% level, whereas the point estimate for f_H is negative, as expected, but strongly insignificant.[16] The average pricing error is 0.16% per month and the model is clearly rejected by the joint significance test (p-value = 0.00%). Furthermore, the cross-sectional R^2 is negative (−23%), which means that the model has less explanatory power than a model with the constant as the sole factor has. Therefore, these results show that the BBGB model is not capable of explaining simultaneously the returns of the three portfolio sorts and hence the corresponding three CAPM anomalies.

The results for the ICAPM scaled by DY show that the model is not rejected by the $\bullet^2$ test at the conventional levels (p-value = 44%), and around 60% of the cross-sectional variance in average excess returns is explained by the model, whereas the average pricing error is 0.11% per month. This fit represents a dramatic improvement relative to the unscaled ICAPM. The point estimate for f_{CFz} is positive and strongly significant (1% level), whereas the estimate for f_{CF} is negative but not statistically significant. On the other hand, the point estimate for f_H has larger magnitude than the corresponding estimate in the BBGB model, being significant at the 5% level for a one-sided test.

When the conditioning variable is *CPI*, the scaled ICAPM also passes the $\bullet^2$ test (p-value = 17%), and the cross-sectional coefficient of determination is only slightly lower than in the version based on DY (47%). The point estimate for f_{CFz} is positive and statistically significant (5% level), which means that a rise in inflation is associated with an increase in the aggregate cash-flow price of risk. On the other hand, the point estimates for both f_{CF} and f_H have the wrong sign; that is, f_{CF} is estimated negatively and f_H is estimated positively, but both coefficients are not significant.

Therefore, the scaled ICAPM is not rejected when tested on portfolios sorted on size, BM, and momentum and is able to explain a significant fraction of the dispersion in average excess returns on these portfolios. In contrast, the two-beta ICAPM does not perform well in pricing jointly those three portfolio groups, being also formally rejected. Moreover, the estimates for the fundamental parameters in the scaled ICAPM indicate that the time-varying or cyclical component of the cash-flow beta/price of risk dominates the corresponding constant or long-term component.

The results for the test with the SM25 portfolios show that the fit of the BBGB model increases substantially in relation to the test with the 30 portfolios, with an OLS R^2 estimate of 34%. The model is not

[16] Notice that CV04 define the beta (covariance) with the negative of discount-rate news (good news); thus, their estimates for f_H are positive.

rejected by the $\bullet^2$ test, although the average pricing error is 0.24% per month, which is economically large. The point estimate for f_{CF} is below one, but largely insignificant, whereas the estimate for f_H (−58) is significantly larger in magnitude than in the test with the 30 portfolios, being significant at the 1% level. Thus, the positive fit of BBGB for the 25 portfolios seems to be a result of the contribution of the discount-rate factor rather than the cash-flow factor. The ICAPM scaled by DY improves substantially the fit of BBGB in pricing SM25, with an average pricing error (0.14% per month) that is nearly half the corresponding value for BBGB, and the coefficient of determination more than doubles to 74%. Moreover, the model also passes the joint significance test, with a p-value marginally above 5%. The point estimate for f_{CFz} has a smaller magnitude than in the test with the 30 portfolios, but it is significant at the 5% level, whereas the magnitude of f_H is nearly twice the corresponding estimate in the test with S10 + BM10 + M10, being significant at the 1% level. When the conditioning variable is *CPI*, the explanatory ratio increases further to 86%, whereas the average pricing error is only 0.11% per month. Despite the large fit, the model is rejected by the $\bullet^2$ test, which should be the result of a poor inversion of the covariance matrix of the pricing errors, Var4$\hat{\boldsymbol{\alpha}}$5. The point estimate for f_{CFz} is very close to the corresponding estimate in the test with the 30 portfolios, being also significant at the 5% level. On the other hand, the point estimate for f_{CF} is negative but not significant at the 5% level, whereas f_H is estimated negatively in contrast to the benchmark test and is significant at the 5% level.

I also estimated the scaled ICAPM for the 10 momentum deciles as the only test assets. Untabulated results show that the version based on DY yields an explanatory ratio of 83% and an average pricing error of 0.11% per month, which represents a significantly larger fit than in the test with the 30 portfolios. In the case of the ICAPM scaled by *CPI*, the coefficient of determination estimate is 86%, which represents about twice the fit obtained in the test with the 30 portfolios. In comparison, the BBGB model explains only 23% of the cross-sectional dispersion in the average returns of the 10 momentum portfolios, and the MAE estimate (0.20% per month) is about twice the magnitude of mispricing for the scaled ICAPM.

Therefore, the benchmark scaled ICAPM explains a large fraction of the dispersion in average excess returns of the momentum portfolios. This represents a major improvement because most models in the empirical asset pricing literature (including the Fama and French 1993 three-factor model) fail to price portfolios sorted on short-term past stock returns (see Fama and French 1996, Cochrane 2007, among others).

3.5. Restricted ICAPM: Estimating the Factor Risk Premia

I estimate the restricted version of the ICAPM in which the (covariance) risk price associated with discount-rate news is fixed at −1; that is, there is no risk-free asset in the reference portfolio of the ICAPM investor (— = 1). The estimation of this version of the model enables one to assess whether allowing for a free risk price for the discount-rate factor plays a major role in driving the fit of the benchmark scaled ICAPM.

The estimation results are displayed in Table 3. In the test with the 30 portfolios the fit of the BBGB model is very similar to the one obtained for the benchmark specification, with a negative explanatory ratio (−25%), whereas the point estimate for f_{CF} is 3.25 and significant at the 1% level. Moreover, the model is rejected by the $\bullet^2$ test (p-value of zero). The ICAPM scaled by DY significantly outperforms the BBGB model, with an R^2 estimate of 40%. This explanatory ratio falls behind the corresponding fit for the benchmark ICAPM scaled by DY, thus showing that the (free) discount-rate risk price has some contribution to the overall fit of the unrestricted model. The cash-flow risk price is estimated positively (in contrast to the benchmark version), but it is strongly insignificant. The point estimate for f_{CFz} has a lower magnitude than in the benchmark model, but it remains significant at the 1% level. When the conditioning variable is *CPI*, the explanatory ratio of the restricted scaled ICAPM is nearly the same as in the benchmark version (46%), and the magnitude of the scaled risk price is very close to the corresponding estimate in the benchmark model, being significant at the 5% level. Both versions of the restricted scaled ICAPM pass the $\bullet^2$ test, with p-values well above 5%.

In the test with the SM25 portfolios, the restricted BBGB performs poorly, as indicated by the negative explanatory ratio (−26%) and the average pricing error of 0.34% per month. Moreover, the model does not pass the $\bullet^2$ test. Thus, a free estimated discount-rate risk price is crucial to drive the positive explanatory power of the benchmark BBGB for the SM25 portfolios, documented above. In the case of the ICAPM scaled by DY, the R^2 estimate is positive (30%), but the average pricing error of 0.26% per month is almost twice the magnitude of the mispricing in the corresponding test of the benchmark scaled ICAPM, leading to the formal rejection of the model (p-value of zero). The point estimate for f_{CF} is positive but largely insignificant as in the test with the 30 portfolios, whereas the magnitude of f_{CFz} is somewhat larger than in the benchmark version, being strongly significant. When the scaling variable is *CPI*, the model's fit is as large as 74%,

Table 3 Factor Risk Premia for Restricted Scaled ICAPM

	γ_{CF}	γ_{CFz}	χ^2	MAE (%)	R^2_{OLS} (%)
		Panel A (S10 + BM10 + M10)			
BBGB	3.25		76.42	0.16	−25.08
	(**3.08**)		(0.00)		
DY	0.50	1,372.93	32.70	0.13	40.29
	(0.24)	(**2.92**)	(0.25)		
CPI	−4.12	552.29	35.72	0.13	46.32
	(−1.09)	(1.67)	(0.15)		
		Panel B (SM25)			
BBGB	3.27		116.19	0.34	−26.39
	(**2.97**)		(0.00)		
DY	0.64	1,491.65	50.88	0.26	30.38
	(0.29)	(**3.18**)	(0.00)		
CPI	−6.20	683.06	35.48	0.16	74.16
	(−1.49)	(2.03)	(0.05)		

Notes. This table reports the estimation and evaluation results for the restricted scaled ICAPM. The models are the BBGB model and the ICAPM scaled by the dividend yield (*DY*) and *CPI* inflation (*CPI*). The testing portfolios are 10 size portfolios, 10 book-to-market portfolios, and 10 momentum portfolios (S10 + BM10 + M10, panel A) and 25 portfolios sorted on both size and momentum (SM25, panel B). The estimation procedure is first-stage GMM with equally weighted moments. γ_{CF} denotes the (covariance) risk price estimate for cash-flow news, whereas γ_{CFz} represents the risk price associated with the scaled factor. Below the parameter estimates are reported asymptotic heteroskedasticity-robust t-statistics. The column χ^2 presents the level (first line) and associated p-values (second line) for the asymptotic χ^2 test. MAE refers to the mean absolute pricing error (in %), and R^2_{OLS} denotes the cross-sectional OLS R^2 (in %). The sample is 1953:05–2009:12. Underlined and bold numbers denote statistical significance at the 5% and 1% levels, respectively.

which is only marginally smaller than the corresponding explanatory ratio in the benchmark model, and the average pricing error is 0.16% per month. Moreover, the model is not (marginally) rejected by the $\bullet^2$ test (p-value = 5%). The magnitude of the scaled risk price is slightly above the counterpart estimate in the benchmark model, being significant at the 5% level. Overall, these results show that the scaled ICAPM significantly outperforms the unscaled ICAPM in pricing both sets of portfolios when the discount-rate risk price is restricted in both models. Secondly, the discount-rate factor does not play a relevant role in the version based on *CPI*, whereas in the case of the version based on *DY*, it has a relevant contribution for the model's global fit, especially in pricing the 25 size-momentum portfolios.

3.6. Comparison with Alternative Factor Models

This subsection puts the results presented above in perspective by comparing the fit of the scaled ICAPM with alternative asset pricing models. Despite the documented failure of the unconditional CAPM in pricing the cross-section of returns, there is prior evidence that the conditional CAPM does a good job in pricing the returns of size/BM portfolios and thus explaining the size and value anomalies (Jagannathan and Wang 1996, Ferson and Harvey 1999, Lettau and Ludvigson 2001, among others). Given that the scaled ICAPM contains a scaled factor (similarly to the conditional CAPM specifications), it should be relevant to compare the two models in pricing both portfolio groups by using the same conditioning variables, *DY* and *CPI*. The conditional CAPM specification is given by[17]

$$\mathrm{E}4R_{i1t+1} - R_{f1t+1}5 = f_M \mathrm{Cov}4R_{i1t+1}1RMRF_{t+1}5 + f_{Mz}\mathrm{Cov}4R_{i1t+1}1RMRF_{t+1}z_t51 \quad (24)$$

where *RMRF* denotes the excess market return. The baseline unconditional CAPM is obtained by imposing $f_{Mz} = 0$. The results presented in Table 4 show that the baseline CAPM cannot price both sets of portfolios, with a coefficient of determination around −24% in both tests. This fit is the same as the BBGB model in the test with the 30 portfolios, but the unscaled ICAPM outperforms significantly in pricing the 25 size-momentum portfolios. On the other hand, the R^2 estimates of the conditional CAPM in the test with the 30 portfolios are 20% and 29% when the conditioning variables are *DY* and *CPI*, respectively. In the test with SM25, the explanatory ratios are 16% and 87% in the versions with *DY* and *CPI*, respectively. Thus, with the exception of the ICAPM scaled by *CPI*

[17] Equation (24) is similar to the model presented in Lettau and Ludvigson (2001), with the difference that there is no place for the separated conditioning variable risk; that is, I explicitly impose the constraint that the conditional covariance of returns with the lagged state variable should be equal to zero.

Table 4 Factor Risk Premia for Conditional CAPM

	γ_M	γ_{Mz}	χ^2	MAE (%)	R^2_{OLS} (%)
		Panel A (S10 + BM10 + M10)			
CAPM	3.10		78.93	0.16	−23.68
	(**3.28**)		(0.00)		
DY	1.91	1,350.14	41.73	0.16	19.62
	(1.07)	(2.21)	(0.05)		
CPI	−1.04	562.81	42.40	0.13	29.38
	(−0.39)	(**2.63**)	(0.04)		
		Panel B (SM25)			
CAPM	3.14		118.90	0.33	−22.77
	(**3.19**)		(0.00)		
DY	2.15	1,429.07	63.57	0.27	15.60
	(1.14)	(**3.33**)	(0.00)		
CPI	−3.56	895.33	27.77	0.10	86.81
	(−1.00)	(2.21)	(0.22)		

Notes. This table reports the estimation and evaluation results for the conditional CAPM. The models are the unconditional CAPM (CAPM) and the CAPM scaled by the dividend yield (*DY*) and *CPI* inflation (*CPI*). The testing portfolios are 10 size portfolios, 10 book-to-market portfolios, and 10 momentum portfolios (S10 + BM10 + M10, panel A) and 25 portfolios sorted on both size and momentum (SM25, panel B). The estimation procedure is first-stage GMM with equally weighted moments. γ_M denotes the (covariance) risk price estimate for the market factor, whereas γ_{Mz} represents the risk price associated with the scaled factor. Below the parameter estimates are reported asymptotic heteroskedasticity-robust t-statistics. The column χ^2 presents the level (first line) and associated p-values (second line) for the asymptotic χ^2 test. MAE refers to the mean absolute pricing error (in %), and R^2_{OLS} denotes the cross-sectional OLS R^2 (in %). The sample is 1953:05–2009:12. Underlined and bold numbers denote statistical significance at the 5% and 1% levels, respectively.

in the test with SM25, both the restricted and unrestricted versions of the scaled ICAPM outperform the conditional CAPM by a good margin, especially in the case of the benchmark version. This finding has two implications: First, the scaled factor associated with aggregate cash-flow news has greater explanatory power for the cross-section of stock returns than does the scaled factor associated with the market return. Second, the unscaled ICAPM beats the unconditional CAPM; that is, assigning different risk prices for the two components of the market return (cash-flow and discount-rate news) helps explain the dispersion in equity premia within the cross section.

I also compare the performance of the scaled ICAPM against the Fama and French (1993) three-factor model (FF3 hereafter):

$$\begin{aligned} \mathrm{E}(R_{i,t+1} - R_{f,t+1}) = {} & f_M \operatorname{Cov}(R_{i,t+1}, RMRF_{t+1}) \\ & + f_{SMB} \operatorname{Cov}(R_{i,t+1}, SMB_{t+1}) \\ & + f_{HML} \operatorname{Cov}(R_{i,t+1}, HML_{t+1}), \end{aligned} \tag{25}$$

and the Carhart (1997) four-factor model,

$$\begin{aligned} \mathrm{E}(R_{i,t+1} - R_{f,t+1}) = {} & f_M \operatorname{Cov}(R_{i,t+1}, RMRF_{t+1}) \\ & + f_{SMB} \operatorname{Cov}(R_{i,t+1}, SMB_{t+1}) \\ & + f_{HML} \operatorname{Cov}(R_{i,t+1}, HML_{t+1}) \\ & + f_{UMD} \operatorname{Cov}(R_{i,t+1}, UMD_{t+1}), \end{aligned} \tag{26}$$

where *SMB*, *HML*, and *UMD* denote the size, value, and momentum factors, respectively. Both of these models have less theoretical background than either the CAPM or ICAPM but are nevertheless empirically successful factor models in the literature. The results reported in Table 5 show that FF3 cannot price the 30 portfolios as indicated by the negative explanatory ratio (−19%), which means that the model performs worse than a model that predicts constant excess returns within the 30 portfolios. This poor fit results from the poor performance of the model in pricing the 10 momentum portfolios because it is well known that it prices well both the size and value premia. In the test with SM25, the R^2 estimate is basically zero (1%), thus confirming that the three-factor model cannot price the momentum anomaly. On the other hand, the four-factor model produces a large fit for the two portfolio groups, with explanatory ratios of 89% in both cases. The large explanatory power of the Carhart (1997) model in pricing the momentum anomaly is not surprising because the UMD factor was specifically designed to price the portfolios sorted on short-term prior returns. What is more interesting is that the scaled ICAPM is almost as successful in explaining the hard-to-price momentum portfolios. Overall, the performance of the scaled ICAPM is quite satisfactory in relation to the alternative models.

3.7. Individual Pricing Errors

To answer the question of which portfolios are priced or not by the scaled ICAPM, I conduct an analysis of the individual pricing errors. Figure 1 plots

Table 5 Factor Risk Premia for Alternative Factor Models

Row	γ_M	γ_{SMB}	γ_{HML}	γ_{UMD}	χ^2	MAE (%)	R^2_{OLS} (%)
			Panel A (S10 + BM10 + M10)				
1	3.10	1.11	2.33		73.05	0.14	−18.65
	(**3.14**)	(0.81)	(1.50)		(0.00)		
2	4.87	1.45	7.66	5.98	39.78	0.05	88.75
	(**4.25**)	(1.00)	(**4.50**)	(**4.36**)	(0.04)		
			Panel B (SM25)				
1	0.73	2.80	−10.49		88.49	0.30	1.44
	(0.58)	(1.68)	(−**3.01**)		(0.00)		
2	5.79	2.33	11.65	7.47	58.86	0.10	88.58
	(**4.14**)	(1.35)	(**3.76**)	(**4.49**)	(0.00)		

Notes. This table reports the estimation and evaluation results for alternative factor models. The models are the Fama and French (1993) three-factor model (row 1) and the Carhart (1997) four-factor model (row 2). The testing portfolios are 10 size portfolios, 10 book-to-market portfolios, and 10 momentum portfolios (S10 + BM10 + M10, panel A), and 25 portfolios sorted on both size and momentum (SM25, panel B). The estimation procedure is first-stage GMM with equally weighted moments. γ_M, γ_{SMB}, γ_{HML}, and γ_{UMD} denote the (covariance) risk price estimates associated with the market, size, value, and momentum factors, respectively. Below the parameter estimates are reported asymptotic heteroskedasticity-robust t-statistics. The column χ^2 presents the level (first line) and associated p-values (second line) for the asymptotic χ^2 test. MAE refers to the mean absolute pricing error (in %), and R^2_{OLS} denotes the cross-sectional OLS R^2 (in %). The sample is 1953:05–2009:12. Underlined and bold numbers denote statistical significance at the 5% and 1% levels, respectively.

the estimated average excess returns (horizontal axis) against the realized average excess returns (vertical axis) associated with both portfolio groups. We can see that the pairs of realized/predicted returns are much closer to the diagonal line (thus indicating higher explanatory power) for the scaled ICAPM in comparison with the BBGB model, which just confirms the higher R^2 estimates discussed above. In the test with the 30 portfolios, the biggest outlier in the graphs for the ICAPM scaled by DY and CPI are the extreme winner (decile 10 among M10) portfolio with pricing errors of 0.38% and 0.41%, respectively, which are still below the corresponding mispricing for BBGB (0.46%). Another difficult portfolio to price is the extreme loser portfolio (decile 1 among M10), for which the BBGB model has a mispricing of −1.01%, whereas the ICAPM scaled by DY and CPI produce much lower pricing errors of −0.21% and −0.38%, respectively. Thus, the outperformance of the scaled ICAPM in pricing the 30 portfolios is closely related to the better explanatory power for the momentum portfolios. The analysis of the SM25 portfolios confirms this finding because the BBGB model has a number of portfolios as significant outliers, with magnitudes of mispricing above 0.30% per month. The portfolio that is the hardest to price for the scaled ICAPM is the loser portfolio within the second size decile (21), with pricing errors of −0.38% and −0.28% in the versions with DY and CPI, respectively, but the corresponding mispricing in the BBGB model is significantly greater (−0.78%). The small-winner portfolio (15) is also difficult to price for the ICAPM scaled by DY (alpha = 0.48%), but the version based on CPI produces a much lower mispricing (0.19%).[18] These results represent additional evidence that the scaled ICAPM clearly outperforms the BBGB model in pricing the momentum portfolios, especially the "extreme" hard to price winner and loser portfolios.

4. Additional Results

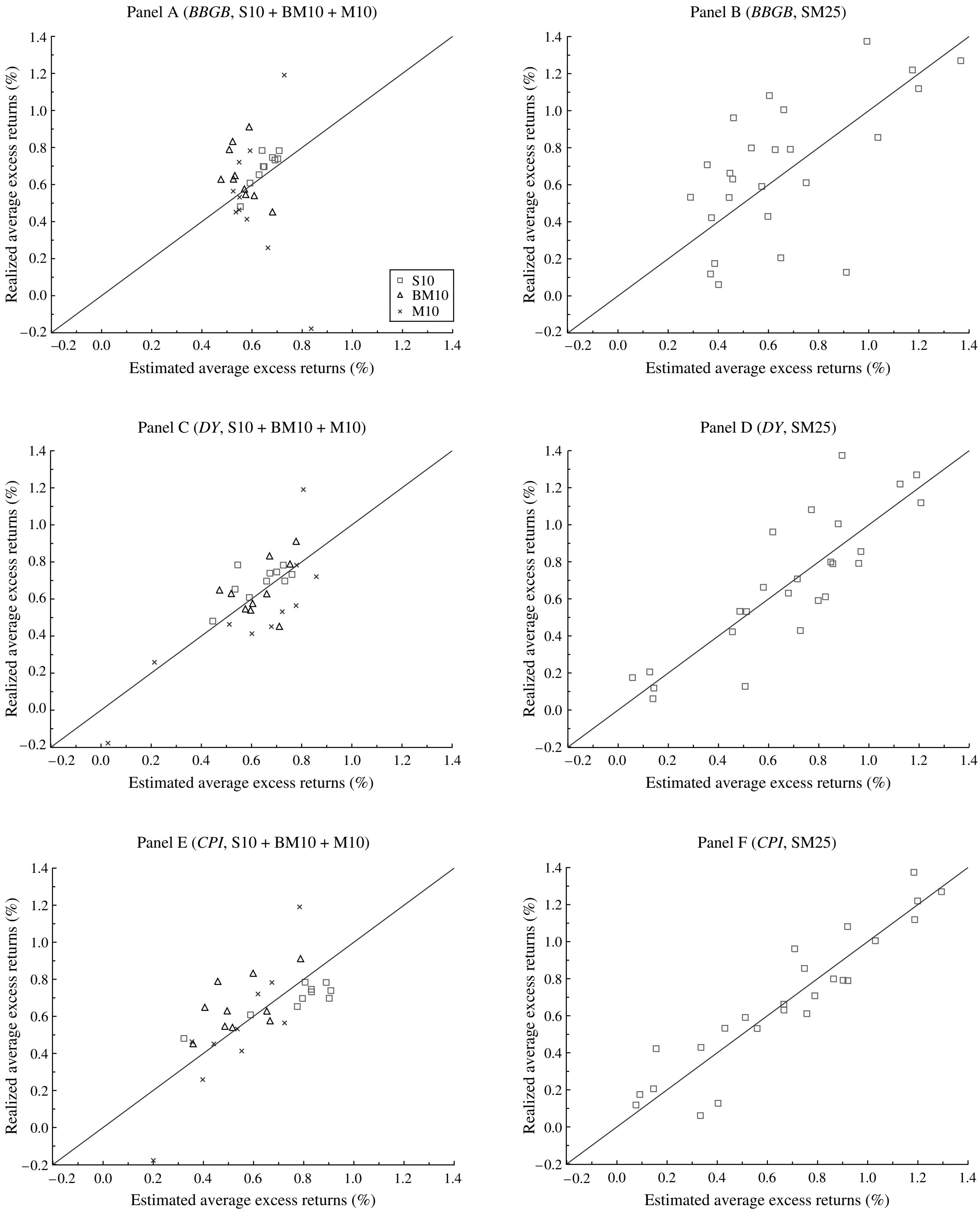
In this section, I conduct several robustness checks to the main results associated with the benchmark (unrestricted) scaled ICAPM in §3. The respective results are presented in the online appendix to this paper (available at http://www.hanken.fi/staff/paulmaio/).

4.1. Alternative VAR Specifications

I use alternative VAR specifications to estimate the components of unexpected market returns—cash-flow and discount-rate news. In the first specification the value spread is excluded from the VAR state vector; that is, $\mathbf{x}_t \equiv [RREL_t, TERM_t, EY_t, r_{mt}]'$. This enables one to assess whether the inclusion of VS drives the fit of the scaled ICAPM because CV04 claim that the fit of the BBGB model is dependent on this variable. The results show that the R^2_{OLS} estimates for the scaled ICAPM are very similar to the corresponding estimates in the benchmark test, varying between 51% (CPI in the test with the 30 portfolios) and 88% (CPI, test with SM25). Thus, the two versions of the scaled ICAPM outperform significantly the BBGB model, which has explanatory ratios of

[18] Hong et al. (2000) document that the momentum profits are larger among small stocks.

Figure 1 Individual Pricing Errors



Notes. This figure plots the average excess returns (in %) of equity portfolios against the corresponding excess returns predicted from alternative ICAPM models. The models are the BBGB model and the ICAPM scaled by the dividend yield (*DY*) and *CPI* inflation (*CPI*). The equity portfolios are 10 size portfolios, 10 book-to-market portfolios, and 10 momentum portfolios (S10 + BM10 + M10; panels A, C, and E) and 25 portfolios sorted on both size and momentum (SM25; panels B, D, and F). The sample is 1953:05–2009:12.

−19% and 48% in the tests with S10 + BM10 + M10 and SM25, respectively. In the version based on *CPI*, the point estimates for f_{CFz} are very close to the corresponding estimates in the benchmark test, whereas in the version with DY the magnitudes are smaller. However, these estimates are statistically significant at the 5% level, the exception being the ICAPM scaled by DY in the test with SM25.

The second alternative VAR specification includes the default spread (*DEF*; Keim and Stambaugh 1986, Fama and French 1989), constructed as the yield spread between BAA and AAA corporate bonds from Moody's; $\mathbf{x}_t \equiv [RREL_t, TERM_t, DEF_t, VS_t, EY_t, r_{mt}]'$. The results show that the fit of the scaled ICAPM is quite close to the benchmark test, with R^2_{OLS} estimates varying between 44% (*CPI*, S10+BM10 + M10) and 85% (*CPI*, SM25). The magnitudes of the scaled factor risk price are marginally larger than in the benchmark case and are statistically significant at the 5% and 1% levels. The point estimates for f_{CF} are negative in all four cases, but only in the version with *CPI* tested with SM25 is the estimate statistically significant. The fit of BBGB is basically the same as in the benchmark case when the equity portfolios are S10 + BM10 + M10 and increases marginally to 44% in the test with SM25, although still lagging behind the scaled ICAPM.

In a third VAR specification, I include the CP factor from Cochrane and Piazzesi (2005), $\mathbf{x}_t \equiv [RREL_t, TERM_t, CP_t, VS_t, EY_t, r_{mt}]'$. Overall, the results are very similar to the benchmark VAR specification.

4.2. Augmented Scaled ICAPM

I estimate an augmented version of the scaled ICAPM in which the price of risk for discount-rate news is also time varying and linear in the state variable:

$$f_{H,t} = f_2 + f_3 z_t, \tag{27}$$

leading to the following four-factor model,

$$\begin{aligned} \mathrm{E}(R_{i,t+1} - R_{f,t+1}) = f_{CF}\,\sigma_{i,CF} + f_{CFz}\,\sigma_{i,CFz} \\ + f_H\,\sigma_{i,H} + f_{Hz}\,\sigma_{i,Hz}, \end{aligned} \tag{28}$$

where $\sigma_{iHz} \equiv \mathrm{Cov}(r_{i,t+1}, r^H_{t+1} z_t)$ denotes the covariance with the new scaled factor. The results show that the point estimates for f_{Hz} are not robust in sign, being positive when the scaling variable is DY and negative when the scaling variable is *CPI*, and in all cases this coefficient is not significant at the 5% level. Moreover, the R^2_{OLS} estimates are only marginally greater than in the benchmark test, varying between 51% (S10 + BM10 + M10) and 90% (SM25) in both cases with *CPI* as the state variable. On the other hand, f_{CFz} is estimated positively and is statistically significant, the exception being the version with DY when the test portfolios are SM25, for which case the coefficient is only marginally significant (at the 10% level). Therefore, the inclusion of the additional scaled factor, $r^H_{t+1} z_t$, does not add relevant explanatory power to the scaled ICAPM.

I also estimate a three-factor model that excludes the scaled cash-flow factor:

$$\mathrm{E}(R_{i,t+1} - R_{f,t+1}) = f_{CF}\,\sigma_{i,CF} + f_H\,\sigma_{i,H} + f_{Hz}\,\sigma_{i,Hz}. \tag{29}$$

The results show that the scaled discount-rate factor is priced (1% level) for both tests when the conditioning variable is DY. In the version based on *CPI*, the estimate for f_{Hz} is significant (5% level) in the test with the 30 portfolios, whereas in the test with SM25 there is statistical significance only at the 10% level. Compared to the benchmark scaled ICAPM from the last section, the new model clearly underperforms in the test with the 30 portfolios with R^2 estimates of 26% and 2% when the state variables are DY and *CPI*, respectively. In the test with SM25 the explanatory ratios are 54% and 57% for DY and *CPI*, respectively, which still lag behind the fit of the benchmark scaled ICAPM by a good margin.

4.3. Expected Return-Beta Representation

I test the scaled ICAPM in expected return-beta form by using the time-series/cross-sectional regressions approach presented in Cochrane (2005, Chap. 12) and Brennan et al. (2004), which enables us to obtain estimates for factor betas and (beta) prices of risk. The factor betas are estimated from the following time-series regressions for each test asset:

$$\begin{aligned} r_{i,t+1} = \alpha_i + \beta_{i,CF}\, r^{CF}_{t+1} + \beta_{i,CFz}\, r^{CF}_{t+1} z_t \\ + \beta_{i,H}\, r^{H}_{t+1} + \varepsilon_{i,t+1}. \end{aligned} \tag{30}$$

The expected return-beta representation is estimated in a second step by the following OLS cross-sectional regression:

$$\overline{R_i - R_f} = \lambda_{CF}\,\beta_{i,CF} + \lambda_{CFz}\,\beta_{i,CF,z} + \lambda_H\,\beta_{i,H} + \alpha_i, \tag{31}$$

which produces estimates for factor (beta) risk prices (λ) and pricing errors (α_i). In the above cross-sectional regression, $\overline{R_i - R_f} \equiv (1/T)\sum_{t=0}^{T-1}(R_{i,t+1} - R_{f,t+1})$ represents the average excess return for asset i. The main difference of this method relative to the GMM estimation is that the factor betas in the cross-sectional regression are multiple-regression betas instead of single-regression betas or covariances; thus, they account for the correlation among the factors. The t-statistics for the factor risk prices and the computation of $\mathrm{Var}(\hat{\alpha})$ are based on Shanken (1992) standard errors, which introduce a correction for the estimation error in the factor betas from the time-series regressions.

The results show that the R^2 and *MAE* estimates are the same as in the GMM estimation of the expected return-covariance representation of the model. Moreover, the estimates for γ_{CFz} are positive in all four cases, and most of these estimates are statistically significant at the 1% level. The sole exception is the ICAPM scaled by DY when the test portfolios are SM25 in which γ_{CFz} is not significant, which represents a signal of multicollinearity. The point estimate for γ_{CF} is significantly positive in the version with *CPI* tested on the 30 portfolios but estimated negatively in the remaining three cases (similarly to the GMM test).

4.4. Alternative Portfolios

I test the scaled ICAPM with alternative portfolio groups that combine momentum with several other stock characteristics. The first three portfolio groups are 25 (5 × 5) portfolios sorted on both momentum and idiosyncratic volatility (MIR25), as in Ang et al. (2006); 25 portfolios sorted on momentum and stock volatility (MSD25), as in Jiang et al. (2005) and Zhang (2006); and 25 portfolios sorted on momentum and credit risk (MDEF25), as in Avramov et al. (2007). The proxy for credit risk is the stock beta relative to the default spread. Idiosyncratic volatility is measured relative to the Fama and French (1993) model, and stock volatility corresponds to the standard deviation of daily stock returns over the last month. The stock return data are from CRSP.

Untabulated results show that the ICAPM scaled by *CPI* outperforms slightly the BBGB model in pricing the MIR25 portfolios with a R^2 estimate of 48% versus 36% for BBGB, whereas the version based on DY has about the same explanatory power as the two-factor ICAPM. When the test portfolios are MSD25, the ICAPM scaled by *CPI* significantly outperforms the BBGB model with an explanatory ratio of 45% (versus 13% for BBGB), whereas the version based on DY has a relatively modest fit ($R^2 = 23\%$). In the test with the MDEF25 portfolios, the explanatory power of the two versions of the scaled ICAPM is very similar (56% and 58% for DY and *CPI*, respectively) and also very close to the fit of the two-factor ICAPM. Thus, the scaled factor does not add explanatory power for the MDEF25 portfolios.

The remaining portfolio groups analyzed are 30 (10 × 3) portfolios sorted on momentum and average daily volume (MV30), as in Lee and Swaminathan (2000), and 15 (5 × 3) portfolios sorted on momentum and age (MAGE15), as in Jiang et al. (2005) and Zhang (2006). Age (like stock volatility above) represents a proxy for information uncertainty and is measured as the difference between the date of the current return observation and the first observation available in CRSP. The results for the test with MV30 show that the fit of the scaled ICAPM (26% and 21% for DY and *CPI*, respectively) is only marginally higher than the BBGB model (17%). In the test with the MAGE15 portfolios, the three models yield a similar explanatory power (R^2 estimates around 74%), showing that the scaled factor is not relevant to price these portfolios.

In all five cross-sectional tests, the discount-rate factor is priced and the associated risk price estimates are positive, which is inconsistent with the negative sign predicted from the ICAPM. Thus, the scaled ICAPM can be interpreted as an empirical factor model that can explain the dispersion in average returns for most of these alternative portfolio groups, although the risk prices are inconsistent with the underlying theory of the Campbell (1993) ICAPM.

4.5. Additional Robustness Checks

I conduct several additional robustness checks, which are presented in the online appendix. Specifically, I conduct the cross-sectional tests by including bond returns in addition to equity portfolios; estimate the model by including an intercept in the pricing equation; estimate the scaled ICAPM for 25 portfolios sorted on size and BM; use an alternative measure of the state variables; and, finally, conduct a bootstrap simulation to provide an empirical distribution for the risk price estimates. Overall, the results indicate that the scaled ICAPM largely outperforms the BBGB model in pricing the cross-section of stock returns.

5. Discussion

In this section, I provide additional discussion and economic intuition for the empirical results reported in the previous sections.

5.1. Factor Loading Estimates

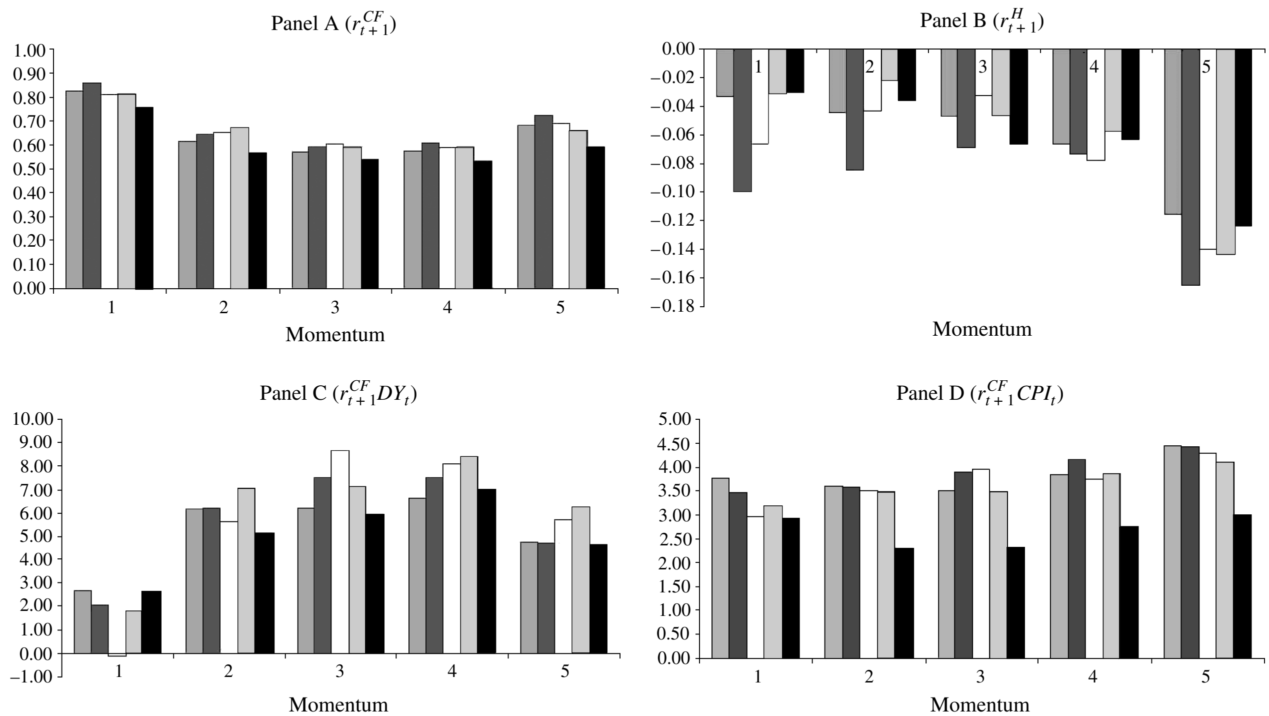
To interpret the mechanism by which the scaled ICAPM explains the cross-section of stock returns, it is convenient to restate the pricing equation in terms of single regression betas:

$$\begin{aligned} \mathrm{E}(R_{i,t+1} - R_{f,t+1}) &= \lambda_{CF}\,\sigma^2_{CF}\,\beta_{i,CF} + \lambda_{CFz}\,\sigma^2_{CFz}\,\beta_{i,CFz} + \lambda_{H}\,\sigma^2_{H}\,\beta_{i,H} \\ &= \gamma_{CF}\,\beta_{i,CF} + \gamma_{CFz}\,\beta_{i,CFz} + \gamma_{H}\,\beta_{i,H}, \end{aligned} \tag{32}$$

where σ^2_{CF}, σ^2_{CFz}, and σ^2_{H} represent the unconditional variances of r^{CF}_{t+1}, $r^{CF}_{t+1}z_t$, and r^{H}_{t+1}, respectively; $\beta_{i,CF}$, $\beta_{i,CFz}$, and $\beta_{i,H}$ are the factor loadings; and the γs denote the (beta) risk prices.

To assess the way the scaled ICAPM explains the momentum portfolios, one needs to analyze the single regression betas associated with the SM25 portfolios, for each of the factors in the model, which are presented in Figure 2. As expected, the betas associated with r^{CF}_{t+1} and r^{H}_{t+1} are, respectively, positive and negative across all portfolios. The cash-flow betas display an approximate u-shaped pattern within each size quintile, with past losers having marginally larger betas than past winners. On the other hand, in the

Figure 2 Single-Regression Beta Estimates for SM25



Notes. This figure plots the single-regression beta estimates associated with the SM25 portfolios from the benchmark scaled ICAPM. The factors are cash-flow news (panel A), discount-rate news (panel B), and the scaled factors when the state variables are the dividend yield (panel C) and CPI inflation (panel D). $i = 1, 2, 3, 4, 5$ designates the momentum (prior return) quintiles, and within each momentum quintile we have the portfolios associated with each size quintile (from small to large). The sample is 1953:05–2009:12.

case of the discount-rate factor the betas for past winners are significantly more negative than the betas for past losers, thus past winners are more sensitive to shocks in future aggregate risk premia. Regarding the first scaled factor, $r^{CF}_{t+1} DY_t$, the betas of past winners have greater magnitudes than do the betas of past losers within each size quintile, although the relation is not monotonic. Regarding the other scaled factor, $r^{CF}_{t+1} CPI_t$, it is also the case that past winners have larger factor loadings than do past losers, and the relation is more close to a monotonic one, with the exception of the larger size quintile, in which case the pattern of betas across the momentum portfolios is relatively flat. This dispersion in the factor loadings associated with the scaled and discount-rate factors among the momentum portfolios should be linked to the capacity of the scaled ICAPM in explaining the momentum anomaly.

Thus, to explain which factors drive the explanatory power of the scaled ICAPM for the SM25 portfolios, I calculate the average factor risk premium (average beta times factor risk price) for each momentum quintile, where the average beta for quintile i is computed from the 5 (of 25) portfolios associated with this quintile. The spread in average excess returns between the first ($Q1$) and fifth momentum quintile ($Q5$) is as large as −1.03% per month. This gap has to be partially matched by the risk premium associated with one or more of the factors in the scaled ICAPM, as shown in the beta pricing equation above. In the case of the BBGB model, the gap in the cash-flow risk premium between $Q1$ and $Q5$ is marginally positive (0.03%), whereas the gap associated with the discount-rate factor is negative (−0.64%). Thus, the fit of the unscaled ICAPM for the momentum quintiles is entirely attributable to the discount-rate factor. However, there is still a spread, $Q1 - Q5$, not explained by the model of −0.42%. In the ICAPM scaled by DY, the gap in risk premia, $Q1 - Q5$, attached to the discount-rate news factor is about −0.55%, whereas the scaled factor also produces a negative gap (−0.29%). This is determinant to obtain a spread, $Q1 - Q5$, in mispricing of only −0.15%, which represents less than half the corresponding mispricing in the BBGB model. In the version with CPI, both $r^{CF}_{t+1} CPI_t$ and r^{H}_{t+1} have an equal contribution in explaining the momentum premium, with risk premia gaps of −0.34% and −0.33%, respectively. On the other hand, the cash-flow factor also helps in explaining the momentum spread, with a gap, $Q1 - Q5$, in risk premium of −0.25% resulting in a spread, $Q1 - Q5$, left unexplained by the model of only −0.12%. Thus, the factors that allow the scaled ICAPM to price momentum are basically discount-rate news and the scaled factor. Specifically, the conditional cash-flow risk premium is greater for winner stocks in comparison to loser stocks, and this explains why the scaled ICAPM can explain momentum significantly better than the BBGB model can. Thus, past winners are riskier than past losers not because they have higher cash-flow risk in average times but because they are more correlated with aggregate cash flows in "bad times," that is, when the realizations of the state variables are above their means.[19] To explore this issue in more detail, and following Lettau and Ludvigson (2001), I calculate average conditional cash-flow betas as

$$\mathrm{E}(\beta_{i,CF,t}) = \beta_{i,CF} + \beta_{i,CFz}\,\mathrm{E}(z_t), \qquad (33)$$

where $\mathrm{E}(z_t)$ represents the average of the scaling variable, z_t, calculated over periods with high and low realizations of the state variable.[20] A period with high realizations of the state variable occurs when the state variable is 2.5 standard deviations above its mean. On the other hand, because of the asymmetric distribution of both DY and *CPI*, a period with low values of the state variable is when z_t is 1.8 standard deviations below its mean. When one considers all the periods, the average conditional cash-flow beta corresponds to the unconditional (multiple-regression) cash-flow beta because the state variable has unconditional zero mean.

Untabulated results show that in periods with low realizations of either DY or *CPI*, past losers have significantly higher cash-flow betas than past winners across all size quintiles. On the other hand, when the state variables are significantly higher than the respective means, it follows that the pattern of cash-flow betas between past losers and past winners is relatively flat; actually, for several size quintiles, winners have larger cash-flow betas than losers. Thus, past winners are riskier than past losers because they have greater cash-flow risk in times of high realizations of either DY or *CPI*.

Why are past winners riskier than past losers in periods with high inflation? A possible explanation is that during economic expansions (which tend to be associated with higher inflation) winners tend to be cyclical firms, which have high cash-flow betas. On the other hand, during recessions (periods with low inflation) winners tend to be noncyclical firms with low cash-flow betas. The changing composition of the momentum portfolios leads to the time variation in the cash-flow betas of these portfolios. These results are consistent with the evidence provided in Daniel and Moskowitz (2011) that momentum profits are associated with time-varying market betas of the winner and loser portfolios. Specifically, after a bear equity market, the market (cash-flow) beta of the momentum factor is low because past winners have low betas (defensive stocks that performed relatively better in the bear market) and past losers have high betas (aggressive or cyclical stocks that underperformed in the bear market). Because in a bear market the inflation rate is usually at low levels, it follows that past losers have high betas when inflation is low and past winners have low betas. On the other hand, in a bull market inflation is at relatively high levels, and thus past winners (those that have outperformed in the bull market) have high cash-flow betas, whereas past losers have low betas. Thus, inflation represents an instrument that signals time variation in market (cash-flow) betas of the winner and loser portfolios as a result of changing market conditions, and hence, of the changing composition of the momentum portfolios.

When the state variable is DY, a possible explanation for why losers have higher cash-flow betas than winners have in periods with low aggregate dividend yield (i.e., periods with low risk premia or low aggregate cash-flow price of risk) is the convexity effect described in Johnson (2002). Under this model, stocks with higher expected growth in the future cash flows are more sensitive to additional shocks in their cash flows. This might explain why past winners, which most likely had positive shocks in cash-flow news in the recent past and have higher expected growth for future cash flows, might be more correlated with aggregate cash-flow news. On the other hand, in periods with low DY (that correspond to low aggregate price of risk or low risk premia) the growth options of past losers become in-the-money dominating the convexity effect. Consequently, past losers are more correlated with aggregate cash flows in periods with low aggregate dividend yield.[21] This intuition is also consistent with the theoretical model proposed by Avramov and Hore (2008), which predicts that the momentum profits are positively correlated with the volatility and persistence of expected cash flows.

To assess how the scaled ICAPM performs in pricing the momentum portfolios during the momentum crash occurred in 2009, as documented by Daniel and Moskowitz (2011), I estimate the model in the test with the SM25 portfolios for the 1953:05–2008:12 period. Untabulated results show that both versions

[19] Koijen et al. (2012) offer a similar business cycle argument to explain the value premium; that is, the cash flows of value stocks are more correlated with aggregate cash flows during recessions, making those stocks riskier than growth stocks.

[20] The multiple-regression betas are calculated from the time-series regressions in (30): $r_{i,t+1} = a_i + \beta_{i,CF} r^{CF}_{t+1} + \beta_{i,CFz} z_t r^{CF}_{t+1} + \beta_{i,H} r^{H}_{t+1} + e_{i,t+1}, \; i = 1, \ldots, N.$

[21] Liu and Zhang (2008) make a similar argument by using industrial production growth as the factor that drives momentum.

of the scaled ICAPM have an explanatory ratio around 60% and an MAE estimate of 0.22% per month. Thus, the fit of the model is slightly smaller for the reduced sample, particularly in the version based on *CPI*, which shows that the scaled ICAPM performs relatively well in pricing the momentum portfolios during the 2009 momentum crash.

5.2. Discount-Rate Prices of Risk

As shown in the derivation of the pricing equation for the scaled ICAPM in (13), the risk price for discount-rate news is given by $-\omega$, where ω represents the weight associated with the stock index in the global portfolio of the ICAPM investor. Thus, from the point estimates of γ_H we can obtain the implied estimates for $\omega = -\gamma_H$. In the case of the benchmark test in Table 2, the point estimates for ω, which are statistically significant, vary between 24 (test with S10 + BM10 + M10) and 50 (SM25). These estimates correspond to portfolio weights for the stock index of 2,400% and 5,000%, or alternatively, weights for the risk-free asset of −2,300% and −4,900%, which represent very large leverage ratios. However, in the test including bond risk premia, presented in the online appendix, the magnitudes of the estimates for ω decrease significantly, varying between 7 and 15 in the test with the 30 portfolios. These estimates correspond to implied weights on the risk-free asset of −600% and −1,400%, respectively. Although these estimates represent high degrees of financial leverage, they are in line with simulation results found in the dynamic portfolio choice literature (e.g., Campbell et al. 2003).

5.3. Time-Varying Risk Aversion

In the analysis conducted in the previous sections, the assumption of a time-varying price of risk for cash-flow news is related to time-varying aggregate risk premia or the notion that the price of risk should increase in "bad times." Alternatively, one can assume time-varying relative risk aversion (RRA) by specifying a state-dependent version of the Epstein and Zin (1989, 1991) utility over consumption (C) that accounts for time-varying RRA, similar to the specification used in Melino and Yang (2003):[22]

$$U_t = \left\{ C_t^{(1-\gamma_t)/\theta_t} + \beta \left[\mathrm{E}_t(U_{t+1}^{1-\gamma_t}) \right]^{1/\theta_t} \right\}^{\theta_t/(1-\gamma_t)}, \qquad (34)$$

where $\theta_t \equiv (1-\gamma_t)/(1-1/\psi)$, ψ is the elasticity of intertemporal substitution (which is assumed to be constant through time), γ_t demotes the time-varying RRA coefficient, and β is a time discount factor. The specification for γ_t is given by $\gamma_t = \gamma_0 + \gamma_1 z_t$. Thus, risk aversion is driven by two distinct fundamental components: a constant or long-term component of risk aversion, γ_0, and a time-varying or cyclical component, $\gamma_1 z_t$, where γ_1 measures the sensitivity of RRA to the state variable.[23] It is possible to show that under these assumptions, and using the same framework as in Campbell (1993) and CV04, one obtains the following log stochastic discount factor:

$$m_{t+1} = \mathrm{E}_t(m_{t+1}) - \gamma_0 r_{t+1}^{CF} - \gamma_1 z_t r_{t+1}^{CF} + r_{t+1}^{H}, \qquad (35)$$

which leads to the expected return-covariance equation of the restricted scaled ICAPM in (12).[24]

Because DY is a countercyclical variable and positive shocks in *CPI* are associated with negative shocks in real wealth, one expects γ_1 to be positive in the versions with both state variables; that is, risk aversion increases in "bad times," similar to the concept of countercyclical risk aversion in Campbell and Cochrane (1999). In the case of the restricted scaled ICAPM, the positive and statistically significant estimates of $\gamma_{CFz} = \gamma_1$ confirm that risk aversion is an increasing function of either DY or *CPI*. On the other hand, the point estimates for $\gamma_{CF} = \gamma_0$, which correspond to the average RRA, $\mathrm{E}(\gamma_t) = \gamma_0 + \gamma_1 \mathrm{E}(z_t) = \gamma_0$, are not statistically different from zero. Hence, the time-varying or cyclical component of RRA largely dominates the average or long-term component.

I compute the implied RRA function over time, $\gamma_t = \gamma_0 + \gamma_1 z_t$, $z = DY$, *CPI*, by truncating the negative values of γ_t to zero. The resulting series have no apparent time trend, although there is a significant variability over time in γ_t. For both state variables there is a sharp decline in risk aversion during the 1990s in line with existent evidence of a possible stock bubble and risk-seeking behavior (Shiller 2000), which extended to the 2000s. On the other hand, the risk aversion estimates are significantly large in the 1970s and early 1980s, and there is also a modest increase in 2008. The implied time-series of RRA from the two state variables are positively correlated, with correlation coefficients of 59% and 58% when the test portfolios are S10 + BM10 + M10 and SM25, respectively. On the other hand, the estimated RRAs are weakly negatively correlated with the NBER dummy, with correlation coefficients of −36% and −29% when the conditioning variables are DY and *CPI*, respectively. Thus, these estimates for RRA seem to have some economic plausibility. However, we cannot interpret the resulting series of RRA as a strict estimate of implied RRA. First, given the linear specification used

[22] In related work Gordon and St-Amour (2000, 2004) specify preferences with explicit state-dependent RRA, in which γ_t evolves according to state Markov processes.

[23] Brandt and Wang (2003) specify a model where log RRA is a nonlinear function of innovations in both log consumption growth and inflation.

[24] The full derivation is available on request.

(required to obtain a linear pricing equation) one cannot constrain the risk aversion to be positive in all periods. Thus, this linear specification might be interpreted as a first-order approximation of a more general nonlinear process, as in Campbell and Cochrane (1999) or Brandt and Wang (2003). Secondly, there are other potentially relevant sources of countercyclical risk aversion (e.g., consumption growth).

6. Conclusion

This paper derives an ICAPM based on a conditional version of the two-beta ICAPM (BBGB) from Campbell and Vuolteenaho (2004). The beta of aggregate cash-flow news (or, alternatively, the respective risk price) is assumed to be time varying, similarly to the conditional CAPM literature. The scaled ICAPM contains three factors: revisions in future aggregate cash flows (cash-flow news), revisions in future market discount rates (discount-rate news), and a scaled factor that corresponds to the interaction of cash-flow news and the lagged state variable. In the empirical implementation of the scaled ICAPM, I use the market dividend yield, which is a popular predictor of the market equity premium and is also a countercyclical variable. The second conditioning variable employed is the CPI inflation, which is used by Brandt and Wang (2003) as a determinant of time-varying risk aversion.

The cross-sectional tests for 10 portfolios sorted on size, 10 portfolios sorted on book-to-market, and 10 portfolios sorted on momentum show that the scaled ICAPM explains relatively well the dispersion in excess returns on the 30 portfolios; thus, the three-factor model accounts for the value and momentum anomalies. The results for an alternative set of equity portfolios (25 portfolios sorted on size and momentum) show that the scaled ICAPM prices particularly well the momentum portfolios, with explanatory ratios varying between 74% and 86%. Furthermore, the risk price estimates show that the risk price for the scaled factor is significantly positive, indicating that the conditional beta/risk price of cash-flow news increases with the state variables; that is, the beta/price of risk rises in "bad times." Moreover, the scaled ICAPM compares favorably with alternative asset pricing models in pricing both sets of equity portfolios—the BBGB model, the unconditional CAPM, the conditional CAPM, the Fama and French (1993) three-factor model, and the Carhart (1997) four-factor model.

The fact that the scaled ICAPM is able to account for momentum profits represents an important innovation to most of the empirical tests of the cross-section of stock returns found in the literature. The scaled factor is decisive to account for the dispersion in average excess returns between past winner and past loser stocks. More specifically, past losers are riskier than past winners in periods of low realizations of the state variable, that is, in times with a low price of risk. Therefore, a time-varying cash-flow beta/price of risk provides a rational explanation for momentum.

Acknowledgments

This paper is a substantial revision from Chap. 1 of the author's doctoral dissertation at Nova School of Business and Economics (Universidade Nova de Lisboa). The author thanks advisors Pedro Santa-Clara and João Amaro de Matos for their guidance and support. He thanks the helpful comments from Levent Akdeniz, Aslihan Altay-Salih, Turan Bali, Parantap Basu, Laurent Calvet, Jean-Pierre Chateau, Engelbert Dockner, João Duque, Hayette Gatfaoui, Paskalis Glabadanidis, Xavier Gomez, Hui Guo, Thomas Hall, Paul Harrison, Jens Hilscher, Ked Hogan, Clemens Kool, Abraham Lioui, Marian Micu, Pierre Monnin, Jesper Rangvid, Patrick Rousseau, Ana Paula Serra, Michael Wickens, Jian Wu, Wei Xiong (the department editor), Kamil Yilmaz, an anonymous associate editor, and two anonymous referees. He also thanks seminar participants at the 2006 Portuguese Finance Network conference; the 2006 Money, Macro, and Finance Research Group Conference; the 2006 German Finance Association Meeting; the 2006 ASSET Meeting; the 2006 Paris Finance International Meeting; the 2007 Campus for Finance Research Conference; conference on "Return Predictability in the short and long run" (Copenhagen); Koç University; Bilkent University; Rouen School of Management; Barclays Global Investors (San Francisco); the 2007 Eastern Finance Association Meeting; and the 2008 FMA Meeting. The author especially thanks John Campbell for detailed comments and suggestions and Danielle Xu for helping in the construction of some of the equity portfolios. He also thanks Kenneth French and Robert Shiller for making available return data on their Web pages. Financial support was provided by Fundação para a Ciência e Tecnologia (Portuguese Government). A previous version circulated with the title "Intertemporal CAPM with Time-Varying Risk Aversion." All errors are the author's own.

References

Ang A, Bekaert G (2007) Stock return predictability: Is it there? *Rev. Financial Stud.* 20(3):651–708.

Ang A, Hodrick R, Xing Y, Zhang X (2006) The cross-section of volatility and expected returns. *J. Finance* 61(1):259–300.

Avramov D, Hore S (2008) Momentum, information uncertainty, and leverage—An explanation based on recursive preferences. Working paper, Hebrew University of Jerusalem, Jerusalem.

Avramov D, Chordia T, Jostova G, Philipov A (2007) Momentum and credit rating. *J. Finance* 62(5):2503–2520.

Bansal R, Dittmar R, Lundblad C (2005) Consumption, dividends, and the cross section of equity returns. *J. Finance* 60(4): 1639–1672.

Brandt M, Wang K (2003) Time-varying risk aversion and unexpected inflation. *J. Monetary Econom.* 50(7):1457–1498.

Brennan M, Wang A, Xia Y (2004) Estimation and test of a simple model of intertemporal capital asset pricing. *J. Finance* 59(4):1743–1775.

Campbell J (1987) Stock returns and the term structure. *J. Financial Econom.* 18(2):373–399.

Campbell J (1991) A variance decomposition for stock returns. *Econom. J.* 101(405):157–179.

Campbell J (1993) Intertemporal asset pricing without consumption data. *Amer. Econom. Rev.* 83(3):487–512.

Campbell J (1996) Understanding risk and return. *J. Political Econom.* 104(2):298–345.

Campbell J, Ammer J (1993) What moves the stock and bond markets? A variance decomposition for long-term asset returns. *J. Finance* 48(1):3–26.

Campbell J, Cochrane J (1999) By force of habit: A consumption-based explanation of aggregate stock market behavior. *J. Political Econom.* 107(2):205–251.

Campbell J, Shiller R (1988) The dividend price ratio and expectations of future dividends and discount factors. *Rev. Financial Stud.* 1(3):195–228.

Campbell J, Vuolteenaho T (2004) Bad beta, good beta. *Amer. Econom. Rev.* 94(5):1249–1275.

Campbell J, Chan Y, Viceira L (2003) A multivariate model of strategic asset allocation. *J. Financial Econom.* 67(1):41–80.

Campbell J, Giglio S, Polk C (2011) Hard times. Working paper, Harvard University, Cambridge, MA.

Carhart M (1997) On persistence in mutual fund performance. *J. Finance* 52(1):57–82.

Chen J (2003) Intertemporal CAPM and the cross-section of stock returns. Working paper, University of California, Davis.

Cochrane J (1996) A cross-sectional test of an investment-based asset pricing model. *J. Political Econom.* 104(3):572–621.

Cochrane J (2005) *Asset Pricing* (Princeton University Press, Princeton, NJ).

Cochrane J (2007) Financial markets and the real economy. Mehra R, ed. *Handbook of the Equity Premium* (Elsevier, Amsterdam), 237–326.

Cochrane J, Piazzesi M (2005) Bond risk premia. *Amer. Econom. Rev.* 95(1):138–160.

Daniel K, Moskowitz T (2011) Momentum crashes. Working paper, Columbia University, New York.

Epstein L, Zin S (1989) Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica* 57(4):937–969.

Epstein L, Zin S (1991) Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis. *J. Political Econom.* 99(2):263–286.

Fama E, French K (1988) Dividend yields and expected stock returns. *J. Financial Econom.* 22(1):3–25.

Fama E, French K (1989) Business conditions and expected returns on stocks and bonds. *J. Finance Econom.* 25(1):23–49.

Fama E, French K (1993) Common risk factors in the returns on stocks and bonds. *J. Finance Econom.* 33(1):3–56.

Fama E, French K (1996) Multifactor explanations of asset pricing anomalies. *J. Finance* 51(1):55–84.

Fama E, Schwert G (1977) Asset returns and inflation. *J. Financial Econom.* 5(2):115–146.

Ferson W, Harvey C (1999) Conditioning variables and the cross section of stock returns. *J. Finance* 54(4):1325–1360.

Gordon S, St-Amour P (2000) A preference regime model of bull and bear markets. *Amer. Econom. Rev.* 90(4):1019–1033.

Gordon S, St-Amour P (2004) Asset returns and state-dependent risk preferences. *J. Bus. Econom. Statist.* 22(3):241–252.

Guo H (2006) Time-varying risk premia and the cross section of stock returns. *J. Banking Finance* 30(7):2087–2107.

Hansen L (1982) Large sample properties of generalized method of moments estimators. *Econometrica* 50(4):1029-1054.

Harvey C (1989) Time-varying conditional covariances in tests of asset pricing models. *J. Financial Econom.* 24(2):289–317.

Hodrick R (1992) Dividend yields and expected stock returns: Alternative procedures for inference and measurement. *Rev. Financial Stud.* 5(3):357–386.

Hong H, Lim T, Stein J (2000) Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. *J. Finance* 55(1):265–295.

Jagannathan R, Wang Z (1996) The conditional CAPM and the cross-section of expected returns. *J. Finance* 51(1):3–53.

Jegadeesh N, Titman S (1993) Returns to buying winners and selling losers: Implications for stock market efficiency. *J. Finance* 48(1):65–91.

Jiang G, Lee C, Zhang Y (2005) Information uncertainty and expected returns. *Rev. Accounting Stud.* 10(2–3):185–221.

Johnson, T. (2002) Rational momentum effects. *J. Finance* 57(2): 585–608.

Keim D, Stambaugh R (1986) Predicting returns in the stock and bond markets. *J. Finance Econom.* 17(2):357–390.

Koijen R, Lustig H, Van Nieuwerburgh S (2012) The cross-section and time-series of stock and bond returns. Working paper, University of Chicago Booth School of Business, Chicago.

Lee C, Swaminathan B (2000) Price momentum and trading volume. *J. Finance* 55(5):2017–2069.

Lettau M, Ludvigson S (2001) Resurrecting the(C)CAPM: A cross-sectional test when risk premia are time-varying. *J. Political Econom.* 109(6):1238–1287.

Lewellen J (1999) The time-series relations among expected return, risk, and book-to-market. *J. Finance Econom.* 54(1):5–43.

Lintner J (1965) The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Rev. Econom. Statist.* 47(1):13–37.

Liu L, Zhang L (2008) Momentum profits, factor pricing, and macroeconomic risk. *Rev. Financial Stud.* 21(6):2417–2448.

Melino A, Yang A (2003) State-dependent preferences can explain the equity premium puzzle. *Rev. Econom. Dynam.* 6(4):806–830.

Merton R (1973) An intertemporal capital asset pricing model. *Econometrica* 41(5):867–887.

Petkova R, Zhang L (2005) Is value riskier than growth? *J. Financial Econom.* 78(1):187–202.

Shanken J (1990) Intertemporal asset pricing: An empirical investigation. *J. Econometrics* 45(1–2):99–120.

Shanken J (1992) On the estimation of beta pricing models. *Rev. Financial Stud.* 5(1):1–34.

Sharpe W (1964) Capital asset prices: A theory of market equilibrium under conditions of risk. *J. Finance* 19(3):425–442.

Shiller R (2000) *Irrational Exuberance* (Princeton University Press, Princeton, NJ).

White H (1980) A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48(4):817–838.

Yogo M (2006) A consumption-based explanation of expected stock returns. *J. Finance* 61(2):539–580.

Zhang L (2005) The value premium. *J. Finance* 60(1):67–104.

Zhang X (2006) Information uncertainty and stock returns. *J. Finance* 61(1):105–138.