Measuring the Time-Varying Risk-Return Relation from the Cross-Section of Equity Returns

Michael W. Brandt
Duke University and NBER

Leping Wang
Silver Spring Capital Management Limited

June 2010

Abstract

We use the structure imposed by Merton’s (1973) ICAPM to obtain monthly estimates of the market-level risk-return relationship from the cross-section of equity returns. Our econometric approach sidesteps the specification of time-series models for the conditional risk premium and volatility of the market portfolio. We show that the risk-return relation is mostly positive but varies considerably over time. It covaries positively with counter-cyclical state variables. The relationship between the risk premium and hedge-related risk also exhibits strong time-variation, which supports the empirical evidence that aggregate risk aversion varies over time. Finally, the ICAPM’s two components of the risk premium show distinctly different cyclical properties. The volatility component exhibits a counter-cyclical pattern whereas the hedging component is less related to the business cycle and falls below zero for extended periods. This suggests the market serves an important hedging role for long-term investors.

*We thank Jules van Binsbergen, seminar participants at the Vienna Symposium on Asset Management, and an anonymous referee for helpful comments.

†Fuqua School of Business, 1 Towerview Drive, Durham, NC 27708, USA. email: mbrandt@duke.edu.

‡Silver Spring Capital Management Limited, Level 69, Central Plaza, 18 Harbour Road, Hong Kong. email: leping.wang@silverspringfunds.com.
1 Introduction

The relation between expected return and risk lies at the heart of asset pricing. Unlike the literature on the cross-section of stock returns, little consensus has been reached on how the conditional expected return and volatility of the aggregate market portfolio covary over time.\(^1\) The existing empirical evidence is inconclusive and leads to inferences that are sensitive to the model specification. We propose a simple new approach that exploits the information contained in the cross-section of stock returns. This cross-sectional information leads to more precise and robust estimates of the intertemporal risk-return relation. We show that the risk-return relation varies considerably through time and covaries positively with counter-cyclical state variables. Furthermore, we find that although the risk premium is mostly positive and significant over time, it was negative in the early 1980s, late 1990s and mid 2000s, reflecting the eagerness of investors to bear stock market risk in these periods.

Empirical studies have focused on the intertemporal relationship:

\[
E_t[r_{m,t+1}] = \lambda \text{Var}_t[r_{m,t+1}] + \gamma \text{Cov}_t[r_{m,t+1}; s_{t+1}],
\]  

where \(r_{m,t+1}\) is the market excess return, \(s_{t+1}\) is a vector of state variables describing the investment opportunity set, and \(\lambda\) and \(\gamma\) relate the conditional expected return to the conditional volatility and hedging component, respectively. Perhaps surprisingly, there is conflicting evidence regarding the sign of the risk-return relation, \(\lambda\). French, Schwert, and Stambaugh (1987), Ghysels, Santa-Clara, and Valkanov (2005), Guo and Whitelaw (2005), and Ludvigson and Ng (2005) find a positive relation between the conditional expected return and volatility. In contrast, Campbell (1987), Breen, Glosten, and Jagannathan (1989), Glosten, Jagannathan, and Runkle (1993), Whitelaw (1994), and Brandt and Kang (2004) document a negative relation.

We argue that this conflicting evidence can be attributed to two important limitations of the existing empirical research designs. First, most studies assume that the risk-return relationship \(\lambda\) is constant through time. However, this assumption is inconsistent with the

\(^1\)Theories on the cross section of expected returns include the Capital Asset Pricing Model (CAPM) (Sharpe, 1964; Lintner, 1965), the intertemporal CAPM (Merton, 1973), and the arbitrage pricing theory (APT) (Ross, 1976). Although there is an ongoing debate about how well certain models fit the data, it is generally accepted that, at any given moment, higher risk, measured by the covariance of return with certain risk factors, should be associated with higher expected return. In contrast, since theory supports both a positive and a negative risk-return relation through time, this intertemporal relationship is primarily an empirical question. See Abel (1988) and Backus and Gregory (1992) for equilibrium models that support a negative risk-return relation.
empirical and theoretical literature arguing that the preferences of investors, which determine the intertemporal risk-return relation, change over the course of the business cycles (e.g., Campbell and Cochrane, 1999; Beber and Brandt, 2006)). Furthermore, using a vector autoregression (VAR) analysis, Whitelaw (1994) shows that the risk-return relation is not stable and that imposing a constant linear relation between return moments may lead to erroneous inferences. Harvey (2001) documents distinct counter-cyclical variation in the parameter relating the conditional mean and volatility. Consistent with this finding, Brandt and Kang (2004) and Ludvigson and Ng (2005) both present evidence of counter-cyclical variation in the Sharpe ratio of the market portfolio.

The second important limitation is that, to estimate the relation (1) using a time-series of market returns, auxiliary assumptions on the dynamics of the conditional moments, $E_t[r_{t+1}]$, $\text{Var}_t[r_{t+1}]$, and/or $\text{Cov}_t[r_{t+1}, s_{t+1}]$, have to be made, since none of these conditional moments are directly observable. French, et al. (1987) and Campbell and Hentschel (1992) employ a GARCH model for the volatility; Whitelaw (1994) and Ludvigson and Ng (2005) model the conditional mean and volatility as linear functions of macroeconomic and financial variables; Ghysels, et al. (2005) forecast monthly volatility with past daily squared returns using the mixed data sampling (MIDAS) method. Pastor, Sinha, and Swaminathan (2005) model the dynamics of the expected return with implied cost of capital computed from analyst forecasts. The hedging component, $\text{Cov}_t[r_{t+1}, s_{t+1}]$, is modeled as an EGARCH process in Scruggs (1998), and it is modeled as a linear function of exogenous state variables in Ghysels, et al. (2005) and Guo and Whitelaw (2005). All these model specifications are empirically and not theoretically motivated. As a consequence, the resulting answers are shaped as much by modelling assumptions as by the data. When the dynamics of one of the return moments is misspecified or when the risk-return relation is time-varying, the resulting inferences may be far from accurate.

We propose a cross-sectional approach that overcomes these problems. Our method builds on Merton’s (1973) ICAPM, and is inspired by the observation that the key parameters capturing the aggregate risk-return relation over time also drive the reward for covariance risk across assets. By using both the time-series and the cross-sectional dimension of the data, we provide a new perspective on the aggregate risk-return relation that has not been exploited by studies that use time series data alone. Furthermore, unlike many previous studies which require distributional assumptions necessary to estimate the parameters by maximum likelihood (MLE), we estimate the parameters in a more versatile and robust

\footnote{Since conditional volatilities are easier to estimate than expected returns, almost all the earlier papers proceed by specifying a conditional volatility process.}
generalized method of moments (GMM) framework. Not having to make distributional assumptions is particularly important in this case because the existing results are so sensitive to modeling assumptions. Although the robustness of the GMM method comes at the cost of statistical inefficiency, in our case this inefficiency is more than compensated by a much larger dataset, namely the whole cross-section of returns.

Using the cross-section of returns in a GMM framework has several advantages. First, as we no longer need to model and estimate the expected return or conditional volatility, we decrease the number of (ad hoc) assumptions. We start with Merton (1973)'s ICAPM and only need to make an assumption on the state variables that drive the investment opportunity set. This reduces the risk of model misspecification and thus leads to more reliable evidence.\(^3\) Second, the coefficient relating the expected return to risk is allowed to be time-varying and its time-variation is easy to estimate.

Our main analysis is based on the Fama-French three-factor model given its impressive performance in explaining the cross-section returns (Fama and French, 1993, 1995, 1996) and the recent evidence that the factor mimicking portfolios SMB (the return difference between small and large stocks) and HML (the return difference between high and low book-to-market ratio stocks) act as state variables describing the changing investment opportunity set (Liew and Vassalou, 2000; Vassalou, 2003). We model the dynamics of the time-varying risk-return relation as linear functions of common state variables, which include the dividend yield, the term spread, the default spread, and the riskfree rate. For all these state variables there is empirical evidence that they forecast returns. To efficiently exploit the cross-sectional information for inferring the intertemporal risk-return relation, we focus on the 25 portfolios sorted on size and the book-to-market ratio because this set of portfolios produces a reasonably large cross-sectional variation in average returns.

We present several important new and interesting findings regarding the intertemporal risk-return relation. First, although we find evidence of a positive risk-return relation under the assumption that this relation stays constant over time, the hypothesis of such a time-invariant risk-return relation is strongly rejected by the data. The positive risk-return relation identified in the studies cited above can only be interpreted as some

\(^3\)Unlike earlier studies based on the static CAPM, our investigation is based on the more general ICAPM. This modeling choice is motivated by the empirical evidence of a time-varying market risk premium (Chen, Roll, and Ross (1986), Shanken (1990), and Ferson and Harvey (1991)). As Scruggs (1998) and Guo and Whitelaw (2005) show, including the hedging terms of the ICAPM (in a time-series setting in their case) leads to a more robust positive relation between expected returns and risk. It needs to be noted, however that if the true asset pricing model deviates substantially from the ICAPM we assume, there still exists a potential risk of model misspecification.
time series average of the potentially changing relation, which may be different from the unconditional risk-return relation. When allowed to vary, the coefficient relating the market risk premium to the conditional market volatility exhibits a counter-cyclical pattern. It depends significantly on the dividend yield, the term spread, the default spread, and the Treasury bill rate. The coefficient is mostly positive and significant in the sample period from April 1953 to December 2008, as we would expect. However, it is significantly negative in the early 1980s, late 1990s and mid 2000s. This suggests that investors became increasingly willing to bear stock market risk during these periods. Furthermore, the time-variation in the coefficients relating the market risk premium to hedge-related risk also appears to be significant, consistent with the evidence on the time-variation of aggregate risk aversion.

Second, consistent with Guo and Whitelaw (2005) we find that the volatility and hedging components of the market risk premium are negatively correlated, with a sample correlation of $-0.52$. However, unlike them, we find that the volatility component exhibits a counter-cyclical pattern whereas the hedge component exhibits little covariation with the business cycle. Note that Guo and Whitelaw (2005) assume a time-invariant risk-return relation whereas we allow this coefficient to be changing over time. As a result, in their study, the market risk premium varies only with the changing risk components whereas in our study, it also varies with the changing risk-return relation. This risk-return relation depends on aggregate risk aversion. We also find that the hedging component of the risk premium takes negative values for extended periods. This stresses the fact that the stock market provides important hedging opportunities for long-term investors.

Third, we argue that the cross-sectional information plays an important role in finding the time-varying risk-return trade-off and that aggregating asset returns in the cross-section could cause significant information loss. This sheds light on why the earlier studies, which are based on time series data only, find it difficult to obtain robust and reliable evidence for a positive risk-return relation.

We check the robustness of our results with a nonparametric method to estimate the time-variation in the risk-return relation, using the innovations in the predictive variables as priced factors in the ICAPM (e.g., Campbell, 1993). The results remain the same.

As mentioned before, our approach is novel in that it adds information from the cross-section of asset returns to examine what is fundamentally a time-series issue. Most earlier studies concentrate on the time-series dimension alone. One study that does exploit cross-sectional data is Ludvigson and Ng (2005) who summarize the economic information of a large dataset of macroeconomic and financial series in a few factors through a principal
component analysis. They then use the extracted factors to estimate the conditional expected return and volatility. They argue that the conventional practice of estimating expected return and conditional volatility by conditioning on a few predetermined instruments may lead to omitted variable bias and thus spurious indications about the risk-return relation. With the effectively larger information set summarized in their estimated factors, they are able to better forecast the return moments and to find a positive risk-return relation. An advantage of their approach is the ability to summarize the desired panel data set in just a few factors regardless of the size of the dataset. The disadvantage is that it is unclear how many estimated factors should be included to sufficiently span the information present in the original dataset. As a consequence, some potentially useful information may still be lost. More importantly, their study does not allow the risk-return relation to be time-varying and relaxing the assumption of a constant risk-return relation is not trivial in their approach because (i) it complicates the time-series estimation and (ii) it introduces more parameters to model the time-variation which then leads to lower statistical power. Our GMM-based approach differs from Ludvigson and Ng (2005) in that it sidesteps the initial step of estimating the expected return and conditional variance. Instead, we estimate the risk-return relation directly by incorporating all the cross-sectional information in asset returns at once in a single step. More importantly, our approach allows for a time-varying risk-return relation, which can be estimated in a straightforward way. A second paper that incorporates cross-sectional information is Bali and Wu (2005). They also impose a time-invariant risk-return relation. Furthermore, the bivariate GARCH(1,1) process they assume for stock returns does not arise directly from a cross-sectional asset pricing restriction (the ICAPM in our case) and thus may be misspecified. In contrast, our GMM-based approach follows very naturally from the ICAPM.

The paper proceeds as follows. We present our cross-sectional approach in Section 2. Section 3 shows the empirical results. Section 3.1 discusses the data and the choice of priced factors. Sections 3.2 and 3.3 present the main results under the assumption of a constant risk-return relation and a time-varying risk-return relation, respectively. Section 3.4 and 3.5 examine the robustness of the results to the factor specification and parametrization. Section 3.6 highlights the importance of using cross-sectional information. Section 4 concludes.
2 Methodology

2.1 Cross-sectional and intertemporal risk-return relation

Merton (1973) analytically derives the intertemporal capital asset pricing model (ICAPM) in a continuous-time economy in which the investment opportunity set is time varying. Merton’s ICAPM generalizes the Sharpe (1964)-Lintner (1965) static CAPM derived in a single-period framework. Given the mounting empirical evidence on stochastic variation in investment opportunities, the ICAPM is a natural candidate to study the cross section of returns. The ICAPM can be written as:

\[
E_t[r^i_{t+1}] = \lambda_t \text{Cov}_t[r^i_{t+1}, r^m_{t+1}] + \gamma'_t \text{Cov}_t[r^i_{t+1}, s_{t+1}], \quad \text{for } i = 1, \ldots, N, \tag{2}
\]

where \( r^i_{t+1} \) is the return on asset \( i \) in excess of the risk-free rate, \( r^m_{t+1} \) is the market excess return, and \( s_{t+1} \) is a \( k \times 1 \) vector of state variables describing the investment opportunities in the economy. \( E_t \) and \( \text{Cov}_t \) are the expectation and covariance operators conditional on the information set at time \( t \), respectively.

By multiplying equation (2) by the asset weight \( w_i \) and summing over \( i \), we get an intertemporal risk-return relation for the aggregate market given by:

\[
E_t[r^m_{t+1}] = \lambda_t \text{Var}_t[r^m_{t+1}] + \gamma'_t \text{Cov}_t[r^m_{t+1}, s_{t+1}], \tag{3}
\]

where the first term on the right-hand side of equation (3) captures the volatility and the second term captures the hedging component of the risk premium. Equation (3) has been the focus of much of the existing literature, as described in the introduction. This literature examines how the expected return and conditional volatility of the aggregate market as well as its hedge-related risk are correlated through time. Some studies assume away the hedging component, \( \gamma'_t \text{Cov}_t[r^m_{t+1}, s_{t+1}] \), following Merton’s argument that the hedging component is negligible if the investment opportunity set is static or if investors have logarithmic utility. As equation (3) is written in aggregate terms (e.g., the market return \( r^m_{t+1} \)), those studies all attempt to infer the nature of \( \lambda_t \) from aggregate time series data, such as market returns or other financial instruments. However, the results from these time-series analyses are subject to debate as they are highly sensitive to the model specification. We provide a new approach that is immune to this issue.

It is important to note that the ICAPM given in equation (2) explains the variation in expected returns across assets, conditional on the information set in time \( t \). The aggregate
model in equation (3), on the other hand, focuses on how the expected return and volatility of the aggregate market as well as its hedging component are related over time. In fact, equation (2) is a conditional factor pricing model with the market return and the state variables as its priced factors. It suggests that assets are priced according to their conditional covariance with the market portfolio as well as their conditional covariance with certain “hedge” portfolios which are correlated with changes in the investment opportunity set. Equation (3) suggests that the market’s compensation for risk is indeed changing over time and depends on the market risk level.

Comparing these two equations, it is interesting to see that, in equation (2) the coefficients $\lambda_t$ and $\gamma_t$ measure the expected excess return per unit of covariance of the asset’s return with the corresponding factors while in equation (3) the same coefficients measure, over time, the change in risk premium for each unit change in the conditional volatility of the market and hedge-related risk. The fact that equations (2) and (3) share the same coefficients provides the key motivation to examine the risk return relation by means of a cross-sectional analysis. Before describing the implementation of the idea, it is important to note that most earlier studies start with an empirically specified intertemporal risk-return relation for the aggregate market and impose some auxiliary assumptions on the dynamics of the market return moments for estimation purpose. In contrast, we derive the relation (3) specifically from the analytical result ICAPM (2) and conduct empirical study based on that. This feature could help us reduce model misspecification concern from the ad hoc empirical assumptions made in the earlier studies. However, to the extent that the true asset pricing model deviates from the ICAPM (2), the true intertemporal risk-return relation for the aggregate market could itself deviate from (3). In such case, one can still perform empirical analysis on (3) as in the earlier studies but it should be pointed out that (3) will be a purely empirical specification and may not represent the true functional form of the intertemporal risk-return relation for the aggregate market.

To implement this idea, note that the cross-sectional model (2) can be written as:

$$E_t[r_{t+1}^i] = \text{Cov}_t[r_{t+1}^i, \lambda_t r_{t+1}^m + \gamma_t^s s_{t+1}].$$

(4)

If we let:

$$m_{t+1} = \alpha_t - \lambda_t r_{t+1}^m - \gamma_t^s s_{t+1},$$

(5)

---

4Model (2) can be equivalently expressed in the expected return-beta form. Let $f_{t+1}$ be a vector with the elements $r_{t+1}^m$ and $s_{t+1}$, and $\delta_t$ the vector of $\lambda_t$ and $\gamma_t$, then (2) can be written as $E_t[r_{t+1}^i] = \text{Cov}_t[r_{t+1}^i, f_{t+1}] \delta_t$. Or equivalently it can be expressed as $E_t[r_{t+1}^i] = B_t \psi_t$ where $B_t = \text{Cov}_t[r_{t+1}^i, f_{t+1}] \text{Var}_t[f_{t+1}]^{-1}$, is the conditional beta matrix, and $\psi_t = \text{Var}_t[f_{t+1}]^{1/2}$ is the factor premium.
where $\alpha_t$ is a normalizing factor such that $E_t[m_{t+1}] = 1$, we have:

$$E_t[m_{t+1}r_{t+1}^i] = 0. \quad (6)$$

Equation (6) shows that the expected return-covariance relation of the ICAPM can also be written as a stochastic discount factor (SDF) model. The SDF approach is simple and universal as it incorporates various modern asset pricing models in a unified framework.\(^5\)

We can now examine the intertemporal risk-return trade-off, measured by $\lambda_t$ and $\gamma_t$, in a GMM estimation framework based on the orthogonality conditions implied by equation (6) using the cross-section of asset returns. To estimate the intertemporal relation (3), the earlier studies proceed by imposing additional dynamics on the conditional moments $E_t[r_{t+1}^m]$, Var$_t[r_{t+1}^m]$, or Cov$_t[r_{t+1}^m, s_{t+1}]$ to have a fully specified model (e.g., the GARCH-in-Mean model used in French, Schwert, and Stambaugh (1987) assumes a GARCH process for Var$_t[r_{t+1}^m]$). In contrast, our econometric approach allows us to avoid making those time-series assumptions for the moments in equation (3) via a cross-sectional analysis based on (6). More importantly, it does not require $\lambda_t$ and $\gamma_t$ to be constant. Note that the ICAPM given in equation (2) does not yield a fully specified SDF in equation (5) due to the time-varying nature of the coefficients. However, the SDF framework does provide an intuitive and simple framework for our estimation approach.

### 2.2 Time-varying risk-return relation

We have shown that $\lambda_t$ and $\gamma_t$ measure not only the aggregate intertemporal return-to-risk ratio, as in equation (3), but also the cross-sectional reward-to-covariability ratio, as in (2). There is ample evidence that the reward-to-covariability, which is closely related to the factor premium in an equivalent expected return-beta formulation, varies through time (e.g., Harvey, 1989). When the CAPM holds, $\lambda_t$ is simply the ratio of the conditional expected excess return on the market portfolio and the conditional variance of the market portfolio. The coefficient $\lambda_t$ thus represents the compensation that the representative investor requires for a one unit increase in the variance of the market return. This compensation strongly depends on the aggregate level of risk aversion (see Merton, 1980).\(^6\) The literature provides empirical support that the time-variation in risk aversion is economically and statistically

---

\(^5\)See Cochrane (2005) for discussions on general equivalence between SDF and beta representations. In particular, standard asset pricing models, including the CAPM, ICAPM, APT, and consumption-based CAPM, can all be expressed in the form of an SDF model.

\(^6\)As Harvey (1989) argues, $\lambda_t$ can be interpreted as the relative risk aversion of the representative agent under some (strong) assumptions about the consumption process, e.g., iid consumption.
significant and relates to business cycles (Brandt and Wang, 2003; Beber and Brandt, 2006). In addition, Whitelaw (1994) shows that the correlation between the first two return moments is not constant over time but varies from large positive to large negative values.

This evidence suggests that at the market level, the expected return, the conditional variance, and the hedging component, are strongly related but in a time-varying fashion. To model the time-variation in \( \lambda_t \) and \( \gamma_t \), we assume for now that they are linear in the state variables (we relax the linearity assumption in Section 3.5). Let \( z_t \) be a vector of \( q \) market-wide state variables in the information set at time \( t \), such that \( \mathbb{E}_t[y_{t+1}] = \mathbb{E}[y_{t+1} | z_t] \) for any random payoff vector \( y_{t+1} \). Following Cochrane (1996), we choose the parametrization:

\[
\alpha_t = a_0 + a' z_t \quad \text{and} \quad (\lambda_t, \gamma_t)' = b_0 + b' z_t,
\]

where \((a_0, a')'\) and \((b_0, b')'\) are a \((q + 1)\) vector and a \((q + 1) \times (k + 1)\) matrix of parameters.

This linear parameterization of the risk-return relation is convenient because the resulting SDF is linear in the parameters, which allows for closed-form GMM parameter estimates. The details of the GMM procedure are as follows. We first write the conditional model in equation (6) and the normalizing condition in a vector representation:

\[
\mathbb{E}_t \left[ m_{t+1} \begin{pmatrix} r_{t+1} \\ 1 \end{pmatrix} \right] = \begin{pmatrix} 0 \\ 1 \end{pmatrix},
\]

where \( r_{t+1} \) is a column vector of the returns \( r_{t+1}^i \), for all \( i \), \( \bar{0} \) denotes a column vector of zeros, and the last equation of this system is the normalizing condition for \( m_{t+1} \). The conditional moments can be written into their equivalent unconditional counterparts by expanding the set of moments with a vector of conditioning variables and applying the law of iterated expectations:

\[
\mathbb{E}[m_{t+1} x_{t+1}] = p,
\]

where

\[
x_{t+1} = \begin{pmatrix} r_{t+1} \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ z_t \end{pmatrix} \quad \text{and} \quad p = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ z_t \end{pmatrix}.
\]

The discount factor can also be written into the following scaled factor representation:

\[
m_{t+1} = b' f_{t+1},
\]
where $b$ is a constant parameter vector and:

$$
\begin{align*}
    f_{t+1} &= \left( \begin{array}{c}
    1 \\
    r_{t+1}^m \\
    s_{t+1}
    \end{array} \right) \otimes \left( \begin{array}{c}
    1 \\
    z_t
    \end{array} \right)
\end{align*}
$$

(12)

are the scaled factors.

Applying standard GMM to estimate the coefficients $b$ using the moment conditions in equation (9), the second-stage estimates and their standard errors are given by:

$$
\hat{b} = (d'S^{-1}d)^{-1}d'S^{-1}E_T[p],
$$

(13)

and

$$
\text{Cov} \left[ \hat{b} \right] = \frac{1}{T} (d'S^{-1}d)^{-1},
$$

(14)

where $E_T[\cdot]$ denotes the sample counterpart of the unconditional expectation (i.e., the time-series average), $d' = E_T(f_x')$, and $S$ is the optimal weighting matrix, which is also replaced by its sample counterpart.\(^7\)

As we mentioned in the introduction, our GMM approach does not impose distributional assumptions on the asset returns or state variables. Beside the ICAPM pricing structure, the only auxiliary assumption we make is the dependence of $\lambda_t$ and $\gamma_t$ on the vector of observable state variables $z_t$. Applied to the same dataset, GMM is generally less efficient than MLE, but this does not mean that our approach leads to less precise inference compared to related studies that rely on MLE. To the contrary, our approach is likely to produce much more accurate estimates. The reason is that we use a much larger dataset than related studies. Our GMM approach allows us to use the entire cross-section of asset returns to estimate the time-series relation between the conditional mean and volatility of the market portfolio, not just the single time-series of market returns.

### 3 Empirical results

#### 3.1 Data

To implement our approach we first need to chose the state variables or risk factors $s_{t+1}$ in the ICAPM relation (2). Since theory provides relatively little guidance in this regard,

\(^7\)See Cochrane (2005) for further details on GMM estimation.
we base our choice of factors on the empirically successful and hence widely applied Fama-French three-factor model. The three factors are the excess return of the market portfolio \((r^m_{t+1})\), the return of a portfolio long in small firms (i.e., low market capitalization) and short in large firms (SMB), and the return of a portfolio long in high book-to-market stocks and short in low book-to-market stocks (HML). Fama and French (1993, 1995, 1996) show that an unconditional version of their three-factor model explains much of the cross-sectional variation in average returns of portfolios sorted by size and book-to-market equity ratio, which suggests that its conditional counterpart provides a reliable basis for our intertemporal study.\(^8\) Fama and French further argue that HML and SMB act as state variables that predict future changes in the investment opportunity set in the context of Merton’s ICAPM, suggesting a risk-based explanation for the model’s empirical success. We mainly conduct our analysis using \(HML_{t+1}\) and \(SMB_{t+1}\) as the state variables \(s_{t+1}\) in the ICAPM relation (2) and its time-series implication (3). We explicitly examine the sensitivity of our results to this choice of factors in Section 3.4.

We capture the cross-section of asset returns through the standard 25 equity portfolios sorted by market capitalization (ME) and book-to-market ratio (BE/ME). Fama and French (1995) show that these portfolios produce a large two-dimensional spread in average realized returns. They also show that sorting on ME and BE/ME subsumes sorts based on leverage, earnings-to-price (EP) and other firm characteristics. As a robustness check, we also examine in Section 3.5 portfolios sorted univariately either by ME or by BE/ME.

Finally, we choose for the state variables \(z_t\), which capture the time-variation, if any, in the return-to-risk ratios \(\lambda_t\) and \(\gamma_t\), the dividend yield, default spread, term spread, and risk-free rate. For all of these four business cycle related variables there exists substantial empirical evidence that they forecast future market returns. The dividend yield is the sum of dividends over the past 12 months divided by the current index value. The default spread is the yield difference between the Moody’s Baa and Aaa rated bonds, the term spread is the yield difference between a ten-year and a one-year government bond, and the risk-free rate is proxied by the 30-day Treasury bill rate.

We obtain the stock return data from CRSP and Ken French’s website. The bond yields are taken from the Federal Reserve Bank reports. The sample period is April 1953 through December 2008 since the term spread is available only from April 1953 onwards. Figure 1 plots the time series of the state variables \(z_t\). NBER business cycle peaks are marked by

\(^8\)See Cochrane (2005) for a comprehensive discussion about the relation between unconditional and conditional versions of factor pricing models. In particular, he shows that even though an unconditional factor model implies a conditional factor model, the reverse implication does not necessarily hold.
dashed lines and business cycle troughs are marked by solid lines. The figure confirms that the time-variation in all four variables is to some extent related to the business cycle.

### 3.2 Constant risk-return relation

Since the literature has focused primarily on a constant risk-return relation, with widely diverging results, we start our analysis by investigating this special case as a benchmark. In other words, we first assume that \( \lambda_t \) and \( \gamma_t \) are constant. The cross-sectional information provides us with sufficient statistical power to identify a positive intertemporal risk-return relation, but it also allows us to strongly reject this special case in favor of a time-varying intertemporal risk-return relation.

If in addition to \( \lambda_t = \lambda \) and \( \gamma_t = \gamma \) we also assume that \( \alpha_t = \alpha \), equation (6) can be conditioned down to obtain:

\[
E[m_{t+1}r_{t+1}^i] = 0, \tag{15}
\]

where \( m_{t+1} = \alpha - \lambda r_{t+1}^m - \gamma s_{t+1} \). This equation can then be rewritten in expected return-covariance form:

\[
E[r_{t+1}^i] = \lambda \text{Cov}[r_{t+1}^i, r_{t+1}^m] + \gamma' \text{Cov}[r_{t+1}^i, s_{t+1}], \tag{16}
\]

which in turn can be estimated by the standard two-stage Fama-Macbeth (1973) procedure.

We apply two methods to estimate the covariance in the first-stage time-series regression. We use either a full-sample regression as in Lettau and Ludvigson (2001), for instance, or a 60-month rolling sample regression, as in Fama and MacBeth (1973). We use the GMM method for estimation and to be consistent with the Fama-Macbeth estimation, we do not use any instruments in the GMM estimation.

Table I presents the estimates of the intertemporal risk-return relation when this relation is assumed to be constant through time. The estimates obtained with all three methods are consistently positive and significant. For example, the Fama-MacBeth estimates for \( \lambda \) are 3.93 using the full-sample and 3.85 using the rolling-sample regressions in the first step. Both of these numbers are statistically significant. Accounting for the error-in-variables problem, the GMM approach leads to an estimate of 4.27 with a \( t \)-statistic of 3.91.

Our evidence on a positive risk-return relation is consistent with a number of earlier studies that draw inference from aggregate time series data with a variety of volatility specifications. Those studies mainly differ in their methods to measure the unobservable
return volatility. However, there are many other studies using yet other return volatility specifications that find the opposite result. From the summaries by Glosten, et al. (1993) and Harvey (2001) we can infer that the evidence about the sign of the risk-return relation shows a wide dispersion of results and is highly sensitive to the specification of the volatility process. Our evidence, however, is not subject to the critique of a potentially mispecified volatility model because our approach does not require any specification for the dynamics of return moments and is solely based on the cross-sectional implications of the ICAPM. As long as our assumed pricing factors are correct, our results provide reliable estimates of the risk-return relation.

The intertemporal relation between the market risk premium and the market covariance with SMB and HML are also estimated to be positive, as measured by $\gamma_1$ and $\gamma_2$. The GMM estimates for $\gamma_1$ and $\gamma_2$ are 2.41 and 8.70 with t-statistics of 1.70 and 5.63, respectively. These estimates are consistent with Petkova (2006), who finds that HML is a significant factor in the cross-section of asset returns but SMB is not. Our results stress the importance of the hedging component in the intertemporal risk return tradeoff.

The estimates for $\lambda$ and $\gamma$ reported above provide evidence on the sign of the aggregate intertemporal risk-return relation if we are willing to assume that the relationship is constant. However, the analytically derived relationship given by (3) imposes no restriction on the time-variation of the coefficients relating risk to return. In fact, Merton (1980) argues that these coefficients relate to the risk aversion of the representative investor, for which significant time-variation has been documented in numerous studies (e.g., Campbell and Cochrane, 1999; Brandt and Wang, 2003; Beber and Brandt, 2006)). Although these studies provide indirect evidence against the hypothesis that the intertemporal risk-return relation stays constant over time, our GMM-based cross-sectional method allows us to directly test this hypothesis using the overidentifying GMM restrictions. We report the $p$-values for this test in the table. The tests decisively reject that the intertemporal relation is constant over time.

---


10For instance, Breen, Glosten, and Jagannathan (1989) and Whitelaw (1994) use instruments such as the riskfree rate to forecast volatilities and find a negative risk-return relation.
3.3 Time-varying risk-return relation

Given the direct and indirect evidence against a constant risk-return relation, our above estimates can only be viewed as some time series average of the true time-varying relation. In this section, we allow $\alpha_t$, $\lambda_t$, and $\gamma_t$ to change over time. The most important result we find is that $\lambda_t$, which measures the risk-return relation, is indeed changing over time and shows a counter-cyclical pattern, but remains positive and significant at most times.

Panel A of Table II presents our results for the general model (3) in which the coefficients are assumed to be linear in the state variables using the GMM approach described above based on the Fama and French’s 25 BE/ME-size portfolios. The $t$-statistics are reported in brackets. The estimates show that $\lambda_t$, $\gamma_{1t}$, $\gamma_{2t}$ are all significantly dependent on at least one state variable: $\lambda_t$ varies significantly with all the four state variables; $\gamma_{1t}$ varies significantly with the dividend yield, default spread and Treasury bill rate; $\gamma_{2t}$ varies significantly with the term spread.

Figure 2 plots the time series of the coefficient $\lambda_t$ (solid lines), which measures the risk-return relation, along with 95 percent confidence intervals (dotted lines). Business cycle peaks and troughs are marked by dashed and solid vertical lines, respectively. The intertemporal relation exhibits significant variation over time. It is estimated to be positive and mostly significant in 467 months, about 70 percent of the total sample period. Among the rest of the sample period, $\lambda_t$ is negative but mostly insignificant, although it occasionally takes significant negative values in the early 1980s, late 1990s, and mid 2000s.

Lewellen, Nagel, and Shanken (2006) argue that the strong covariance structure of the Fama-French 25 BE/ME-size portfolios can cause misleading results in asset pricing tests. To help alleviate this problem, they advocate to expand the Fama-French 25 BE/ME-size portfolios to include 30 industry portfolios. Motivated by their argument, we also perform our cross-sectional analysis using this expanded portfolio set. We report the corresponding results in panel B of Table II and panel B of Figure 2. We find that, in our context of the intertemporal risk-return relation, the results are not sensitive to the inclusion of the industry portfolios. The parameter estimates in panel B of Table II are largely similar in magnitude to those in panel A, still implying significant time-variations in $\lambda_t$, $\gamma_{1t}$, and $\gamma_{2t}$. The plots in panel B of Figure 2 are also qualitatively similar to those in panel A. We therefore focus our subsequent discussions on the results for the Fama-French 25 BE/ME-size portfolios.

Our evidence confirms that the aggregate market risk premium and the aggregate market risk level are positively correlated, however the link between them is time-varying. The evidence is new and has not been documented in the literature. It is, however, consistent
with our intuition for time-variation in aggregate risk aversion. From Merton’s derivation of the ICAPM, $\lambda_t$ is related to the aggregate risk tolerance. If the aggregate risk aversion remains the same over time, it is generally expected that the equilibrium expected return on the market is an increasing function of the risk of the market and thus $\lambda_t$ is anticipated to be positive and constant. However, if either changes in preferences or changes in the distribution of wealth are such that aggregate risk aversion is lower when the market is riskier, then a higher market risk level can be associated with a lower risk premium. This can, in principle, explain the occasional negative values of $\lambda_t$ shown in Figure 2.

It is often argued in the literature that the level of aggregate risk aversion is time-varying.\(^{11}\) In a habit formation model, the representative agent’s risk aversion changes with the difference between consumption and the habit-level of consumption. This habit-level is based on past consumption. Brandt and Wang (2003) document a positive and significant correlation between aggregate risk aversion and unexpected inflation. They point out an alternative explanation for time-varying aggregate risk aversion based on heterogeneous preferences and changes in the cross-sectional distribution of real wealth due to inflation shocks. Since aggregate risk aversion relates to consumption growth and inflation, both of which exhibit a business cycle pattern, it is not surprising that risk aversion exhibits a business cycle pattern as well. Indeed, Brandt and Wang (2003) find that periods of strong economic conditions to be associated with low or falling risk aversion while recessions are associated with high or rising risk aversion.

Given (i) the counter-cyclical behavior of aggregate risk aversion, and (ii) the strong dependence of the risk premium per unit of market risk on the contemporaneous change in aggregate risk aversion, we expect that the coefficient $\lambda_t$ tends to be low or decreasing when economic conditions are strong and high or increasing during recessions. We indeed find in Figure 2 that $\lambda_t$ varies counter-cyclically. During the economic contractions 1957.8-1958.4, 1980.1-1980.7 and 2001.3-2001.11, $\lambda_t$ rises sharply. During the economic expansions 1954.5-1957.8, 1970.11-1973.11, 1980.7-1981.7, 1991.3-2001.3, 2001.11-2007.12, $\lambda_t$ drops sharply. In the remaining contractions and expansions, $\lambda_t$ does not change much. The significantly negative value of $\lambda_t$ during the early 1980s, late 1990s and mid 2000s suggests that investors become increasingly willing to bear stock market risk in these periods. Further, the counter-cyclical pattern of $\lambda_t$ is consistent with the counter-cyclical behavior of the conditional Sharpe ratio (the ratio of conditional expected excess return to conditional standard deviation) as in Brandt and Kang (2003), Lettau and Ludvigson (2003), and Ludvigson and Ng (2005).

\(^{11}\)There are a number of papers which model changing risk aversion. Examples include Constantinides (1990), Campbell and Cochrane (1999), and Brandt and Wang (2003).
It is also consistent with the finding that the time-variation in expected asset returns is strongly counter-cyclical (Fama and French (1989)), whereas the conditional volatility seems to exhibit a much weaker counter-cyclical pattern (Ludvigson and Ng (2005)).

Panel A of Figure 2 also shows time-variation in the relation between the market premium and the market covariances with the factors SMB and HML. Even though the estimation errors are larger, the estimates of \( \gamma_{1t} \) and \( \gamma_{2t} \) are mostly positive and are occasionally significant. The time series averages of \( \gamma_{1t} \) and \( \gamma_{2t} \) are 1.98 and 6.52, respectively, which are close to the estimates in Table I where we assume a constant relation.

As Fama and French argue, SMB and HML proxy for the state variables that describe the time-variation in the investment opportunity set, which suggests a risk-based explanation for their three-factor model. This point is also supported by the empirical evidence presented in Petkova (2006) and Campbell and Vuolteenaho (2004).\(^{12}\) Thus, depending on a long-term investor’s risk aversion, an increase in the covariance of the stock market returns with SMB or HML would make the stock market less (or more) attractive as a hedging tool,\(^{13}\) in which case a higher (or lower) market premium would be demanded. So, the time-variation in \( \gamma_{1t} \) and \( \gamma_{2t} \) in Figure 2 also suggests that aggregate risk aversion changes over time.

Scruggs (1998), Ghysels, et al. (2005), and Guo and Whitelaw (2005) also take into account the hedging component when investigating the risk-return relation and find that it plays an important role. However, they all impose the restriction that hedge-related risk has a constant market premium over time. Our results in Figure 2 and Table II shows that the market’s reward for hedge-related risk depends significantly on the state of the economy.

Given the estimates of the time-varying coefficients (\( \hat{\lambda}_t \) and \( \hat{\gamma}_t' \)) in (3), we can construct estimates of the volatility component (\( \hat{\lambda}_t \hat{\text{Var}}_t[r_{t+1}^m] \)) and hedging component (\( \hat{\gamma}_t' \hat{\text{Cov}}_t[r_{t+1}^m, s_{t+1}] \)) of the market risk premium. In other words, we estimate the conditional volatility, \( \hat{\text{Var}}_t[r_{t+1}^m] \), and the conditional covariances, \( \hat{\text{Cov}}_t[r_{t+1}^m, s_{t+1}] \) in the conventional way through linear regressions (e.g., Campbell (1987) and Whitelaw (1994)). In particular, we first obtain \( \hat{E}_t[r_{t+1}^m] \) by regressing \( r_{t+1}^m \) on the state vector \( z_t \). The conditional return volatility, \( \hat{\text{Var}}_t[r_{t+1}^m] \), can then be estimated by regressing \( [r_{t+1}^m - \hat{E}_t[r_{t+1}^m]]^2 \) on the state variables. We

\(^{12}\) Vassalou (2003) shows that SMB and HML appear to contain mainly news related to future GDP growth and that accounting for macroeconomic risk reduces the informational content of SMB and HML. Liew and Vassalou (2000) show that SMB and HML have additional forecasting power for future GDP growth, independent of the information contained in the market factor, even in the presence of popular business cycle variables.

\(^{13}\) A long-term dynamic investor with power utility and risk aversion larger than one typically allocates more to assets whose returns are negatively correlated with the future investment opportunity set. The converse holds for an investor with power utility and a risk aversion lower than one.
estimate the conditional covariances in a similar way. This approach has been used in Shanken (1990) and is also closely related to Whitelaw (1994), who models the conditional standard deviation as a linear function of the state variables.

Figure 3 plots the time series of the estimated volatility component, $\hat{\lambda}_t \text{Var}_t[r_{t+1}^m]$, and hedging component, $\hat{\gamma}_t \text{Cov}_t[r_{t+1}^m, s_{t+1}]$. Consistent with Guo and Whitelaw (2005), we find that the two series are negatively correlated, with a sample correlation of $-0.52$, suggesting that omitting the hedging component, as done by earlier studies (e.g., French, et al., 1987), leads to a downward bias on the coefficient estimate of the volatility component. However, unlike them, we find the volatility component to exhibit a counter-cyclical pattern, which is not surprising given the counter-cyclical pattern of aggregate risk aversion and the weak cyclical variation in the conditional volatility $\text{Var}_t[r_{t+1}^m]$ (Ludvigson and Ng, 2005). Note that in Guo and Whitelaw (2005) the volatility component of the risk premium changes only with the conditional volatility itself due to their assumption that $\lambda_t$ is constant. In our study it also varies with $\lambda_t$, which is related to the aggregate level of risk aversion. The hedging component exhibits a much weaker business cycle pattern. It falls below zero for extended periods which suggests that during these times the stock market becomes more attractive to long-term investors because of the hedging opportunities it provides. Occasionally, the sum of these two components is negative, leading to negative estimates of the expected excess return. Although a negative risk premium is not impossible from a theoretical point of view according to Whitelaw (2000), these negative estimates could certainly also be caused by the inevitable estimation errors that is common in linear empirical specifications (e.g., Harvey, 2001; Ludvigson and Ng, 2005).

Because of the dual roles of $\lambda_t$ and $\gamma_t$ in equations (2) and (3), all the above discussions about the time-variation in $\lambda_t$ and $\gamma_t$ are equally applicable to the time-varying reward-to-covariabilities ratios, or covariance premium, for cross-sectional returns. Since the present paper attempts to resolve the conflicting evidence on the intertemporal risk-return relation, we mainly focus on the interpretation of $\lambda_t$ and $\gamma_t$ as measures for the aggregate intertemporal risk-return relation.

### 3.4 Are the results sensitive to the factor specification?

Campbell (1996) points out that the pricing factors in the ICAPM should not be selected according to important macroeconomic variables, but instead, should relate to innovations in the state variables that forecast stock market returns. Given the critical role of the pricing model to the reliability of our results, we conduct in this section an analysis based on an
alternative specification of the state variables used in the ICAPM. We find that the resulting evidence is consistent with our earlier results based on the Fama-French model.

Stock return predictability is illustrated in the literature with a predictive regression of the form:

\[
r_{m,t+1} = v + \delta z_t + \varepsilon_{m,t+1},
\]

where \( z_t \) is a \( q \times 1 \) vector of predictive variables that are observed at time \( t \), \( v \), and \( \delta \) are coefficients, and \( \varepsilon_{m,t+1} \) is the innovation in the market return. The predictive vector \( z_t \) is in general stochastic and may be correlated with past return innovations. For example, the dividend yield, which is often used as a predictor in many studies, is likely to be negatively correlated with the return innovation contemporaneously. A common approach to allowing for the stochastic properties of the predictors is to assume that \( z_t \) obeys a first-order vector autoregressive (VAR) process (e.g., Kandel and Stambaugh (1996) and Barberis (2000)). That is, we assume:

\[
z_{t+1} = C_0 + C z_t + \varepsilon_{t+1}^z,
\]

where \( C_0 \) is a \( q \times 1 \) vector, \( C \) is a \( q \times q \) matrix, and \( \varepsilon_{t+1}^z \) is a vector of innovations to the predictor variables. Like before, we assume that the predictive vector, \( z_t \), includes the dividend yield, term spread, default spread, and Treasury bill rate. As the evolution of \( z_t \) fully determines the distribution of the return \( r_{m,t} \), the state vector \( z_t \) fully describes the investment opportunity set.

Table III presents the estimation results for the predictive regression. Each row of panel A corresponds to a different equation of the model. The first five columns report coefficients on the five predictive variables: a constant and the lagged values of the dividend yield, term spread, default spread, and Treasury bill rate. The standard errors are reported in parentheses below the estimates, and the \( R^2 \) values are presented for each equation in the model.\(^{14}\) Panel B reports the correlation matrix of the innovations, with the standard deviations on the diagonal.

The first row of panel A shows the monthly forecasting equation for market returns. The results are consistent with those found in the literature. The risk-free interest rate negatively predicts the market return while the dividend yield, term spread, and default spread all enter with positive signs. Both the dividend yield and the risk-free rate exhibit significant forecasting power for future market returns. The forecasting equation has a modest \( R^2 \) of three percent, which is reasonable for a monthly model. The remaining rows of panel A

\(^{14}\) The results are subject to the estimation biases pointed out in Stambaugh (1999).
summarize the dynamics of the predictive variables. The predictors are highly persistent, with persistence coefficients of 0.98, 0.96, 0.99, and 0.92, for the dividend yield, term spread, default spread, and risk-free rate, respectively. The $R^2$ values are all above 90 percent for the evolution of the predictors. The standard deviation of market return shocks is about four percent per month. The shocks to the dividend yield and the risk-free rate are much less volatile. Further, there is a large negative correlation between the innovations to the dividend yield and the market return.

According to Campbell (1993), if the investment opportunity set varies over time, an asset’s exposure to these changes are important determinants, in addition to the market beta, of expected returns. That is, $\varepsilon_{t+1} \equiv (\varepsilon^m_{t+1}, \varepsilon^z_{t+1})$ should be the pricing factor. We can write the ICAPM given in (2) as:

$$E_t[r^i_{t+1}] = \lambda_t \text{Cov}_t[r^i_{t+1}, \varepsilon^m_{t+1}] + \gamma_i' \text{Cov}_t[r^i_{t+1}, \varepsilon^z_{t+1}].$$

(19)

A regression setting essentially similar to (19) is used in Petkova (2006). As shown earlier, equation (19) translates into a SDF:

$$m_{t+1} = \alpha_t - \lambda_t \varepsilon^m_{t+1} - \gamma_i \varepsilon^z_{t+1}.$$  

(20)

First we rewrite the SDF as $m_{t+1} = \tilde{\alpha}_t - \lambda_t r^m_{t+1} - \gamma_i' z_{t+1}$, where $\tilde{\alpha}_t = \alpha_t + \lambda_t E_t[r^m_{t+1}] + \gamma_i' E_t[z_{t+1}]$. Again we use the parametrization $\tilde{\alpha}_t = a_0 + d' z_t$ and $(\lambda_t, \gamma_i)' = b_0 + b' z_t$, where $(a_0, a')'$ and $(b_0, b')'$ are a $(q+1) \times 1$ vector and a $(q+1) \times (q+1)$ matrix, respectively. As noted before, this yields a linear factor model with observable factors, which permits closed-form GMM estimation. Due to the increase in the number of factors and therefore the number of free parameters, more information is needed to obtain reliable estimators. We therefore conduct our estimation based on the 100 portfolios sorted by size and BE/ME, available from Ken French’s website. This set of portfolios has also been used in Ludvigson and Ng (2005), where the information contained in those portfolio returns as well as other macroeconomic and financial series is summarized by a few estimated factors via dynamic factor analysis. They find that the estimated factors recover much information contained in those financial series, which helps identify a strongly positive conditional risk-return relation. Different from their factor-augmented specification, we directly exploit the cross-sectional information in the 100 size and BE/ME sorted portfolios with the GMM-based approach and are able to identify not only the positive risk-return relation on average but more importantly the significant time-variation in this relation.
Table IV presents the estimation results. Using the innovations in the predictors as pricing factors leads to similar results as those obtained from using the Fama-French factors in Table II. Both $\lambda_t$ and $\gamma_t$ are strongly dependent on the state variables and exhibit profound time variation. As in Table II, $\lambda_t$ is positively correlated with the dividend yield and term spread and negatively correlated with the default spread. It also exhibits significant negative dependence on the riskfree rate. Recall that $\gamma_t$ measures the relation between the risk premium and the hedge-related risk. It also shows substantial variation with most state variables, although the state variables used in this case differ from those used in Table II. Figure 4 plots the time series of $\lambda_t$ for comparison with the results in plot A of Figure 2. It shows a similar counter-cyclical pattern as plot A in Figure 2 although the variation appears more profound. Again, $\lambda_t$ is estimated to be mostly positive and significant, but takes negative values in the early 1980s, late 1990s and mid 2000s. The time series of $\gamma_t$ are not presented because the hedge-related risk in this case is captured by different variables from the case with the Fama-French model and thus there is no meaningful comparison of $\gamma_t$ between the results from these two models.

The fact that both sets of state variables (i.e., the innovations to the predictive variables and the Fama-French factors) lead to similar results for the risk-return relation suggests that SMB and HML are proxies for the state variables that drive the investment opportunity set. This is consistent with an ICAPM explanation for the empirical success of the Fama-French model. Petkova (2006) shows that SMB and HML are correlated with innovations to predictive variables and that incorporating the changes to the dividend yield, term spread and the riskfree rate reduces the explanatory power of SMB and HML. Our study does not look at the cross-sectional pricing performance of different factors but instead focuses on their implication for the intertemporal risk-return relation.

3.5 Are the results sensitive to the parametrization?

We have assumed that the coefficients relating risk to return are linear functions of the state variables. This linear specification is parsimonious and commonly used in empirical studies since it provides a first-order approximation of the true relationship. However, since the validity of our empirical results depends crucially on the validity of such assumptions, it is important to examine the robustness of our findings to these model specifications.

$^{15}$Given the concern of small-sample biases, we reconstruct the datasets of the same length as the original data by resampling contemporaneously from the residuals of the cross-sectional model (2) and the predictive regression (17) and (18), and then re-estimate the model 500 times. The resulting t-statistics from the bootstrap method are reported in the brackets in Table IV. The results remain the same qualitatively.
To this end, we employ a nonparametric method to estimate the time-varying relation mainly because of the empirical success of the conditional Fama-French model shown in the nonparametric framework of Wang (2003). When we relax the parametric assumption, our results remain broadly unaffected.

The time-variation in $\alpha_t$, $\lambda_t$ and $\gamma_t$ are determined by the cross-sectional asset returns following (6) and the associated normalizing condition $E_t[m_{t+1}] = 1$. That is,

$$\alpha_t E_t[r_{t+1}^i] - \lambda_t E_t[r_{t+1}^m r_{t+1}^i] - \gamma_t'E_t[s_{t+1}r_{t+1}^i] = 0, \text{ for all } i,$$

which can be written in matrix form as:

$$B_t \theta_t = e,$$

where $\theta_t = (\alpha_t, \lambda_t, \gamma_t')'$, $e = (0, ..., 0, 1)'$ is an $(N+1) \times 1$ vector, and $B_t$ is an $(N+1) \times (k+2)$ matrix whose first $N$ rows are given by $(E_t[r_{t+1}^i], -E_t[r_{t+1}^m r_{t+1}^i], -E_t[s_{t+1}r_{t+1}^i])$ for $i = 1, ..., N$, and whose last row is given by $(1, -E_t[r_{t+1}^m], -E_t[s_{t+1}])$.

The matrix $B_t$ can be estimated nonparametrically by replacing each element with its corresponding kernel regression estimate:

$$\widehat{B}_t \theta_t = e,$$

where $y_{t+1}$ represents $r_{t+1}^i$, $r_{t+1}^m r_{t+1}^i$, $s_{t+1}r_{t+1}^i$, $r_{t+1}^m$, or $s_{t+1}$. The resulting estimator $\widehat{B}_t$ is the Nadaraya-Watson kernel estimator and converges to $B_t$ asymptotically under certain regularity conditions on the kernel function $K(\cdot)$. The estimation takes into account the cross-sectional correlation in returns as its approach is quite general and makes no ad hoc assumption on the return dynamics.

Let $\widehat{\xi}_t(\theta_t) = \widehat{B}_t \theta_t - e$, then the parameters $\theta_t$ can be consistently estimated with:

$$\widehat{\theta}_t = \arg\min_{(\theta_t)} \widehat{\xi}_t(\theta_t)'\widehat{\xi}_t(\theta_t) = (\widehat{B}_t'^{-1}\widehat{B}_t)^{-1}\widehat{B}_t'e.$$
In choosing the kernel we follow the common practice to use an independent multivariate normal density function $K(u) = \prod_{i=1}^{q} \varphi_i(u_i)$, where $\varphi_i$ is the univariate normal density with mean zero and variance $k\sigma_i^2$, and $\sigma_i^2$ is the variance of the $i$th state variable and replaced by its sample variance in estimation. The kernel $K(\cdot)$ can be viewed as a rescaled kernel with the bandwidth $\sqrt{k}\sigma_i$ from the standard normal density. We determine the optimal value of the scaler $k$ using the leave-one-out cross-validation method (Härdle (1990)). The leave-one-out method minimizes the average squared error in out-of-sample prediction. The nonparametric method has clear advantages and disadvantages. The advantage is its robustness. The estimator is less biased than incorrectly specified parametric estimators, and is consistent. However, this comes at an efficiency cost. As a result, the variance of the estimator tends to exceed the variance of a correctly specified parametric estimator.

Table V shows the nonparametric estimates of the intertemporal relation based on the Fama-French factors using the 25 portfolios sorted by size and BE/ME. The estimates are reported for the corresponding conditioning variables taking its 10th, 25th, 50th, 75th, and 90th percentiles while other variables are fixed at their medians. The t-statistics given between brackets are produced with a bootstrap from 2000 simulated realizations. The bootstrap method samples asset returns, risk factors, and state variables, and obtains the t-statistics by computing the standard errors of the parameter estimates from the simulated realizations.

The estimates of $\lambda_t$ vary with all the conditioning variables, and remain significant and positive across different percentiles of investment states, except at the 10th percentiles of the term spread and 90th percentile of the T-Bill rate, where the estimate is positive but insignificant. The estimated intertemporal relation between the market premium and the market covariance with the state variables is also estimated positive most of the time. The hedging components related to $HML$ are significant in all the states whereas those related to $SMB$ are insignificant.

We can also see that $\lambda_t$ first decreases and then increases when the default spread rises. $\gamma_{1t}$ appears to be non-monotonic in the dividend yields, term spread, and default spread, and $\gamma_{2t}$ is non-monotonic in the dividend yield. To test the nonlinearity of $\lambda_t$ in the dividend yield, for example, we can estimate the ratio difference $\frac{\lambda^3 - \lambda^2}{div^3 - div^2} - \frac{\lambda^2 - \lambda^1}{div^2 - div^1}$, where $div^1$, $div^2$, and $div^3$ denote the 10th, 50th, and 90th percentiles for the dividend yield, respectively, and $\lambda^1$, $\lambda^2$, and $\lambda^3$ are the values of $\lambda_t$ corresponding to $div^1$, $div^2$, and $div^3$ while other variables are fixed at their medians. A nonzero value for this ratio difference would suggest nonlinearity of $\lambda_t$ in the dividend yield. Similarly we can test the nonlinearity of $\lambda_t$, $\gamma_{1t}$ and $\gamma_{2t}$ in other conditioning variables. The last row of Table V reports the bootstrap t-statistics
of the ratio difference. The results show some nonlinearity of $\lambda_t$ and $\gamma_{2t}$ in the state variables. The nonlinearity should not be surprising as there is no theoretical reason to expect $\lambda_t$, $\gamma_{1t}$, or $\gamma_{2t}$ to be linear in the conditioning variables.

Figure 5 plots the time series of the nonparametric estimates based on the Fama-French model. Although nonparametric estimation leads to lower efficiency than the parametric estimation, the results are largely consistent with the results in Figure 2, showing time-variation in $\lambda_t$, $\gamma_{1t}$, and $\gamma_{2t}$. The estimates for $\lambda_t$ vary counter-cyclically and are mostly positive and significant. The estimation errors are large during the early eighties because of the boundary effect of the kernel regression. Note in Figure 1 that all the conditioning variables, the dividend yield, the term spread, the default spread, and the T-Bill rate, take values near the boundary of their corresponding observation intervals. Note that in kernel regressions, fewer observations can be averaged at the boundary, leading to lower accuracy of the estimates.

### 3.6 The importance of the information from the cross-section

Our success in identifying the time-varying risk-return relation stands in sharp contrast to the failure of many earlier studies to find a positive risk-return relation under the more restrictive setting of a constant relation. It shows the importance of the cross-sectional information to our success of finding the time-varying risk-return trade-off, and more importantly that the cross-sectional aggregation of asset returns could cause significant information loss, making it difficult to identify the dynamics of the intertemporal risk-return relation. It also indicates that the standard 25 portfolios sorted on size and BE/ME captures much more of the cross-sectional information in asset returns, which is consistent with Fama-French’s justification for using the 25 portfolios in asset pricing tests. Fama and French (1993) show that average returns tend to decrease with size in the BE/ME quintile, and tend to increase with BE/ME in the size quintile.

### 4 Conclusions

We used the structure imposed by Merton’s (1973) ICAPM to obtain monthly estimates of the market-level risk-return relationship from the cross-section of equity returns. Our econometric approach sidesteps the specification of time-series models for the conditional risk premium and volatility of the market portfolio. We showed that the risk-return relation
is mostly positive but varies considerably over time. It covaries positively with counter-cyclical state variables. The relationship between the risk premium and hedge-related risk also exhibits strong time-variation, which supports the empirical evidence that aggregate risk aversion varies over time. Finally, the ICAPM’s two components of the risk premium show distinctly different cyclical properties. The volatility component exhibits a counter-cyclical pattern whereas the hedging component is less related to the business cycle and falls below zero for extended periods. This suggests the market serves an important hedging role for long-term investors.
References


Table I: Constant risk-return relation

This table shows the estimates of the intertemporal relations when the relations are assumed to be time-invariant. The intertemporal relationship is captured by the equation:

\[ E_t[r_{t+1}^m] = \lambda \text{Var}_t[r_{t+1}^m] + \gamma_1 \text{Cov}_t[r_{t+1}^m, s_{1,t+1}] + \gamma_2 \text{Cov}_t[r_{t+1}^m, s_{2,t+1}], \]

where \( \lambda \) measures the intertemporal relation between the market risk premium and conditional market volatility, \( \gamma_1 \) and \( \gamma_2 \) measure the intertemporal relation between the market risk premium and the conditional market covariance with the state variables \( s_{1,t+1} \) and \( s_{2,t+1} \), respectively, which capture the hedging components. The SMB and HML return spreads are taken as \( s_{1,t+1} \) and \( s_{2,t+1} \). The estimates are obtained with cross-sectional analysis using three different methods. GMM denotes the results from the GMM approach, and FMR\textsuperscript{full} (FMR\textsuperscript{rolling}) denotes those from the Fama-Macbeth method with full (rolling) sample used in the first-stage time series regression. The t-statistics are in the parentheses. The p-values from the GMM test are reported for testing the hypothesis that all the relations stay invariant over time.

<table>
<thead>
<tr>
<th></th>
<th>( \lambda )</th>
<th>( \gamma_1 )</th>
<th>( \gamma_2 )</th>
<th>( p )-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GMM</td>
<td>4.27</td>
<td>2.41</td>
<td>8.70</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(3.91)</td>
<td>(1.70)</td>
<td>(5.63)</td>
<td></td>
</tr>
<tr>
<td>FMR\textsuperscript{full}</td>
<td>3.93</td>
<td>2.23</td>
<td>8.74</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.94)</td>
<td>(1.58)</td>
<td>(5.54)</td>
<td></td>
</tr>
<tr>
<td>FMR\textsuperscript{rolling}</td>
<td>3.85</td>
<td>3.27</td>
<td>7.59</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.55)</td>
<td>(1.48)</td>
<td>(3.28)</td>
<td></td>
</tr>
</tbody>
</table>

29
Table II: Time-varying risk-return relation using Fama-French factors

This table shows the parameter estimates of the model when the relation is allowed to be time-varying. The intertemporal relationship is captured by the equation:

\[
E_t[r_{t+1}^m] = \lambda_t \text{Var}_t[r_{t+1}^m] + \gamma_{1t} \text{Cov}_t[r_{t+1}^m, s_{1,t+1}] + \gamma_{2t} \text{Cov}_t[r_{t+1}^m, s_{2,t+1}],
\]

where \( \lambda_t \) measures the intertemporal relation between the market risk premium and conditional market volatility, \( \gamma_{1t} \) and \( \gamma_{2t} \) measure the intertemporal relation between the market risk premium and the conditional market covariance with the state variables \( s_{1,t+1} \) and \( s_{2,t+1} \), respectively. The last two terms on the right-hand side of the equation capture the hedging components. The SMB and HML return spreads are taken as \( s_{1,t+1} \) and \( s_{2,t+1} \). \( \lambda_t, \gamma_{1t}, \text{and} \gamma_{2t} \) are linear in conditioning variables, given by the dividend yield, term spread, default spread, and 30-day Treasury Bill rate. The parameter estimates are obtained with cross-sectional data using the GMM approach with the Fama and French’s 25 BE/ME-size portfolios used alone (panel A) or together with their 30 industry portfolios (panel B). The parameter estimates are reported and the t-statistics are in the parentheses.
### A) Using FF 25 BMSZ

<table>
<thead>
<tr>
<th></th>
<th>$\lambda_t$</th>
<th>$\gamma_{1t}$</th>
<th>$\gamma_{2t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-13.83</td>
<td>5.85</td>
<td>-0.87</td>
</tr>
<tr>
<td></td>
<td>(-3.06)</td>
<td>(1.01)</td>
<td>(-0.22)</td>
</tr>
<tr>
<td>DIV</td>
<td>638.62</td>
<td>-434.55</td>
<td>233.93</td>
</tr>
<tr>
<td></td>
<td>(4.59)</td>
<td>(-2.55)</td>
<td>(1.24)</td>
</tr>
<tr>
<td>TERM</td>
<td>7.93</td>
<td>-2.35</td>
<td>6.16</td>
</tr>
<tr>
<td></td>
<td>(3.92)</td>
<td>(-0.94)</td>
<td>(2.64)</td>
</tr>
<tr>
<td>DEF</td>
<td>-16.09</td>
<td>20.22</td>
<td>-3.43</td>
</tr>
<tr>
<td></td>
<td>(-4.24)</td>
<td>(3.27)</td>
<td>(-0.55)</td>
</tr>
<tr>
<td>RF</td>
<td>1994.46</td>
<td>-2497.63</td>
<td>1088.16</td>
</tr>
<tr>
<td></td>
<td>(2.43)</td>
<td>(-2.17)</td>
<td>(0.80)</td>
</tr>
</tbody>
</table>

### B) Using FF 25 BMSZ + 30 industry

<table>
<thead>
<tr>
<th></th>
<th>$\lambda_t$</th>
<th>$\gamma_{1t}$</th>
<th>$\gamma_{2t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-24.17</td>
<td>8.66</td>
<td>-23.90</td>
</tr>
<tr>
<td></td>
<td>(-6.06)</td>
<td>(1.81)</td>
<td>(-4.93)</td>
</tr>
<tr>
<td>DIV</td>
<td>801.55</td>
<td>-223.54</td>
<td>371.20</td>
</tr>
<tr>
<td></td>
<td>(7.34)</td>
<td>(-1.57)</td>
<td>(2.66)</td>
</tr>
<tr>
<td>TERM</td>
<td>10.58</td>
<td>-4.05</td>
<td>13.25</td>
</tr>
<tr>
<td></td>
<td>(6.45)</td>
<td>(-2.33)</td>
<td>(7.42)</td>
</tr>
<tr>
<td>DEF</td>
<td>-17.91</td>
<td>20.11</td>
<td>-16.39</td>
</tr>
<tr>
<td></td>
<td>(-5.60)</td>
<td>(6.16)</td>
<td>(-4.12)</td>
</tr>
<tr>
<td>RF</td>
<td>1365.74</td>
<td>-3770.64</td>
<td>5591.03</td>
</tr>
<tr>
<td></td>
<td>(4.09)</td>
<td>(-4.14)</td>
<td>(6.15)</td>
</tr>
</tbody>
</table>
Table III: Predictive regressions

This table presents the estimation results for the predictive regression:

\[ r_{t+1}^m = \nu + \delta z_t + \varepsilon_{t+1}^m, \]
\[ z_{t+1} = C_0 + C z_t + \varepsilon_{t+1}^z, \]

where \( r_{t+1}^m \) is the market excess return and \( z_t \) is a vector of predictive variables including the dividend yields (DIV), term spread (TRM), default spread (DEF), and one-month T-Bill rate (RF). Panel (A) shows the parameter estimates and reports their standard errors in parentheses. It also reports the \( R^2 \) values for each equation in the predictive regression. Panel (B) presents the correlation matrix of the shocks with the shock standard deviations on the diagonal.

<table>
<thead>
<tr>
<th>Panel A: Parameter estimates</th>
<th>Constant</th>
<th>DIV</th>
<th>TRM</th>
<th>DEF</th>
<th>RF</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>MKT</td>
<td>-0.0097</td>
<td>0.5427</td>
<td>0.0027</td>
<td>0.0018</td>
<td>-2.5478</td>
<td>0.03</td>
</tr>
<tr>
<td>(0.0061)</td>
<td>(0.1709)</td>
<td>(0.0021)</td>
<td>(0.0054)</td>
<td></td>
<td>(1.0652)</td>
<td></td>
</tr>
<tr>
<td>DIV</td>
<td>0.0006</td>
<td>0.9835</td>
<td>-0.0001</td>
<td>-0.0003</td>
<td>0.0562</td>
<td>0.98</td>
</tr>
<tr>
<td>(0.0002)</td>
<td>(0.0060)</td>
<td>(0.0001)</td>
<td>(0.0002)</td>
<td></td>
<td>(0.0376)</td>
<td></td>
</tr>
<tr>
<td>TRM</td>
<td>-0.0545</td>
<td>-0.3835</td>
<td>0.9632</td>
<td>0.0896</td>
<td>3.1287</td>
<td>0.93</td>
</tr>
<tr>
<td>(0.0382)</td>
<td>(1.0681)</td>
<td>(0.0134)</td>
<td>(0.0339)</td>
<td></td>
<td>(6.6569)</td>
<td></td>
</tr>
<tr>
<td>DEF</td>
<td>0.0100</td>
<td>0.3397</td>
<td>-0.0136</td>
<td>0.9851</td>
<td>-0.1274</td>
<td>0.94</td>
</tr>
<tr>
<td>(0.0149)</td>
<td>(0.4170)</td>
<td>(0.0052)</td>
<td>(0.0132)</td>
<td></td>
<td>(2.5990)</td>
<td></td>
</tr>
<tr>
<td>RF</td>
<td>0.0003</td>
<td>0.0010</td>
<td>-0.0001</td>
<td>0.0001</td>
<td>0.9168</td>
<td>0.91</td>
</tr>
<tr>
<td>(0.0001)</td>
<td>(0.0029)</td>
<td>(0.0000)</td>
<td>(0.0001)</td>
<td></td>
<td>(0.0178)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Correlations/standard deviations</th>
<th>MKT</th>
<th>DIV</th>
<th>TRM</th>
<th>DEF</th>
<th>RF</th>
</tr>
</thead>
<tbody>
<tr>
<td>MKT</td>
<td>0.0427</td>
<td>-0.9095</td>
<td>0.1086</td>
<td>-0.0598</td>
<td>-0.1428</td>
</tr>
<tr>
<td>DIV</td>
<td>-0.9095</td>
<td>0.0015</td>
<td>-0.1657</td>
<td>0.0238</td>
<td>0.2103</td>
</tr>
<tr>
<td>TRM</td>
<td>0.1086</td>
<td>-0.1657</td>
<td>0.2669</td>
<td>0.1663</td>
<td>-0.3255</td>
</tr>
<tr>
<td>DEF</td>
<td>-0.0598</td>
<td>0.0238</td>
<td>0.1663</td>
<td>0.1042</td>
<td>0.0179</td>
</tr>
<tr>
<td>RF</td>
<td>-0.1428</td>
<td>0.2103</td>
<td>-0.3255</td>
<td>0.0179</td>
<td>0.0007</td>
</tr>
</tbody>
</table>
Table IV: Time-varying risk-return relation using predictor innovations as factors

This table shows the parameter estimates of the model:

\[
E_t[r_{t+1}^i] = \lambda_t \text{Cov}_t[r_{t+1}^i, \varepsilon_{t+1}^m] + \gamma'_t \text{Cov}_t[r_{t+1}^i, \varepsilon_{t+1}^z],
\]

where \( \varepsilon_{t+1}^m \) and \( \varepsilon_{t+1}^z \) measure the innovations in market return and predictive variables from the predictive regression, respectively. \( \lambda_t, \gamma_{1t} \) and \( \gamma_{2t} \) are linear in conditioning variables, given by the dividend yield, term spread, default spread, and 30-day Treasury Bill rate. The estimation is conducted with cross-sectional data using the GMM approach with GMM t-statistics given in the parenthesis. The numbers in the brackets give the t-statistics from bootstrap method. Business cycle peaks are marked by dashed lines, business cycle troughs are marked by solid lines.

<table>
<thead>
<tr>
<th></th>
<th>( \lambda_t )</th>
<th>( \gamma_{1t} )</th>
<th>( \gamma_{2t} )</th>
<th>( \gamma_{3t} )</th>
<th>( \gamma_{4t} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-11.99 (3.06)</td>
<td>-94.74 (-0.41)</td>
<td>-1.36 (-2.41)</td>
<td>-18.64 (-11.54)</td>
<td>1513.23 (2.89)</td>
</tr>
<tr>
<td>DIV</td>
<td>1811.50 (11.91)</td>
<td>3965.28 (16.26)</td>
<td>0.34 (0.82)</td>
<td>279.80 (7.48)</td>
<td>-1467.73 (-0.46)</td>
</tr>
<tr>
<td>TERM</td>
<td>8.52 (4.92)</td>
<td>33.07 (2.05)</td>
<td>-0.16 (-2.45)</td>
<td>2.39 (7.05)</td>
<td>-318.60 (-5.50)</td>
</tr>
<tr>
<td>DEF</td>
<td>-30.01 (-11.67)</td>
<td>-351.39 (-6.41)</td>
<td>-2.03 (-6.16)</td>
<td>-1.18 (-5.28)</td>
<td>-59.18 (-0.94)</td>
</tr>
<tr>
<td>RF</td>
<td>-4941.05 (-6.40)</td>
<td>63337.07 (8.45)</td>
<td>336.04 (4.01)</td>
<td>1037.75 (5.05)</td>
<td>-205894.91 (-7.25)</td>
</tr>
</tbody>
</table>
Table V: Nonparametric estimates of time-varying risk-return relation

This table shows the nonparametric estimates of the intertemporal relation as a function of the conditioning variables. The intertemporal relationship is captured by the equation

$$E_t[r^m_{t+1}] = \lambda_t \text{Var}_t[r^m_{t+1}] + \gamma_{1t} \text{Cov}_t[r^m_{t+1}, s_{1,t+1}] + \gamma_{2t} \text{Cov}_t[r^m_{t+1}, s_{2,t+1}],$$

where $\lambda_t$ measures the intertemporal relation between the market risk premium and conditional market volatility, $\gamma_{1t}$ and $\gamma_{2t}$ measure the intertemporal relation between the market risk premium and the conditional market covariance with the state variables $s_{1,t+1}$ and $s_{2,t+1}$, respectively. The SMB and HML return spreads are taken as $s_{1,t+1}$ and $s_{2,t+1}$. The estimates, obtained with kernel regressions, are reported corresponding to the different percentiles of the conditioning variable (dividend yield, term spread, default spread, or 30-day Treasury Bill rate), while other variables are fixed at their medians. The bootstrap t-statistics are reported in the parentheses. The last row of each table gives the bootstrap t-statistics for testing the nonlinearity of $\lambda_t$, $\gamma_{1t}$ and $\gamma_{2t}$ in the conditioning variable.

<table>
<thead>
<tr>
<th>Percentile</th>
<th>Dividend yield</th>
<th>Term spread</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\lambda_t$</td>
<td>$\gamma_{1t}$</td>
</tr>
<tr>
<td>10%</td>
<td>4.10 (2.18)</td>
<td>1.60 (0.57)</td>
</tr>
<tr>
<td>25%</td>
<td>4.12 (2.75)</td>
<td>0.69 (0.30)</td>
</tr>
<tr>
<td>50%</td>
<td>4.17 (2.87)</td>
<td>0.38 (0.17)</td>
</tr>
<tr>
<td>75%</td>
<td>4.29 (2.74)</td>
<td>0.75 (0.32)</td>
</tr>
<tr>
<td>90%</td>
<td>4.42 (2.44)</td>
<td>1.68 (0.68)</td>
</tr>
<tr>
<td>ratio difference</td>
<td>1.95</td>
<td>0.97</td>
</tr>
<tr>
<td>Percentile</td>
<td>Default spread</td>
<td>T-Bill rate</td>
</tr>
<tr>
<td>-----------</td>
<td>----------------</td>
<td>-------------</td>
</tr>
<tr>
<td></td>
<td>$\lambda_t$</td>
<td>$\gamma_{1t}$</td>
</tr>
<tr>
<td>10%</td>
<td>4.59</td>
<td>1.17</td>
</tr>
<tr>
<td></td>
<td>(3.00)</td>
<td>(0.49)</td>
</tr>
<tr>
<td>25%</td>
<td>4.37</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>(2.99)</td>
<td>(0.34)</td>
</tr>
<tr>
<td>50%</td>
<td>4.17</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>(2.88)</td>
<td>(0.17)</td>
</tr>
<tr>
<td>75%</td>
<td>4.14</td>
<td>-0.05</td>
</tr>
<tr>
<td></td>
<td>(2.48)</td>
<td>(-0.02)</td>
</tr>
<tr>
<td>90%</td>
<td>4.46</td>
<td>1.03</td>
</tr>
<tr>
<td></td>
<td>(1.94)</td>
<td>(0.34)</td>
</tr>
<tr>
<td>ratio difference</td>
<td>2.91</td>
<td>1.32</td>
</tr>
</tbody>
</table>
Figure 1: State variables
This figure plots the time series of the dividend yield, term spread, default spread, and T-bill rate for the sample period from April 1953 to December 2008. Business cycle peaks are marked by dashed lines, business cycle troughs are marked by solid lines.
Figure 2: Time-varying risk-return relation using Fama-French factors
This figure plots the time series of the intertemporal relation using parametric estimation (solid lines) along with its 95 percent confidence intervals (dotted lines). The intertemporal relationship is captured by equation (3). The SMB and HML return spreads are taken as $s_{1,t+1}$ and $s_{2,t+1}$. $\lambda_t$, $\gamma_{1t}$, and $\gamma_{2t}$ are assumed to be linear in conditioning variables. The estimation is conducted with cross-sectional data using the GMM approach. Panel A reports estimation results based on the Fama and French’s 25 BE/ME-size portfolios and Panel B reports results based on the expanded portfolio set by adding the 30 industry portfolios. Business cycle peaks are marked by dashed lines, business cycle troughs are marked by solid lines.
Panel B

Plot A: $\lambda_t$

Plot B: $\gamma_{1t}$

Plot C: $\gamma_{2t}$
Figure 3: Volatility and hedging components of the risk premium
This figure plots the time series of the estimated volatility and hedging components of market risk premium for the sample period. Business cycle peaks are marked by dashed lines, business cycle troughs are marked by solid lines.

Plot A: Volatility component

Plot B: Hedging component
Figure 4: Time-varying risk-return relation using predictor innovations as factors
This figure plots the time series of the intertemporal relation using parametric estimation (solid lines) based on the standard 25 portfolios formed on size and book-to-market equity, along with its 95 percent confidence intervals (dotted lines). The intertemporal relationship is captured by the equation

$$E_t[r_{t+1}^i] = \lambda_t \text{Cov}_t[r_{t+1}^i, \varepsilon_{t+1}^m] + \gamma_t' \text{Cov}_t[r_{t+1}^i, \varepsilon_{t+1}^z].$$

where $\varepsilon_{t+1}^m$ and $\varepsilon_{t+1}^z$ measure the innovations in market return and predictive variables from the predictive regression, respectively. $\lambda_t$, $\gamma_{1t}$ and $\gamma_{2t}$ are linear in conditioning variables, given by the dividend yield, term spread, default spread, and 30-day Treasury Bill rate. The estimation is conducted with cross-sectional data using the GMM approach. Business cycle peaks are marked by dashed lines, business cycle troughs are marked by solid lines. The sample period is from April 1953 to December 2008.
Figure 5: Nonparametric estimates of the time-varying risk-return

This figure plots the time series of the intertemporal relation using nonparametric estimation (solid lines) along with its 95 percent confidence intervals (dotted lines). The intertemporal relationship is captured by equation (3). The SMB and HML return spreads are taken as $s_{1,t+1}$ and $s_{2,t+1}$. $\lambda_t$, $\gamma_{1t}$ and $\gamma_{2t}$ are assumed to be linear in conditioning variables. The estimation is conducted with cross-sectional data using the GMM approach. Business cycle peaks are marked by dashed lines, business cycle troughs are marked by solid lines.