Why Do Discount Rates Vary?*

Serhiy Kozak[†], **Shrihari Santosh**[‡]

December 8, 2016

Abstract

We argue that the price of "discount-rate" risk reveals whether increases in equity risk premia represent "good" or "bad" news to rational investors. We employ a new empirical methodology and find that the price is negative, contrary to previous estimates. This finding supports equilibrium models with stochastic technology or preferences as the drivers of time-varying expected returns, but is inconsistent with canonical models of sentiment. Our approach relies on using future realized market returns to consistently estimate covariances of asset returns with the market risk premium. Covariances drive observed patterns in the broad cross-sections of stock and bond expected returns.

*We thank Nick Barberis, Mike Chernov, John Cochrane, Julien Cujean, Eugene Fama, Stefano Giglio, Valentin Haddad, Lars Hansen, John Heaton, Bryan Kelly, Ralph Koijen, Stefan Nagel, Diogo Palhares, Jay Shanken, Ken Singleton, two anonymous referees, participants at the European Finance Association Meetings, Finance Down Under Conference, SFS Finance Cavalcade, and seminar participants at Chicago, Emory, Maryland, Michigan, UCLA, UIUC, Wisconsin for helpful comments and suggestions. This paper was previously circulated under the title "Linking Cross-Sectional and Aggregate Expected Returns".

[†]Ross School of Business, University of Michigan. Email: sekozak@umich.edu.

[‡]R.H. Smith School of Business, University of Maryland. Email: shrihari@umd.edu.

1 Introduction

Discount rates (expected excess returns) unquestionably vary significantly over time. Yet there is little agreement as to why[1]. Is technology highly volatile? Are preferences particularly sensitive to economic conditions? Or are markets prone to fads and panics? The question seems almost too vague given the nearly limitless number of potential explanations for return predictability. The typical approach to addressing this question is to postulate a particular model of beliefs and preferences and test how well that specific model fits with empirical regularities. In this paper we approach the problem from a new perspective: we categorize models based on whether rational investors (within each model) "like" or "dislike" states of the world with high expected returns. This dichotomy translates into the premium or discount accruing to an asset based on its return covariance with discount-rate shocks. We develop a new "model-free" empirical approach which overcomes the difficulty in precisely estimating shocks to aggregate expected returns and does not rely on assumptions about investors' information set. We find that positive covariance with such shocks leads to lower expected returns, implying states of the world with high discount rates are "bad". This finding supports models with stochastic technology or preferences as the drivers of time-varying expected returns and is inconsistent with many canonical models of sentiment (biased expectations).

We consider two broad classes of popular asset pricing models. The first class, which we loosely label as "rational", consists of models in which agents have objective beliefs about the distribution of shocks in the economy. Time-varying risk premia (expected excess returns) are due to stochastic technology or preferences. This class includes standard consumption based models such as long-run risks with stochastic volatility (e.g., Bansal and Yaron, 2004) and habits (e.g., Campbell and Cochrane, 1999, Constantinides, 1990), as well as models featuring stochastic risk aversion (Dew Becker, 2011, Kozak, 2015) or ambiguity-averse agents (e.g., Drechsler, 2013), etc. The second group, which we label as "behavioral", consists of models in which some of investors have subjective (biased) beliefs. From the perspective of rational[2]

[1] "[D]iscount-rate variation is the central organizing question of current asset-pricing research." (Cochrane, 2011)

[2] We consider the perspective of an unconstrained investor with objective beliefs. For this agent, the Euler equation holds with equality under the objective measure, which we can learn about from historic data.

investors, expected returns vary "exogenously". Such models include Campbell and Kyle (1993), Kim and Omberg (1996), Campbell and Viceira (1999), Barberis et al. (2015) etc[3].

The key to our estimation is that both types of models typically posit the existence of an unconstrained[4] investor with objective beliefs. The models differ in how such investors view discount rate shocks. In the "rational" class, expected returns are high during "bad times": states of the world with high marginal utility of consumption, volatility, ambiguity, risk-aversion, or low surplus consumption ratio. Investors are willing to pay to hedge against these states. This implies a negative price of risk for discount rate shocks. Rational investors prefer to hold assets which pay off when risk premia increase. Consequently, they bid up the prices of such assets so that in equilibrium, they have low expected returns.

In the "behavioral" class, rational investors consider states with high expected returns as "good times". They view discount rate variation as essentially exogenous. An increase in risk premia (and associated drop in asset values) lowers utility due to the direct income effect but is more than offset by the increase in utility due to better investment opportunities (substitution effect)[5]. Rational investors in these models dislike states of the world with *low* discount rates and are willing to pay to hedge against these states. This implies a *positive* price of risk for discount rate shocks, contrary to the predictions of the rational class of models[6].

Our "model-free" estimation uses a broad cross-section of return anomalies and reveals a striking pattern. We find that most of the variation in expected returns is explained by covariance with our risk-premium factor, which commands an unambiguously *negative* price of risk. This evidence suggests that *an unconstrained investor with objective beliefs dislikes states of the world with high aggregate risk premia.* This result contrasts with much of the

[3] Kim and Omberg (1996) and Campbell and Viceira (1999) solve for a rational agent's optimal investment policy (and implicitly her SDF) assuming that expected excess returns follow an AR(1) or Ornstein-Uhlenbeck process. Adding a group of sentiment investors whose belief bias follows an AR(1) leads to results obtained in these papers. Campbell and Kyle (1993) and Barberis et al. (2015) explicitly model the trading/beliefs of sentiment investors and solve analytically for the rational arbitrageurs' value function. It is trivial to derive the SDF given this closed form solution.

[4] We use unconstrained to mean the investor faces no binding hard constraints.

[5] Kim and Omberg (1996) analytically show this result for HARA utility with $\gamma < 1$. This includes exponential utility and power utility with RRA > 1. Campbell and Viceira (1999) numerically show that for Epstein-Zin preferences with power aggregator, it obtains if RRA > 1 and also IES $\neq 1$.

[6] We recognize it easy to write both "rational" and "behavioral" models with arbitrary sign of price of risk (depending on the covariance of various shocks). Our analysis considers canonical models and captures the economic intuition of these frameworks.

prior literature, in which the evidence is mixed but, perhaps, tilts towards a positive price of risk[7]. Our findings thus imply that the time-variation in risk premia is an important economic risk which affects asset prices and provide strong support for a rational basis (technology or preference shocks) for time-varying expected market returns.

Our empirical investigation starts with a general two-factor SDF which contains shocks to the market return and the market risk premium. Such an ICAPM representation (Merton, 1973) obtains for both "rational" and "behavioral" classes of models which we discussed previously. We show how to condition down the pricing model implied by such a pricing kernel in a general way. We do so under two different settings.

In the first approach we assume that conditional covariances are constant (but study the robustness of this assumption via simulation later in the paper). We show that when discount rates follow an AR(1) process, it is possible to replace conditional covariance with respect to shocks to discount rates with unconditional covariances with respect to *levels* of discount rates. This substitution results in a "twist" of the unconditional prices of risk relative to their conditional expectations, but we show that the sign of the price of discount-rate risk is preserved under very mild and economically realistic conditions.

We derive similar results in a setup with constant prices of risk, but with stochastic volatility. Following Campbell et al. (2015), we assume a single state variable drives all variances. We obtain similar unconditional results in this framework. Again, the sign of the price of risk is preserved for reasonable calibrations[8]. Therefore, our results apply equally to models with time-varying prices or quantities of risk.

Finally, we argue that the covariance of asset returns with the level of discount rates can be consistently estimated by computing the covariance with future realized market returns. This methodological insight allows us to estimate the pricing equation consistently, without relying on any particular model or assumptions about the information set that investors use (predictive variables).

This is not the first paper to estimate an ICAPM representation of expected asset returns which includes shocks to discount rates. Where we differ from the prior literature is

[7]As a technical aside, we estimate the price of shocks to expected excess returns, but much of the literature uses shocks to expected total returns. We prefer excess returns since they are inherently "real" and thus avoid issues of changes in expected inflation.

[8]Bansal et al. (2014), Campbell et al. (2015) find the opposite price of risk because they rely on a predictive VAR in their estimation and do not impose certain restrictions implied by their models.

our approach to estimating covariances. Most previous papers such as Bansal et al. (2014), Brennan et al. (2004), Campbell and Vuolteenaho (2004), Campbell et al. (2015), Koijen et al. (2015) use predictive regressions, in the form of a vector auto-regression (VAR), to estimate shocks to discount rates. We believe this approach is often unreliable. Different sample periods (Chen and Zhao, 2009a) and alternative definitions of the state vector (Bansal et al., 2014, Campbell and Vuolteenaho, 2004, Campbell et al., 2015, Chen and Zhao, 2009a) deliver conflicting evidence on the price of discount-rate risk. Indeed, the sign of the estimated price of discount rate shocks is inconsistent across, and even within, these studies. To circumvent this issue, we use future realized returns to proxy for expected future returns (see Section 2.3), yielding consistent estimates of covariances without the need to estimate expected returns.

Further evidence of the relation between marginal utility and the market risk premium comes from bonds. Long-term Treasury bonds have higher covariance with innovations to market discount rates than do short-term bonds. This is consistent with a commonly held view that long-maturity bonds are good hedges during times of market stress (when risk premia are high). We decompose the average return differential between long and short term government bonds into a large positive differential due to loadings on "level risk" (interest-rate risk) and a large negative spread due to loadings on our "risk premium" factor. These net to a slightly upward sloping term structure of expected bond returns. Our results suggest that analyzing fixed income securities in isolation can lead to erroneous conclusions about bond risks and risk premia.

2 Empirical Framework

2.1 Conditional Pricing Model

We start by assuming that risk premia are earned due to exposure to only two sources of risk: market risk (as in the CAPM) and "discount-rate" risk (risk due to time-variation in expected excess market returns). We therefore postulate the following reduced-form two-factor representation for expected returns:

$$\mathbb{E}_t\left[r_{t+1}^i\right]+\frac{V_t^i}{2}=\delta_t^m C_t^{i,m}+\delta_t^\lambda C_t^{i,\lambda} \tag{1}$$

where r_{t+1}^i denotes log returns on an asset i at time $t+1$ in excess of risk-free rate, δ_t^m and δ_t^λ are prices of market and discount-rate risk, respectively, $V_t^i=\mathrm{var}_t\left(r_{t+1}^i\right)$, and $C_t^{i,m}$ and $C_t^{i,\lambda}$ are conditional covariances of asset returns with shocks to market returns r_t^m and shocks to discount rates $\lambda_t=\mathbb{E}_t\left(R_{t+1}^m\right)\equiv\mathbb{E}_t\left(r_{t+1}^m\right)+\frac{V_t^m}{2}$, respectively[9]:

$$C_t^{i,m}=\mathrm{cov}_t\left[r_{t+1}^i,\, r_{t+1}^m\right] \tag{2}$$

$$C_t^{i,\lambda}=\mathrm{cov}_t\left[r_{t+1}^i,\, u_{t+1}^\lambda\right] \tag{3}$$

where $u_{t+1}^\lambda=\left(\mathbb{E}_{t+1}-\mathbb{E}_t\right)\lambda_{t+1}$. Such a representation obtains in standard consumption based models such as long-run risks with stochastic volatility (e.g., Bansal and Yaron, 2004) and habits (e.g., Campbell and Cochrane, 1999, Constantinides, 1990), models featuring stochastic risk aversion (Dew Becker, 2011, Kozak, 2015) or ambiguity-averse agents (e.g., Drechsler, 2013), as well as extrapolative expectations models, such as Campbell and Kyle (1993), Kim and Omberg (1996), Campbell and Viceira (1999), Barberis et al. (2015) etc.

The typical approach in the literature is to condition down the Eq. 1 and test the resulting unconditional pricing equation in the cross-section of equity returns. We proceed in this manner in the following section.

[9] We assume the approximation $\mathbb{E}_t\left(r_{t+1}^i\right)+\frac{V_t^i}{2}\approx\exp\left[\mathbb{E}_t\left(r_{t+1}^i\right)+\frac{V_t^i}{2}\right]-1$ holds.

2.2 Unconditional Pricing Model

Below we derive properties of the unconditional representation of Eq. 1 under two settings: (i) homoskedasticity with time-varying risk prices (δs); and (ii) heteroskedasticity with constant risk prices. In general, both approaches yield an unconditional pricing equation:

$$\mathbb{E}\left[r_t^i\right] + \frac{V^i}{2} = \tilde{\delta}^m \mathrm{cov}\left(r_t^i, \left(\mathbb{E}_t - \mathbb{E}_{t-1}\right) r_t^m\right) + \tilde{\delta}^\lambda \mathrm{cov}\left(r_t^i, \left(\mathbb{E}_t - \mathbb{E}_{t-1}\right) \lambda_t\right) \tag{4}$$

where $V^i = \mathbb{E}\left(V_t^i\right)$. The expression requires estimating covariances with *innovations* to discount rates. Discount rates innovations, however, are notoriously difficult to estimate because of our limited ability to forecast levels of discount rates. We show below that under certain assumptions we can substitute the second covariance in Eq. 4 with covariance with respect to *levels* of discount rates, cov $\left(r_t^i, \lambda_t\right)$. Furthermore, if we consider the cross-section of assets with zero exposure to the market (market-neutral), the unconditional pricing equation becomes

$$\mathbb{E}\left[rx_t^i\right] + \frac{V^i}{2} = \hat{\delta}^\lambda \mathrm{cov}\left(rx_t^i, \lambda_t\right) \tag{5}$$

which is much easier to work with in practice. We use rx_t^i notation to refer to *market-neutral* excess returns.

2.2.1 Time-varying Prices of Risk

In this section we derive an unconditional representation and its properties in a homoskedastic setting.

Assumption 1. *All second moments are constant.*

Technically, we can allow for "idiosyncratic" returns (uncorrelated with the two pricing factors) to be arbitrarily heteroskedastic. Homoskedasticity means Eq. 1 takes the simpler form:

$$\mathbb{E}_t\left[r_{t+1}^i\right] + \frac{V^i}{2} = \delta_t^m C^{i,m} + \delta_t^\lambda C^{i,\lambda} \tag{6}$$

In addition, we rely on the following simplifying assumption in our analysis:

Assumption 2. *Expected excess returns on the market portfolio follow an autonomous* $AR(1)$ *process:*

$$\lambda_t = \bar{\lambda} + \phi\left(\lambda_{t-1} - \bar{\lambda}\right) + u_t^{\lambda} \tag{7}$$

where $u_t^{\lambda} \sim \mathbb{N}(0, \sigma_{\lambda}^2)$.

Consider the cross-section of market-neutral excess returns rx_t^i , i.e. assets for which

$$\mathbb{E}_t\left[rx_{t+1}^i\right] + \frac{V^i}{2} = \delta_t^{\lambda} C^{i,\lambda} \tag{8}$$

holds. We link the unconditional covariance of asset returns with the expected market return and the conditional covariance $C^{i,\lambda}$ using the formula of total covariance:

$$\begin{aligned} \text{cov}\left(rx_t^i, \lambda_t\right) &= C^{i,\lambda} + \text{cov}\left(\delta_t^{\lambda} C^{i,\lambda}, \phi\lambda_t\right) \\ &= \left[1 + \phi \text{cov}\left(\delta_t^{\lambda}, \lambda_t\right)\right] \times C^{i,\lambda} \end{aligned} \tag{9}$$

Finally, by substituting this expression into Eq. 8 and taking unconditional expectations we obtain the unconditional relation:

$$\mathbb{E}\left[rx_t^i\right] + \frac{V^i}{2} \quad \equiv \quad \hat{\delta}^{\lambda} \times \text{cov}\left(rx_t^i, \lambda_t\right) \tag{10}$$

where

$$\hat{\delta}^{\lambda} = \bar{\delta}^{\lambda}\left[1 + \phi \text{cov}\left(\delta_t^{\lambda}, \lambda_t\right)\right]^{-1} \tag{11}$$

and $\bar{\delta}^{\lambda} = \mathbb{E}\left[\delta_t^{\lambda}\right]$.

The next two theorems provide conditions under which we can learn about sign $\left[\bar{\delta}^{\lambda}\right]$ from sign $\left[\hat{\delta}^{\lambda}\right]$.

Theorem 1. *If the sign of expected price of discount-rate risk is equal to the sign of the covariance between the price of discount-rate risk and the level of market discount rates* $\left(sign\left[\bar{\delta}^{\lambda}\right] = sign\left[\text{cov}\left(\delta_t^{\lambda}, \lambda_t\right)\right]\right)$*, a negative unconditional price of risk in equation Eq. 10 implies a negative expected conditional price, i.e.,* $\hat{\delta}^{\lambda} < 0 \implies \bar{\delta}^{\lambda} \equiv \mathbb{E}\left[\delta_t^{\lambda}\right] < 0$*. Additionally,* $\bar{\delta}^{\lambda} > 0 \implies \hat{\delta}^{\lambda} > 0$*, i.e., a positive expected conditional price of risk implies a positive unconditional price in equation Eq. 10.*

The condition that $\text{sign}\left[\bar{\delta}^{\lambda}\right] = \text{sign}\left[\text{cov}\left(\delta_t^{\lambda}, \lambda_t\right)\right]$ is economically motivated. It implies that compensation for discount-rate risk tends to increase in times of high aggregate discount rates. Proof of Theorem 1 in this case is straightforward and obtains directly from Eq. 10 and Eq. 11.

Below we provide an alternative argument under a different condition.

Theorem 2. *Absence of near-arbitrage opportunities (when Sharpe Ratios are "reasonably" bounded) implies that the signs of unconditional and expected conditional prices of discount-rate risk are the same,* $\text{sign}\left[\hat{\delta}^{\lambda}\right] = \text{sign}\left[\bar{\delta}^{\lambda}\right]$ *(the sign is fully preserved).*

The sign is fully preserved if and only if $1 + \phi\text{cov}\left(\delta_t^{\lambda}, \lambda_t\right) > 0$. Equation (8) can be rewritten as

$$\text{SR}_t^i = \delta_t^{\lambda} \rho_{i,\lambda} \sigma_{\lambda}$$

where $\text{SR}_t^i = \frac{\mathbb{E}\left[rx_t^i\right] + \frac{V^i}{2}}{\sigma_i}$ is the conditional Sharpe ratio on asset i, $\rho_{i.\lambda} = \text{corr}_t\left(rx_{t+1}^i, \lambda_{t+1}\right)$, and $\sigma_{\lambda}^2 = \text{var}_t\left(\lambda_{t+1}\right)$. The maximum (absolute) Sharpe ratio obtains for the the λ-mimicking portfolio ($\rho_{i,\lambda} = 1$) and is equal to:

$$\text{SR}_t^{\max} = \delta_t^{\lambda} \sigma_{\lambda}$$

The standard deviation of this Sharpe ratio is

$$\sigma\left(\text{SR}_t^{\max}\right) = \sigma\left(\delta_t^{\lambda}\right) \sigma_{\lambda}$$

Use this to rewrite the sign condition,

$$\begin{aligned}
0 <& 1 + \phi\text{cov}\left(\delta_t^{\lambda}, \lambda_t\right) \\
=& 1 + \phi\text{corr}\left(\delta_t^{\lambda}, \lambda_t\right) \sigma\left(\delta_t^{\lambda}\right) \sigma\left(\lambda_t\right) \\
=& 1 + \frac{\phi}{\sqrt{1-\phi^2}}\text{corr}\left(\delta_t^{\lambda}, \lambda_t\right) \sigma\left(\delta_t^{\lambda}\right) \sigma_{\lambda} \\
=& 1 + \frac{\phi}{\sqrt{1-\phi^2}}\text{corr}\left(\delta_t^{\lambda}, \lambda_t\right) \sigma\left(\text{SR}_t^{\max}\right)
\end{aligned}$$

We used the relationship between the conditional and unconditional variance of λ_t ($\sigma_{\lambda}^2 = \left(1-\phi^2\right) \sigma^2\left(\lambda_t\right)$) in the derivation above. Since $\text{corr}\left(\delta_t^{\lambda}, \lambda_t\right) \geq -1$ (-1 is the "worst case" setting), a sufficient condition is

$$\sigma\left(\mathrm{SR}_t^{\max}\right) < \sqrt{\frac{1}{\phi^2} - 1}$$

Consider an annual calibration. To be conservative, let $\phi = 0.7$ — a value which allows for slightly more high frequency variation in discount rates than that implied by only the variation in D/P ratio, but is similar in magnitude to some of the "de-trended" and "de-biased" estimates (Sabbatucci, 2015). For simplicity assume the mean max Sharpe ratio is equal to $2 \times \sigma\left(\mathrm{SR}_t^{\max}\right)$ so that the expected mimicking portfolio return is unlikely to change sign. Such values imply that an annualized max SR on market-neutral strategies can reach values $4 \times \sqrt{\frac{1}{0.7^2} - 1} \approx 4.1$ — above most plausible bounds on the maximum SR or the variance of the SDF (e.g., Kozak et al., 2015). Therefore, the condition above is satisfied under very weak assumptions that only rule out near-arbitrage opportunities.

2.2.2 Time-varying Quantities of Risk

We show how to condition down Eq. 1 in the presence of heteroskedasticity and constant risk prices. We rely on the following two assumptions.

Assumption 3. *All risk prices (δs) are constant.*

Eq. 1 in this case takes the simpler form:

$$\mathbb{E}_t\left[r_{t+1}^i\right] + \frac{V_t^i}{2} = \delta^m C_t^{i,m} + \delta^\lambda C_t^{i,\lambda} \tag{12}$$

Additionally, as in Campbell et al. (2015), we assume that a single state variable drives all variances.

Assumption 4. *The conditional variance of the market follows a square-root (discrete-time Cox-Ingersoll-Ross) process:*

$$\begin{aligned} \sigma_{t+1}^2 &= \overline{\sigma^2} + \phi_\sigma\left(\sigma_t^2 - \overline{\sigma^2}\right) + u_{t+1}^\sigma \\ u_{t+1}^\sigma &\sim \mathcal{N}\left(0, k^2\sigma_t^2\right) \end{aligned} \tag{13}$$

where k^2 is a constant relating the the conditional variance of variance to the conditional variance of the market (σ_t^2). For any asset i, $\mathrm{var}_t\left[r_{t+1}^i\right] = k_i^2\sigma_t^2$*. Further, all correlations are constant.*

Note that aggregate discount rates are proportional to the conditional variance of the market,

$$\lambda_t \equiv \mathbb{E}_t\left[r_{t+1}^m\right] + \frac{V_t^m}{2} = \phi_{m,\sigma}\sigma_t^2$$

where $\phi_{m,\sigma} = \delta^m + \delta^\sigma k\rho_{m,\sigma}$ and $\rho_{m,\sigma} = \mathrm{corr_t}\left(r_{t+1}^m, u_{t+1}^\sigma\right)$. Therefore, we can combine Eq. 12 and Assumption 4 to recast the pricing equation as:

$$\mathbb{E}_t\left[r_{t+1}^i\right] + \frac{V_t^i}{2} = \delta^m C_t^{i,m} + \delta^\sigma C_t^{i,\sigma} \tag{14}$$

with the following equations describing the mapping from Eq. 7 and Eq. 12 to Eq. 13 and Eq. 14:

$$\lambda_{t+1} = \overline{\lambda} + \phi_{m,\sigma}\left(\lambda_t - \overline{\lambda}\right) + u_{t+1}^\lambda \tag{15}$$

$$C_t^{i,\lambda} \equiv \phi_{m,\sigma}C_t^{i,\sigma} \tag{16}$$

where $\overline{\lambda} = \phi_{m,\sigma}\overline{\sigma^2}$ and $u_{t+1}^\lambda = \phi_{m,\sigma}u_{t+1}^\sigma$.

Again, consider the cross-section of market-neutral log excess returns rx_t^i, i.e., any asset for which for any t: $C_t^{i,m} = 0$. Further consider assets which are re-scaled to have constant variance so $V_t^i = V^i$ and $C_t^{i,j} = \sigma_i\sigma_{j,t}\rho_{i,j}$. Solve for the unconditional covariance, $\mathrm{cov}\left(rx_t^i, \lambda_t\right)$:

$$\begin{aligned}
\mathrm{cov}\left(rx_t^i, \lambda_t\right) &= \mathbb{E}\left[\mathrm{cov}_t\left(rx_t^i,\ \phi_{m,\sigma}u_t^\sigma\right)\right] + \mathrm{cov}\left[\mathbb{E}_{t-1}\left(rx_t^i\right),\ \mathbb{E}_{t-1}\left(\phi_{m,\sigma}\sigma_t^2\right)\right] && (17)\\
&= \phi_{m,\sigma}\mathbb{E}\left[C_t^{i,\sigma}\right] + \mathrm{cov}\left[\delta^\sigma C_t^{i,\sigma},\ \phi_{m,\sigma}\phi_\sigma\sigma_t^2\right] && (18)\\
&= \phi_{m,\sigma}\mathbb{E}\left[C_t^{i,\sigma}\right] + \mathrm{cov}\left[\delta^\sigma k\sigma_i\rho_{i,\sigma}\sigma_t,\ \phi_{m,\sigma}\phi_\sigma\sigma_t^2\right] && (19)\\
&= \phi_{m,\sigma}\rho_{i,\sigma}\sigma_i k\mathbb{E}\left[\sigma_t\right] + \phi_{m,\sigma}\rho_{i,\sigma}\sigma_i k \times \delta^\sigma\phi_\sigma\mathrm{cov}\left[\sigma_t,\ \sigma_t^2\right] && (20)\\
&= \mathbb{E}\left[C_t^{i,\lambda}\right]\left(1 + \delta^\sigma\phi_\sigma\frac{\mathrm{cov}\left[\sigma_t,\ \sigma_t^2\right]}{\mathbb{E}\left[\sigma_t\right]}\right) && (21)
\end{aligned}$$

where $\mathbb{V}\left[\sigma_t^2\right] \equiv \text{var}\left(\sigma_t^2\right)$. Taking unconditional expectations of the conditional model gives:

$$\mathbb{E}\left[rx_t^i\right] + \frac{V^i}{2} = \delta^\lambda \times \mathbb{E}\left[C_t^{i,\lambda}\right] = \hat{\delta}^\lambda \times \text{cov}\left(rx_t^i, \lambda_t\right) \tag{22}$$

where

$$\hat{\delta}^\lambda = \delta^\lambda \left(1 + \delta^\sigma \phi_\sigma \frac{\text{cov}\left[\sigma_t,\, \sigma_t^2\right]}{\mathbb{E}\left[\sigma_t\right]}\right)^{-1}. \tag{23}$$

The next theorem gives conditions under which we can learn about sign $\left[\delta^\lambda\right]$ from sign $\left[\hat{\delta}^\lambda\right]$.

Theorem 3. *Absence of near-arbitrage opportunities implies that* sign $\left[\hat{\delta}^\lambda\right]$ = sign $\left[\delta^\lambda\right]$, *i.e., the sign is fully preserved.*

The sign is fully preserved if and only if $1 + \delta^\sigma \phi_\sigma \frac{\text{cov}\left[\sigma_t, \sigma_t^2\right]}{\mathbb{E}\left[\sigma_t\right]} > 0$. Consider a monthly calibration. We estimate $\phi_\sigma = 0.79$, $\text{cov}\left[\sigma_t,\, \sigma_t^2\right] = 8.4 \times 10^{-5}$, and $\mathbb{E}\left[\sigma_t\right] = 5.7\%$ from historical VIX data[10 11]). Substituting these values reduces the condition to $\delta^\sigma > -864$. How should we interpret this restriction on the risk price? The annualized conditional Sharpe ratio of the volatility-mimicking portfolio is $|\delta^\sigma k \sigma_t|$. First we estimate k using the relationship $k^2 = \frac{\mathbb{E}\left[u_{\sigma,t+1}^2\right]}{\bar{\sigma}^2}$ and recover $k \approx 4\%$. Similar to the argument in Theorem 2, consider the 97.5th percentile of σ_t still based on historic VIX). $\delta^\sigma = -864$ gives a "max" Sharpe ratio of 13 — implausibly high. Again, the condition above is satisfied under very weak assumptions that only rule out near-arbitrage opportunities.

2.3 New Measure of the Covariance with Discount Rates

Aggregate discount rates, λ_t, are not directly observable by an econometrician. A vast literature (Bansal et al., 2014, Campbell and Vuolteenaho, 2004, Campbell et al., 2015) employs a VAR setup with macroeconomic and financial variables to predict λ_t and back out the corresponding shocks. A major limitation of such an approach is that it restricts the information set to a small number of variables. Since investors are presumed to condition on all available information, the forecasts from a predictive regression will not equal market expectations. Additionally, even a small bias in the levels of risk premia forecasts often translates

[10] VIX technically gives risk neutral integrated variance over the next month which could differ arbitrarily from the physical expectation.

[11] In annualized easier to understand units, $\mathbb{E}\left[\sigma_t\right] = 20\%$ and $\text{s.d.}\left(\sigma_t\right) = 8\%$.

into significant misspecification of shocks inferred from a VAR. Indeed, Bansal et al. (2014) and Campbell et al. (2015) estimate similar VARs with slightly different state variables and obtain conflicting predictions. Chen and Zhao (2009b) show that VAR estimates are highly sensitive to small changes in specification.

We employ a novel methodology designed to circumvent the issue. We use future realized returns as an unbiased estimator of the current risk premia required by investors. The following theorem illustrates this idea.

Theorem 4. *Using the level of future realized market returns* $R^m_{t+1} = \exp\left(r^m_t\right) - 1$ *in place of* λ_t *delivers a consistent estimate of the covariance, i.e.,*

$$\operatorname{cov}\left(r^i_t, R^m_{t+1}\right) = \operatorname{cov}\left(r^i_t, \lambda_t\right) \tag{24}$$

in population, where $\lambda_t = \mathbb{E}_t\left[r^m_{t+1}\right] + \frac{V^m_t}{2}$.

Proof. For any information set $\mathcal{F}_t$ at time t such that $r^i_t \in \mathcal{F}_t$, $R^m_{t+1} = \mathbb{E}\left[R^m_{t+1} \left|\mathcal{F}_t\right.\right] + \epsilon^m_{t+1} \approx \lambda_t + \epsilon^m_{t+1}$. ϵ^m_{t+1} is the projection error and thus by definition is orthogonal to any information at time t, $\mathbb{E}\left[\epsilon^m_{t+1} \left|\mathcal{F}_t\right.\right] = 0$. Unconditional covariances are therefore equal in population,

$$\operatorname{cov}\left(r^i_t, R^m_{t+1}\right) = \operatorname{cov}\left(r^i_t, \mathbb{E}\left[R^m_{t+1} \left|\mathcal{F}_t\right.\right]\right) + \operatorname{cov}\left(r^i_t, \epsilon^m_{t+1}\right) = \operatorname{cov}\left(r^i_t, \lambda_t\right) \tag{25}$$

□

Theorem 4 allows us to simply substitute R^m_{t+1} in place of λ_t and shows that this approach delivers a consistent estimate of $\operatorname{cov}\left(r^i_t, \lambda_t\right)$ in population. The in-sample estimator $\widehat{\operatorname{cov}}\left(r^i_t, R^m_{t+1}\right)$ of this covariance is, however, noisy since $\operatorname{var}\left(\epsilon^m\right) \gg \operatorname{var}\left(\epsilon^\lambda\right)$. We find that, provided that λ_t is persistent, we can increase the power of the estimator by proxying for λ_t with the sum of future realized excess returns, $\hat{\lambda}_{t:t+T} = \sum_{j=0}^{T} R^m_{t+j+1}$. We establish via simulation that estimation is robust to varying T, with the optimum around 12 months. In Appendix D.2 we verify this result empirically. Further, under Assumption 2, the following identity holds in population:

$$\operatorname{cov}\left(r^i_t, \hat{\lambda}_{t:t+T}\right) = \left(\frac{1 - \phi^{T+1}}{1 - \phi}\right) \times \operatorname{cov}\left(r^i_t, \lambda_t\right) \tag{26}$$

and hence the price of risk corresponding to the covariance $\operatorname{cov}\left(r_t^i, \hat{\lambda}_{t:t+T}\right)$ is scaled by a positive constant $\left(\frac{1-\phi^{T+1}}{1-\phi}\right)^{-1}$ relative to the price of risk implied by Eq. 4.

Naturally, $\hat{\lambda}_{t:t+T} = \sum_{j=0}^{T} R_{t+j+1}^m$ is a rather imprecise estimator of the *level* of risk premia λ_t; however, it proves to be informative for the purpose of estimating unconditional covariances $\operatorname{cov}(r_t^i, \lambda_t)$. Therefore, we are able to estimate Eq. 4 without directly estimating the level of market risk premium λ_t. We establish statistical and economic significance of our estimator in Section 3. The idea behind using future realized returns as a proxy for contemporaneous discount rates is related to Kelly and Pruitt (2013). In that paper the authors use the Partial Least Squares (PLS) algorithm to construct a predictor of equity returns using the cross-section of dividend-price ratios. The methodology condenses the cross-section of price-dividend ratios according to covariance with the forecast target – future realized returns on the stock market. They therefore use a projection of future realized market returns onto a set of dividend-price ratios as their predictor. Since we are only interested in estimating the covariance with aggregate discount rates rather than predicting the market in the time series, we can side-step the projection step and use future returns directly and maintain consistency of the estimate.

2.4 Empirical Specification

Combining the previous results in Section 2.2 and Section 2.3:

$$\mathbb{E}\left[rx_t^i\right] + \frac{V^i}{2} = \tilde{\delta}^{\hat{\lambda}} \times \operatorname{cov}\left(r_t^i, \hat{\lambda}_{t:t+T}\right) \tag{27}$$

where $\hat{\lambda}_{t:t+T} \equiv R_{t+1:t+T+1}^m = \sum_{j=0}^{T} R_{t+j+1}^m$ are future market excess returns cumulated over $T+1$ periods and $\tilde{\delta}^{\hat{\lambda}} = \hat{\delta}^{\lambda}\left(\frac{1-\phi^{T+1}}{1-\phi}\right)^{-1}$.

We can easily accommodate the case when the cross-section of assets is not market-neutral by adding the market factor to our specification:

$$\mathbb{E}\left[R_t^i\right] = \hat{\delta}^m \times \mathbb{E}\left[C_t^{i,m}\right] + \tilde{\delta}^{\hat{\lambda}} \times \operatorname{cov}\left(r_t^i, \hat{\lambda}_{t:t+T}\right) \tag{28}$$

where $R_t^i \equiv \left[\exp\left(r_t^i\right) - 1\right]$ is the level of excess returns on an asset i. We provide required derivations in Appendix A. Note that the first term in Eq. 28 requires computing as asset's expected conditional covariance with market returns while the second term contains the unconditional covariance with the level of discount rates.

All quantities in Eq. 28 are observable or easily computable and thus the relation can be estimated by standard methods. Furthermore, we argued in Theorem 1, Theorem 2, and Theorem 3 that a negative estimate of the unconditional price of the discount-rate risk $\tilde{\delta}_{\hat{\lambda}}$ implies a negative value of the expected conditional price of risk, $\mathbb{E}\left[\delta_t^{\lambda}\right] < 0$ in Eq. 1.

3 Empirical Link Between Cross-Sectional and Aggregate Expected Returns

We estimate the expected return relation of Eq. 28 using three sets of test assets. The first is the canonical 25 portfolios formed by a two-way sort of firms on market capitalization (ME) and book-to-market ratio (BE/ME), available at Ken French's website.[12] Lewellen et al. (2010) highlight a key issue in estimating and testing asset pricing models. When the test assets have a strong factor structure that captures much of the time-series variation as well as the cross-sectional variation in expected returns, a spurious model with many factors may still produce a remarkably good cross-sectional fit as long as the spurious factors are correlated with the "true" factors. This result is not due to sampling variation; it holds in population. A solution they propose is to add assets which increase the "dimensionality" of the test asset space.

In addition to the canonical 25 portfolios, we construct a second alternative set of test assets. We include fifteen portfolios consisting of five value-weighted quintile portfolios each from independent sorts on size, book-to-market ratio, and momentum (prior 2-12)[13]. The momentum factor, UMD (Carhart, 1997), is nearly uncorrelated with the size factor, SMB, and is negatively correlated with the book-to-market factor, HML (Fama and French, 1996). Further, sorting firms based on prior performance produces a reliable spread in average returns subsumed by neither the size effect nor the book-to-market effect (Fama and French, 2008). Therefore, including momentum sorted portfolios as test assets makes it decidedly more difficult for a model to fit the cross-section of expected returns. Our preferred estimation uses these fifteen portfolios; for robustness and for comparison with the literature, we

[12] `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`
[13] Also available at Ken French's website.

perform all estimation using the Fama-French 25 portfolios as well.

Third, we estimate the model using 15 long-short portfolios which capture many prominent features (anomalies) in the cross-section of returns.[14] As shown in Kozak et al. (2015), these returns have weaker factor structure than the Fama-French 25 portfolios, resulting in many more "effective" test assets. Finally, we always include the value-weight market and risk-free returns.

Following the spirit of Merton (1973), we use portfolio returns measured at daily frequency[15]. All returns are measured over the period 01-Aug-1966 to 31-Dec-2013. Using daily returns, rather than monthly, reduces the approximation error due to linearization of the exponential function that we rely on in deriving Eq. 28. As noted in Campbell and Vuolteenaho (2004), "July 1963 is when COMPUSTAT data become reliable and most of the evidence on the book-to-market anomaly is obtained from the post-1963 period". Furthermore, in the pre-1963 sample, the "CAPM explains the cross-section of stock returns reasonably well" (Campbell and Vuolteenaho, 2004). Since the beta arbitrage strategy can only be constructed from 01-Aug-1966, we start our sample on that date[16].

As a proxy for the excess return on the wealth portfolio, r_t^m, we use the log excess return on the value-weight portfolio of all common equity traded on the NYSE, AMEX, and NASDAQ. Of course the standard critique applies that there exist many assets, both traded (foreign securities) and non-traded (real-estate, human capital) that are not included in this portfolio (Roll, 1977). As discussed above, we construct $\hat{\lambda}_t = \sum_{i=1}^{H} r_{t+i}^m$. For our preferred specification, we set $H = 126$ trading days, or one-half year. Our results are quantitatively robust across various choices of H, using daily or monthly frequency of returns (see Appendix D).

Table 1 shows the estimated covariances of asset returns with the factors. Panel A shows $\text{cov}\left(r_t^i, r_t^m\right)$; Quintile 1 represents large firms, growth firms, and recent losers in relation to the dimensions, size, book-to-market, and momentum, respectively. Analogously, Quintile 5 represents small firms, value firms, and recent winners. The column to the right of Quintile 5

[14]We include the most of the anomalies listed in Novy-Marx and Velikov (2014) but exclude strategies which: (1) which cannot be constructed from 1963 onwards, (2) are high turnover (such as monthly rebalanced net issuance), and (3) gross margins and asset turnover, since they are subsumed by gross profitability, as shown in Novy Marx (2013).

[15]We replicate the analysis at monthly frequency and obtain very similar results (see Appendix D).

[16]To sort portfolios on $\beta_{i,m}$, we use three years of daily data to estimate pre-ranking values.

Table 1: Covariances

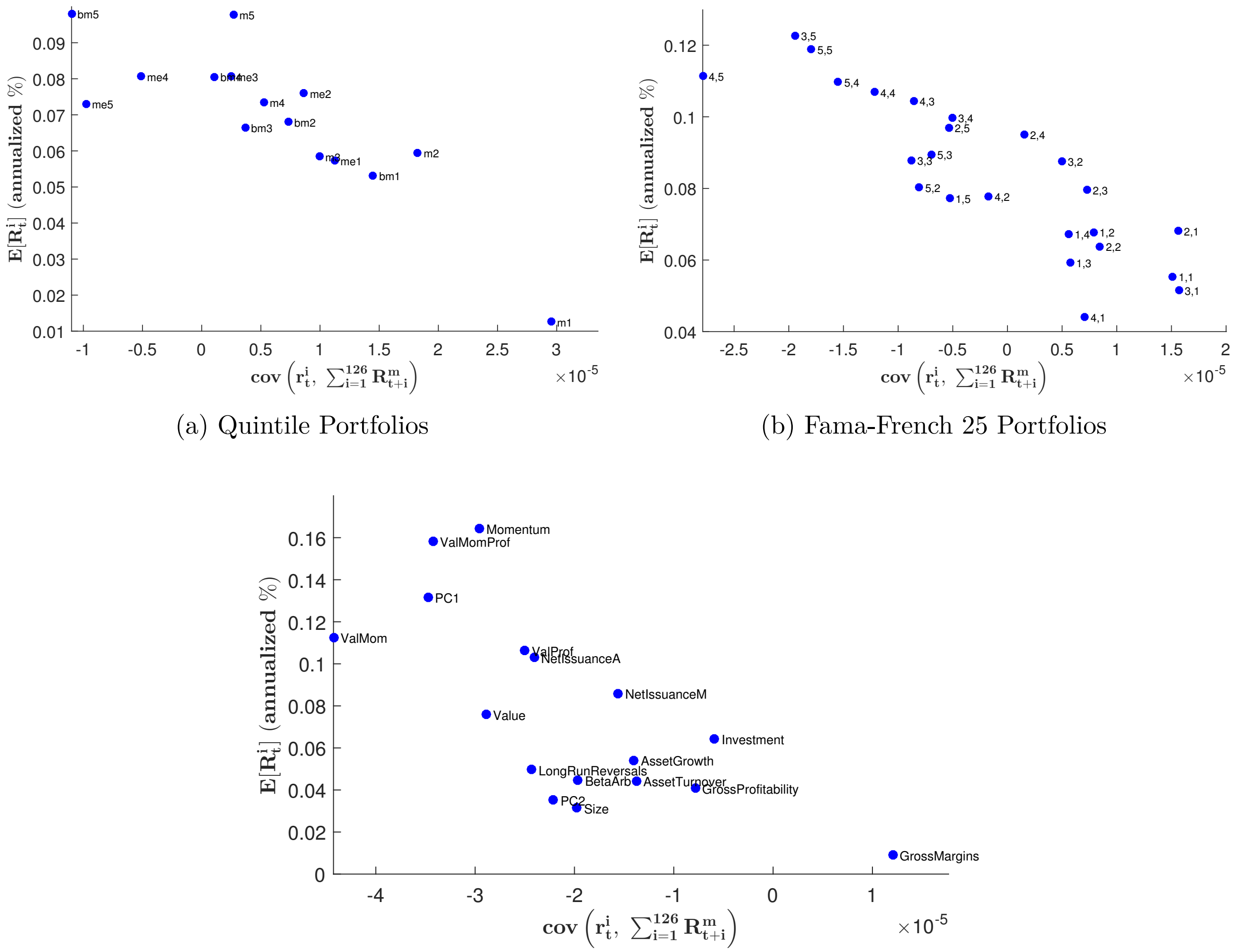
This table shows covariances and annualized mean returns estimated over 01-Aug-1966 to 31-Dec-2013. Panel A lists the covariances of portfolio returns with the market return, $C_{i,m} = \text{cov}\left(rx_t^i, rx_t^m\right)$. Panel B depicts the covariances of portfolio returns with the risk premium factor, $C_{i,\lambda} = \text{cov}\left(rx_t^i, \hat{\lambda}_t\right)$. Panel C shows the the annualized expected excess returns on each portfolio, $\mathbb{E}\left[R_i^e\right]$. The column "FFC" represents the Fama-French-Carhart portfolios: SMB, HML, and UMD. The last column represents the Q5-Q1 spread portfolio. t-statistics (in parentheses) are adjusted for serial correlation using Newey-West procedure with 252 lags (1 year). All covariances are scaled by the variance of the daily market excess returns.

	Q1	Q2	Q3	Q4	Q5	FFC	Q5-Q1
			Panel A: $\mathbb{E}\left[C_t^{i,m}\right]$				
ME	1.02	0.97	0.95	0.94	0.80	-0.08	-0.21
BE/ME	1.05	0.96	0.93	0.90	0.95	-0.14	-0.10
Prior 2-12	1.20	1.00	0.93	0.92	1.06	-0.09	-0.14
			Panel B: $C^{i,\lambda}$				
ME	0.11	0.08	0.02	-0.05	-0.09	-0.11	-0.20
	(1.2)	(0.7)	(0.2)	(-0.4)	(-0.7)	(-1.9)	(-2.2)
BE/ME	0.14	0.07	0.03	0.01	-0.10	-0.20	-0.24
	(1.5)	(0.7)	(0.3)	(0.1)	(-1.0)	(-2.7)	(-2.9)
Prior 2-12	0.28	0.17	0.09	0.05	0.03	-0.18	-0.26
	(2.0)	(1.8)	(1.1)	(0.5)	(0.2)	(-1.8)	(-2.1)
			Panel C: $\mathbb{E}\left[R_t^i\right]$				
ME	5.73	7.60	8.06	8.07	7.31	1.86	2.28
BE/ME	5.32	6.80	6.65	8.06	9.81	4.58	5.11
Prior 2-12	1.26	5.95	5.84	7.35	9.77	7.79	10.70

represents the Q5-Q1 spread portfolio. The FF column gives the estimates for the canonical Fama-French-Carhart (FFC) factors, SMB, HML, and UMD. The covariances match the well known pattern in market betas – their inability to explain the cross-section of size, book-to-market, and momentum sorted portfolios.

Panel B reports $\text{cov}\left(r_t^i, \hat{\lambda}_t\right)$ for the same portfolios with Newey-West t-statistics in parentheses.[17] In all three dimensions (size, book-to-market, and momentum), $\text{cov}\left(r_t^i, \hat{\lambda}_t\right)$ de-

[17]Moving block bootstrap gives similar standard errors.



(a) Quintile Portfolios

(b) Fama-French 25 Portfolios

(c) Anomaly Portfolios

Figure 1: Univariate fit. The top left plot shows sample values of $\mathbb{E}\left[R_i^e\right]$ vs $\text{cov}\left(rx_t^i, \hat{\lambda}_t\right)$ for the 15 quintile portfolios: 5 size (me), 5 book-to-market (bm) and 5 momentum (m) sorted portofios. The plot in the top right panel depicts same results for the 25 Fama-French portfolios. The first number in portfolio labels refers to ME quintile (1=large; 5=small); the second number corresponds to BE/ME quintile (1=growth; 5=value). The bottom plot is for anomaly long-short portfolios. PC1 and PC2 are the first two principal components of anomaly returns. The sample is from 01-Aug-1966 to 31-Dec-2013.

creases from left to right. That is to say, when the "risk premium", λ_t, rises, small stocks are expected to fall more than large stocks, value stocks are expected to fall more than growth stocks, and recent winners are expected to fall more than recent losers. Though using realized market returns in place of expected returns produces consistent covariance estimates, they are less precisely estimated due to the noise present in realized returns. Still, the covariances of the spread portfolios with $\hat{\lambda}_t$ are statistically significantly different from zero and the covariances follow a reliable pattern, suggesting that our results are not spurious. Panel C shows sample average returns which increase monotonically from left to right across quintiles, consistent with the well known size, value, and momentum phenomena. Panels B and C suggest a strong relationship between $\mathrm{cov}\left(r_t^i, \hat{\lambda}_t\right)$ and $\mathbb{E}\left[R_i^e\right]$, which can be clearly seen in Figure 1.

Figure 1 (a) plots sample values of $\mathbb{E}\left[R_i^e\right]$ vs $\mathrm{cov}\left(r_t^i, \hat{\lambda}_t\right)$ for the 15 quintile portfolios. Figure 1 (b) is the same plot for the 25 Fama-French portfolios and Figure 1 (c) shows the various anomalies. The graphs confirm that $\mathrm{cov}\left(r_t^i, \hat{\lambda}_t\right)$ and $\mathbb{E}\left[R_i^e\right]$ line up well in the cross-section of assets, suggesting the λ_t risk factor rationalizes the size, value, momentum effects, and the various anomalies. The downward sloping pattern of average returns vs. $\mathrm{cov}\left(r_t^i, \hat{\lambda}_t\right)$ strongly suggests a negative price of risk.

3.1 Estimation Results

We estimate the risk price vector $\delta = \left[\delta_m \ \delta_\lambda\right]'$ using GMM with a pre-specified block-diagonal weighting matrix (Cochrane, 2001, Chapter 11.5). It is equivalent to the standard two-stage estimation procedure. Covariances $C_{i,\lambda} \equiv \mathrm{cov}\left(r_t^i, \hat{\lambda}_t\right)$ and $C_{i,m} \equiv \mathbb{E}\left[\mathrm{cov}\left(r_t^i, r_t^m\right)\right]$ are estimated in the first stage by just-identified GMM, which yields the standard formulas for sample covariance. In the second stage, we estimate risk prices (SDF coefficients) via an OLS regression of sample mean returns on the covariances estimated from the first stage. In addition to the two-factor model, we estimate unconditional (standard) versions of the Sharpe-Lintner CAPM and well as the Fama-French-Carhart (FF augmented with the UMD "momentum" factor of Carhart, 1997)[18]. For ease of comparison, all models are written and estimated in terms of covariances instead of regression βs. Below is a summary of the pricing equations, where δs are interpreted as risk prices (coefficients in the SDF):

$$\text{2-Factor model: } \mathbb{E}\left[R_i^e\right] = \alpha + C_{i,m}\delta_m + C_{i,\lambda}\delta_\lambda \tag{29}$$

$$\text{CAPM: } \mathbb{E}\left[R_i^e\right] = \alpha + C_{i,m}\delta_m \tag{30}$$

$$\begin{aligned} \text{4-Factor FFC: } \mathbb{E}\left[R_i^e\right] &= \alpha + C_{i,m}\delta_m + C_{i,smb}\delta_{smb} \\ &\quad + C_{i,hml}\delta_{hml} + C_{i,umd}\delta_{umd} \end{aligned} \tag{31}$$

[18] For these models, $C_{i,m} \equiv \mathrm{cov}\left(r_t^i, r_t^m\right)$

where $C_{i,X} \equiv \text{cov}\left(rx_t^i,\ X_t\right)$. Recall, the key parameter is δ_λ; rational models such as long-run risks or habits predict $\delta_\lambda < 0$ whereas sentiment models of irrational expectations predict $\delta_\lambda > 0$. We also estimate the models under the restriction that the zero-beta rate equals the risk-free rate ($\alpha = 0$).

Estimated risk prices are given in Table 2 along with sample R^2 and mean absolute pricing errors. We follow Campbell and Vuolteenaho (2004) and define R^2 as $1 - \frac{\sum_i e_i^2}{\mathbb{V}\left(\mathbb{E}\left[R_i^e\right]\right)}$, where e_i is the difference between model expected and actual average return for asset i. This allows for negative R^2 in the case of restricted models. Standard errors are calculated using a moving block bootstrap (Horowitz, 2001) and are consistent across various choices of block size.[19] Quantitatively, the two-factor model fits the cross-section of average returns nearly as well as the 4-factor FFC model. The estimated intercept is nearly zero, both statistically and economically. While a good fit is not strictly necessary for the purpose of interpreting δ_λ, it alleviates a concern that we might be estimating a parameter from a badly misspecified model. Though $\text{cov}\left(r_t^i, \hat{\lambda}_t\right)$ is not very well estimated for any individual test asset, the cross-sectional spread in covariances is strong enough to yield precise estimation of δ_λ. $H_0 : \delta_\lambda = 0$ is rejected for all conventional significance levels, implying that discount rate shocks are economically important and lead to significant hedging demand.Covariance with the risk premium factor is able to capture a large portion of the cross-sectional variation in average returns due to the size, book-to-market, and momentum effects. Our estimate of δ_λ is unambiguously negative, providing support for "rational" models of time-varying expected returns. The cross-sectional fit of our 2-factor and 4-factor Fama-French-Carhart model is shown in Figure 2. The graphs plot model implied mean excess returns on the horizontal and sample average returns on the vertical axis. The 45° line represents a model with perfect in-sample fit (100% R^2).

Lewellen et al. (2010) document that when test assets have a strong factor structure, OLS R^2 is a poor guide to model fit, both in sample and (even in) population. Their first suggestion for mitigating this concern is to include more portfolios, "sorted [by] other characteristics" (besides size and B/M) as test assets. We include momentum sorted portfolios, which are notoriously difficult to fit. Second, they suggest to impose model implied restrictions on risk-premia. We already present results with zero-beta rates set to the risk-free rate. Further imposing that the model expected return on the market equals its average

[19] We bootstrap the entire GMM system, so uncertainty in the first-stage covariance estimates is fully incorporated in the standard errors of risk prices.

Table 2: Risk Price Estimates

This table shows risk prices estimates for the two-factor model, CAPM, and Fama-French-Carhart model. The test assets are value-weighted quintile portfolios sorted on ME, BE/ME, and Prior2-12. α is annualized (in %) and "-" indicates that the value is restricted to zero. MAPE is average absolute pricing error, annualized. Moving block bootstrap t-statistics are in parentheses. The sample is from 01-Aug-1966 to 31-Dec-2013.

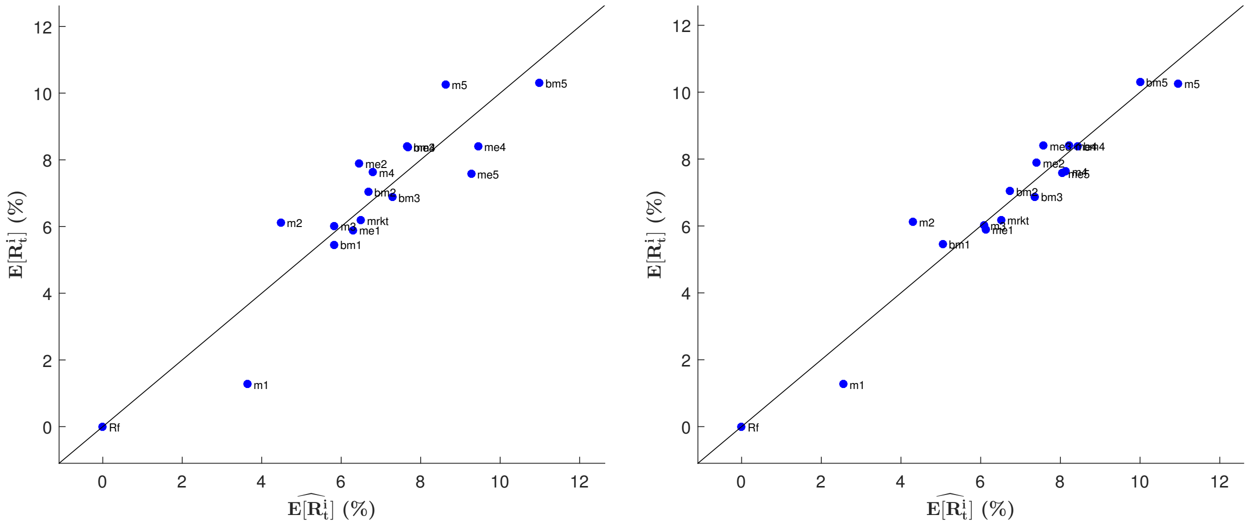
	α	δ^m	δ^λ	δ^{smb}	δ^{hml}	δ^{umd}	R^2	$MAPE$
2-Factor model	-	3.21	-8.73	-	-	-	84	0.816
		(3.2)	(-3.7)					
	0.196	3.13	-8.68	-	-	-	84	0.824
	(0.5)	(3.2)	(-3.8)					
CAPM	-	2.63	-	-	-	-	21.8	1.61
		(2.8)						
	2.1	1.81	-	-	-	-	26.3	1.64
	(3.3)	(1.9)						
4-Factor FFC	-	4.57	-	2.11	9.58	7.2	93.8	0.467
		(3.9)		(0.7)	(2.9)	(4.1)		
	0.184	4.49	-	2.09	9.52	7.16	93.8	0.478
	(1.0)	(4.0)		(0.7)	(2.9)	(4.1)		

leaves our results quantitatively unchanged (unreported). Because the risk-premium factor is not a traded return, we cannot similarly restrict its price of risk. Finally, they recommend reporting confidence intervals for R^2 rather than just point estimates. Bootstrap simulation rejects the null $H_0: R^2 = 0$ with $p \approx 5\%$.

3.2 Anomaly Portfolios

To further test the ability of Eq. 28 to summarize expected returns, we estimate it using a cross-section of anomaly portfolios defined in Novy-Marx and Velikov (2014)[20], which represent a broad set of empirical regularities with seemingly very different fundamental drivers. However, as shown in Kozak et al. (2015), a pricing model using the first few principal components of returns produces high R^2 in fitting the cross-section of average

[20]We construct our own portfolios at daily frequency using their definitions.



(a) 2-Factor model

(b) FFC model

Figure 2: Performance of our 2-factor and 4-Factor Fama-French-Carhart models using quintile portfolios. The plot shows sample average, $\mathbb{E}\left[R_i^e\right]$, vs model expected , $\widehat{\mathbb{E}\left[R_i^e\right]}$, excess returns. The 2-Factor model with restricted intercept on the left, and 4-factor FFC model with restricted intercept on the right. m1-m5 correspond to momentum quintile portfolios (losers to winners). bm1-bm5 correspond to book-to-market quintiles (growth to value). me1-me5 correspond to size quintiles (large to small). The sample is from 01-Aug-1966 to 31-Dec-2013.

returns. This gives us hope that one (or a few) basic economic mechanism is responsible for the variety of anomalies.

Table 3 shows parameters of the pricing models in Equations (29)-(31), estimated using the long-short anomaly returns[21]. We also include two principal components (labeled PC1 and PC2) as test assets. The estimated risk prices are similar to those in Table 2. Now, however, the 2-factor model significantly outperforms the 4-factor FFC model in fitting the cross-section of average returns (with restricted intercept). Notably, the t-statistics on risk prices are much higher than in Table 2. This is due to the weaker factor structure in the anomaly portfolios as compared to the test assets used before (quintile portfolios sorted on ME, BE/ME, and Prior2-12). Weaker cross-sectional correlation of returns results in more

[21]Because the anomaly portfolios are long-short, they tend to have CAPM βs near zero. Sample β estimates are very noisy and yield unreliable estimates of δ_M. To address this issue, we orthogonalize each anomaly return against the market portfolio before estimating asset pricing models. This procedure is equivalent to giving infinite weight to the market portfolio in the second stage estimation (like GLS). Without this restriction, the model R^2 slightly improves, at the cost of very poor fit for $\mathbb{E}\left[R^m\right]$.

Table 3: Risk Price Estimates using Anomaly Portfolios

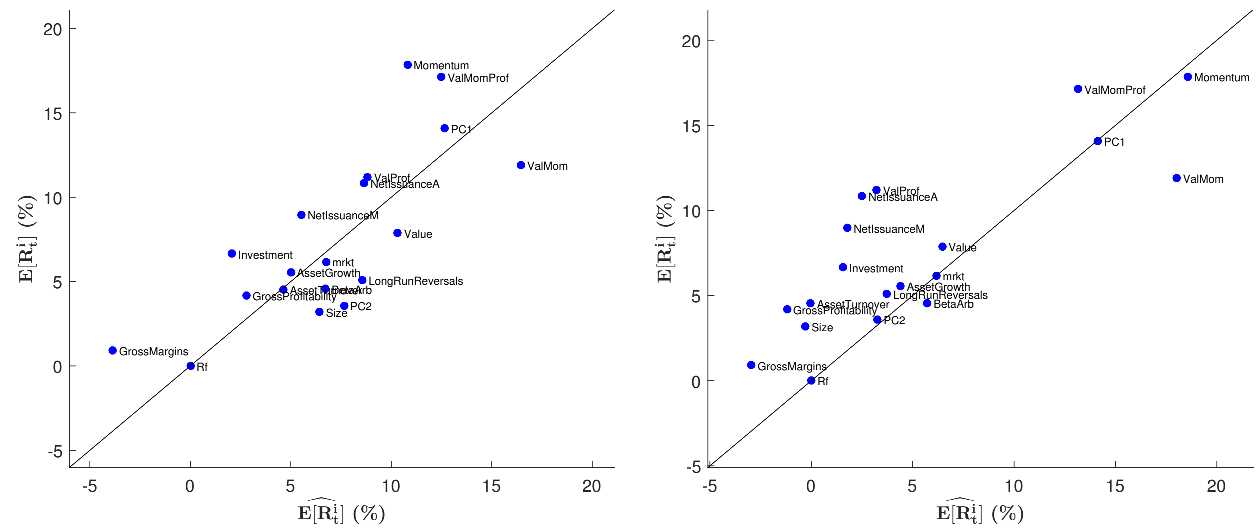
This table shows risk price estimates for our two-factor, CAPM, and FFC models. Test assets are long-short anomaly portfolios. α is annualized and "-" indicates that the intercept is restricted to zero. MAPE is average absolute pricing error, annualized. Moving block bootstrap t-statistics are in parentheses. The sample is from 01-Aug-1966 to 31-Dec-2013.

	α	δ^m	δ^λ	δ^{smb}	δ^{hml}	δ^{umd}	R^2	$MAPE$
2-Factor model	-	3.75	-13.6	-	-	-	56.4	2.6
		(3.7)	(-6.6)					
	1.92	2.77	-10.9	-	-	-	61.5	2.48
	(1.8)	(2.4)	(-5.1)					
CAPM	-	2.21	-	-	-	-	-226	7.31
		(2.2)						
	7.26	-0.319	-	-	-	-	5.43	3.78
	(6.3)	(-0.3)						
4-Factor FFC	-	5.12	-	-2.63	13.1	11.7	25.4	3.15
		(4.2)		(-0.7)	(3.2)	(6.0)		
	4.41	2.13	-	-3.14	6.73	7.7	72.9	1.82
	(4.2)	(1.6)		(-0.8)	(1.7)	(3.7)		

"effective" test assets, improving statistical power. Figure 3 shows graphically the fit of our 2-factor and 4-factor FFC models (with restricted zero-beta rate). The 2-factor model is visibly superior to the 4-factor FFC model in fitting the average anomaly returns. As before, we find $\delta^\lambda < 0$, providing evidence for "rational" models in which time of high aggregate risk premia are "bad" (high marginal utility states).

4 Bond Risks and Risk Premia

In the previous section we showed that relatively "safe" portfolios, such as growth, large, and recent losers stocks, have high covariance with market discount rates and thus tend to be good hedges against equity market turmoils. Government bonds are perhaps the most obvious and likely alternative hedging instruments. Indeed, interest rates tend to fall during times of market turmoils, and bond prices tend to rise. Such an embedded "flight-to-quality" mechanism arises in both types of models: rational and behavioral. Both types imply that



(a) 2-Factor model

(b) FFC model

Figure 3: Performance of our 2-factor and 4-Factor Fama-French models using anomaly portfolios. The plot shows sample average, $\mathbb{E}\left[R_i^e\right]$, vs model expected , $\widehat{\mathbb{E}\left[R_i^e\right]}$, excess returns. The 2-Factor model with restricted intercept on the left, and 4-factor FFC with restricted intercept on the right. Assets are anomaly long-short portfolios from 01-Aug-1966 to 31-Dec-2013. PC1 and PC2 denote two largest principal components of anomaly returns.

bond returns *positively* covary with aggregate equity discount rate shocks.

In this section we study bond and stock risk premia jointly, with particular emphasis on bonds. Numerous papers explore risk premia either for equities or for fixed income securities, yet few study these assets in a unified framework.[22] Our empirical findings are consistent with models' predictions: bond returns covary positively with aggregate equity discount rate shocks and bonds command an unambiguously negative risk premium with respect to discount-rate shocks due to their hedging value. Analyzing bonds in isolation, and thereby ignoring bonds' hedging value with respect to changing *equity* risk premia, misses a substantial component of bond return premia.

4.1 Data

We use zero-coupon treasury yields from Gürkaynak et al. (2006) (GSW), which provides a daily constant maturity yield curve from 1961 onward. Though the data are smoothed by

[22]Recent work in this area includes Koijen et al. (2015).

the use of a Svensson polynomial (extension of Nelson-Siegel), the yields are usually very close to the unsmoothed yields derived using the methodology of Fama and Bliss (1987) and "for many purposes the slight smoothing in GSW data may make no difference" (Cochrane and Piazzesi, 2008). The advantage of GSW yields is the daily observation frequency, which we have argued in Section 3 is important to our empirical strategy. Prior to 1971, the GSW yields only include maturities up to seven years. Post-1971 they includes maturities to 30 years, though there is some question of the reliability of the very long maturity yields. To match the timing of our stock data, we use maturities up to seven years over the same sample period as for stocks, 01-Aug-1966 to 31-Dec-2013. To construct zero-coupon bond returns from the GSW yields, we use the daily parameter estimates available online.[23] This allows us, for example, to recover the yield on a bond with 364 days to maturity. This yield is necessary for calculating the daily return on a one-year bond. For excess returns, we subtract the return on a one month T-bill, the same procedure we use for excess stock returns. We use the 3-month zero-coupon yield[24] as our proxy for the short rate, r_t^f.

4.2 Estimating the price of "level risk"

Studying bonds in isolation, Cochrane and Piazzesi (2008) conclude that expected excess bond returns are well captured by the single-factor model, $E_t\left[r_{i+t}\right] = C_{i,B}\delta_{B,t}$, where $C_{i,B} \equiv \text{cov}_t\left[r_{i,t+1},\, \Delta level_{t+1}\right]$ and $level_{t+1}$ is the "level of interest rates". We proxy for $level_{t+1}$ with r_t^f, the 3-month T-bill rate. We estimate the model using returns on zero-coupon bonds with maturities one to seven years and recover $\delta_B \approx -57$. The cross-sectional R^2 is 95% with 0.08% annualized mean absolute pricing error. Figure 4 shows graphically the good fit of the level model for bonds.

Given our findings in Section 3 and the "flight-to-quality" mechanism in both rational and behavioral models, we argue that δ_B, the price of "level risk", is underestimated in bond-only models. The problem is a classic case of omitted variables bias. Eq. 28 suggests at least two such missing variables, $C_{i,\lambda} = \text{cov}\left(rx_t^i, \sum_{i=1}^H rx_{t+i}^m\right)$ and $C_{i,M} = \text{cov}\left(rx_t^i, rx_t^m\right)$. Table 4 shows $C_{i,B}$, $C_{i,\lambda}$ and $C_{i,m}$ across maturities. First note that $C_{i,M} \approx 0$ for all maturities (bonds have almost zero market βs). More importantly, $\forall i,\; C_{i,\lambda} \approx -15 \times C_{i,B}$. Cross-

[23] http://www.federalreserve.gov/pubs/feds/2006/200628/200628abs.html

[24] GSW note that yields on bonds with maturities shorter than three months are strongly affected by liquidity issues.

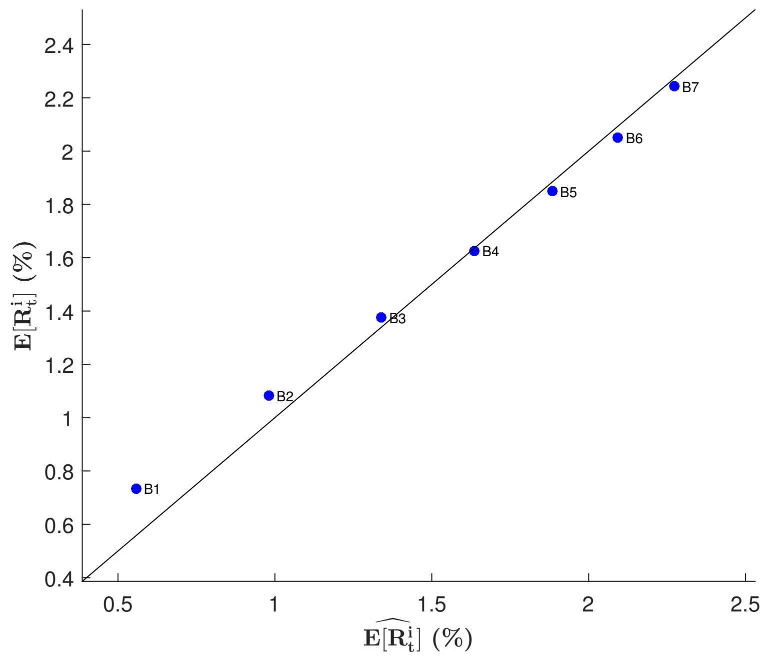


Figure 4: Univariate Bond Pricing. The plot shows sample average, $\mathbb{E}\left[R_i^e\right]$, vs model expected, $\widehat{\mathbb{E}\left[R_i^e\right]}$, excess returns from a single factor bond model.

sectionally, corr $(C_{i,B}, C_{i,\lambda}) \approx 1.0$. Since we know from Section 3.1 that $\delta_\lambda \neq 0$, the univariate level model suffers greatly from omitted variables bias. Using the estimate of $\delta_\lambda = -9$, a back-of-the-envelope calculation suggests the true $\delta_B = -57 - 15 \times 8.8 = -189$. In other words, the required compensation for bearing level risk is much higher than is estimated from a univariate model of bond expected returns. Treasury bonds, in addition to loading positively on level risk, also provide investors a hedge against increases in the risk premium on stocks. Thus, bonds earn lower average excess returns than in a hypothetical economy where the expected excess market return is constant.

This intuition is formalized by estimating the 3-factor bond model given by Eq. 34. Table 5 gives estimated risk prices (δs) from the following models:

$$\text{2-Factor model: } \mathbb{E}\left[R_i^e\right] = C_{i,m}\delta_m + C_{i,\lambda}\delta_\lambda \tag{32}$$

$$\text{Univariate Level Risk: } \mathbb{E}\left[R_i^e\right] = C_{i,B}\delta_B \tag{33}$$

$$\text{3-Factor bond model: } \mathbb{E}\left[R_i^e\right] = C_{i,m}\delta_m + C_{i,\lambda}\delta_\lambda + C_{i,B}\delta_B \tag{34}$$

Table 4: Bond Covariances

This table shows covariances of government bonds with the bond "level" factor, $C_{i,B} = \text{cov}\left(rx_t^i, \Delta r_t^f\right)$; the stock excess returns market factor, $C_{i,m} = \text{cov}\left(rx_t^i, rx_t^m\right)$; and the discount rate factor, $C_{i,\lambda} = \text{cov}\left(rx_t^i, \sum_{i=1}^{H} rx_{t+i}^m\right)$, estimated over 01-Aug-1966 to 31-Dec-2013. t-statistics in parentheses are adjusted for serial correlation using the Newey-West procedure with 252 lags (1 year). All covariances are scaled by the variance of the market daily excess returns. The sample is from 01-Aug-1966 to 31-Dec-2013.

	1Y	2Y	3Y	4Y	5Y	6Y	7Y
$C_{i,B}$	-0.004	-0.006	-0.009	-0.011	-0.012	-0.014	-0.015
$C_{i,M}$	0.002	0.002	0.003	0.004	0.006	0.008	0.010
$C_{i,\lambda}$	0.051	0.094	0.130	0.162	0.191	0.218	0.242
	(3.6)	(3.7)	(3.8)	(4.0)	(4.1)	(4.1)	(4.2)

where $C_{i,B} \equiv \text{cov}\left[rx_t^i, \Delta r_t^f\right]$ and δ_B is the price of the "level" risk. All models are estimated with the intercept restricted to zero. The 2-factor model is estimated using only the stock portfolios from Section 3 (15 quintile portfolios; both stocks and bonds are used as test assets) and hence the risk price estimates are the same as in Section 3.1. The univariate Level Risk model is estimated using only bond excess returns; bonds are also the only test assets. The 3-factor bond model is estimated using all assets, stock portfolios (15 quintile portfolios) as well as bonds. Estimated values for δ_m and δ_λ are essentially unchanged in the 3-factor bond model (relative to the 2-factor estimates). The R^2 of the 2-factor model is so low because bonds are included as test assets. Importantly, δ_B in the 3-factor bond model is $-187 \ll -54$. This is nearly equal to the back of the envelope prediction given above. Table 6 gives annualized percent returns by maturity in sample, as implied by the univariate Level Risk model, and as implied by the 3-factor bond model. It shows the joint stock and bond performs on par with the level risk model in pricing bonds.

Figure 5 shows average returns vs our 3-factor bond model expected returns for bonds and stock portfolios, with model implied mean excess returns on the horizontal and sample average returns on the vertical axes. The 45° line represents a model with perfect in-sample fit (100% R^2). Stocks fit almost as well as in Figure 2 (using our 2-factor model) and bonds fit quite well. It is worth emphasizing that this result is not merely mechanical. Given two factor models, each fitting either cross-section of stocks or bonds, a combined model with all

Table 5: Risk Price Estimates

This table shows premia estimates for our 2-factor model (estimated using 15 stock quintile portfolios only; both bonds and stocks included as test assets), the Level Risk model (estimated using bond returns; only bonds used as test assets), and the 3-factor bond model (estimated using both quintile stock portfolios and bonds to price both). Model intercepts are restricted to zero. MAPE is average absolute pricing error, annualized. Moving block bootstrap t-statistics are in parentheses. The sample is from 01-Aug-1966 to 31-Dec-2013.

	δ_M	δ_λ	δ_B	R^2	$MAPE$
2-Factor model	3.21	-8.73	-	-9.23	2.22
(est. stocks only; pricing bonds & stocks)	(3.2)	(-3.8)			
Level Risk	-	-	-56.8	97.7	0.0613
(bonds only)			(-1.8)		
3-Factor bond model	2.29	-7.61	-187	95.3	0.496
(bonds and stocks)	(2.4)	(-3.8)	(-4.3)		

Table 6: Bond Expected Returns

This table shows bond annualized percent returns by maturity in sample (second column), as implied by the univariate Level Risk model (third column), and as implied by our 3-Factor ICAPM (last column). The sample is from 01-Aug-1966 to 31-Dec-2013.

	Sample Mean	Level Risk	3-Factor bond model
1-year bond	0.73	0.56	0.7
2-year bond	1.1	0.98	1.1
3-year bond	1.4	1.3	1.5
4-year bond	1.6	1.6	1.8
5-year bond	1.8	1.9	2
6-year bond	2	2.1	2.1
7-year bond	2.2	2.2	2.1

factors need not fit the joint cross-section of bonds and stocks (see Koijen et al. 2015).

Figure 6 decomposes the expected excess return on the various bonds. The premium due to market risk, $C_{i,M}$, is excluded since it is negligible for bonds. Bonds earn a large premium for loading on the "level risk", whereas they command a large negative premium for loading on the "risk premium" factor. This is consistent with a "flight-to-quality" (Caballero and Krishnamurthy, 2008) interpretation where investors' appetite for risk falls and they attempt to rebalance their portfolios towards safer securities (like U.S. government debt and "good

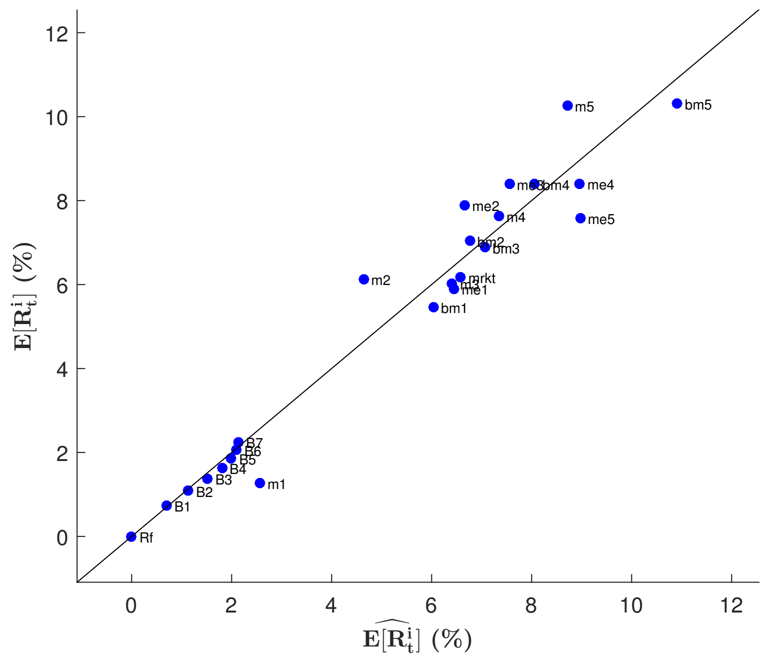


Figure 5: Joint Bond and Stock Pricing. The figure shows average returns vs our 3-factor bond model expected returns for bonds and 15 quintile stock portfolios. The line represents a model with perfect in-sample fit (100% R^2).

companies"). Since it is impossible for everyone to rebalance in this way at the same time, prices adjust instead of quantities. The prices of "risky" assets fall relative to the prices of "safer" assets.

Koijen et al. (2015) have a seemingly similar decomposition, albeit with a very different interpretation. Our 3-factor bond model as well as their model both feature a level factor and a market factor. Instead of our expected stock return factor, they use an expected bond return factor (CP from Cochrane and Piazzesi, 2005). Whereas bond returns load positively on our factor, λ, they load negatively on CP. Koijen et al. (2015) find a positive price of CP risk whereas we find a negative price of λ risk. The product of loading $\times$ risk price yields a negative number in both cases, and hence the pictures look quite similar, but with different interpretation. We find that bonds hedge against increases in expected stock returns but Koijen et al. (2015) find that bonds respond negatively to increases in expected bond returns. Finally, our estimated model produces a term structure of expected returns which is consistent with the data (Table 6). In contrast, the estimates in Koijen et al. (2015) result in a flat term structure.

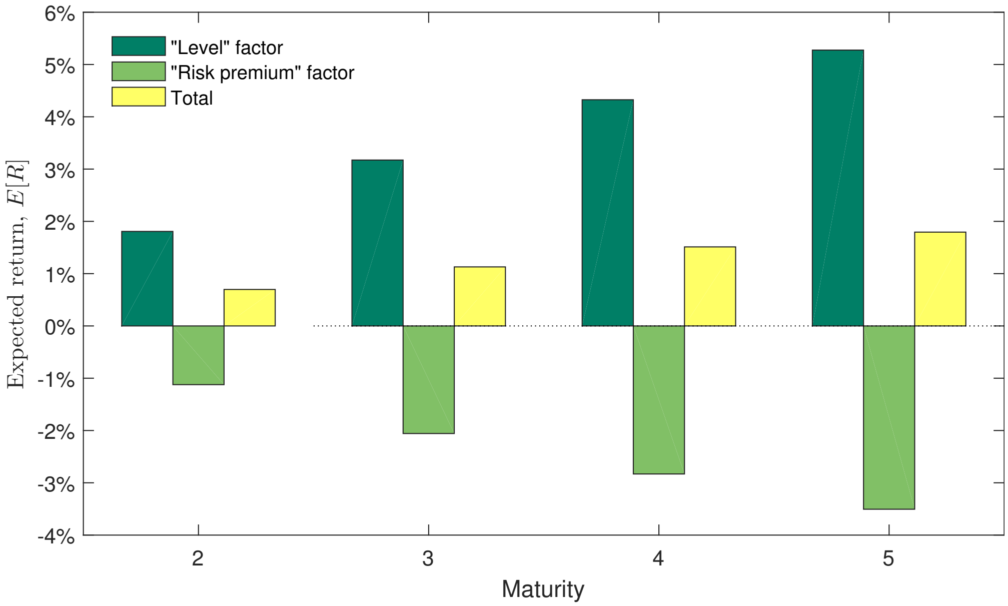


Figure 6: Decomposition of Bond Risk Premia. The figure decomposes the total expected excess return on bonds of maturities 2-5 years (yellow rightmost bars) onto two components: risk premia earned due to the loading on the "level" factor (dark-green leftmost bars), $C_{i,B}$; and risk premia earned due to the loading on the "risk premium" factor (light-green middle bars), $C_{i,\lambda}$. The premium due to market risk, $C_{i,M}$, is excluded since it is negligible for bonds.

5 Predicting the Future Market using Cross-Section

In our empirical methodology, we use future realized excess returns as a proxy for today's market expectation of future excess returns. We further show that this proxy is key in explaining the cross section of stock returns. This observation can be viewed from the reverse perspective. If time-varying expected returns manifest in the cross-section, the cross-section of stock returns can provide information about expected future returns. Indeed, "priced factors ... are innovations in state variables that predict future returns." (Brennan et al., 2004). It is therefore natural to ask whether cross-sectional variables can predict future returns and to what extent. Few recent papers have looked at this question. Kelly and Pruitt (2015) use the cross-section of dividend-price ratios and show that they indeed predict future returns significantly better than the aggregate dividend-price ratio alone.

Our aim is not to construct the optimal predictor; we only want to show that predictability is indeed present and use it as a robustness check of our methodology. If future returns

Table 7: Time-series predictability of market excess returns using Fama-French factors

The table shows time-series predictability of the stock market risk premium using returns on SMB, HML, and UMD. We estimate the following regression:

$$\sum_{t+1}^{t+k} rx^M = a + [DP_t \; MRKT_{t-63:t} \; SMB_{t-63:t} \; HML_{t-63:t} \; UMD_{t-63:t}]\, b + \varepsilon_{M,t+1}.$$

Daily sample from 01-Aug-1966 to 31-Dec-2013. Newey-West t-statistics in parentheses.

	DP	$MRKT$	SMB	HML	UMD	R^2
3 months	0.018	0.049	-	-	-	0.01
	(1.3)	(0.8)				
6 months	0.036	0.015	-	-	-	0.014
	(1.2)	(0.2)				
9 months	0.051	0.0043	-	-	-	0.019
	(1.2)	(0.0)				
12 months	0.065	-0.025	-	-	-	0.023
	(1.2)	(-0.3)				
3 months	0.019	0.064	-0.3	-0.21	-0.18	0.071
	(1.5)	(0.8)	(-2.8)	(-2.8)	(-3.0)	
6 months	0.037	0.0095	-0.38	-0.38	-0.24	0.072
	(1.5)	(0.1)	(-2.4)	(-2.4)	(-2.1)	
9 months	0.052	-0.032	-0.37	-0.44	-0.31	0.069
	(1.4)	(-0.3)	(-2.1)	(-2.0)	(-2.1)	
12 months	0.065	-0.046	-0.41	-0.35	-0.39	0.067
	(1.3)	(-0.3)	(-1.6)	(-1.5)	(-2.4)	

help to explain the cross-section, the cross-section of returns themselves mechanically must predict future returns. We want to ensure the covariances reported in Table 1 and Figure 1 are *economically* significant. As such, we use the returns on SMB, HML, and UMD portfolios to forecast future market returns:

$$\sum_{i=1}^{H} rx^M_{t+i} = a + [DP_t \; MRKT_{t-63:t} \; SMB_{t-63:t} \; HML_{t-63:t} \; UMD_{t-63:t}]\, b + \varepsilon_{M,t+1} \qquad (35)$$

Each of the $MRKT, SMB, HML, UMD$ factors is computed using the past 63 trading days

Table 8: Time-series predictability of market excess returns using anomalies

The table shows time-series predictability of the stock market risk premium using the first principal component of anomaly long-short portfolios. We estimate the following regression:

$$\sum_{t+1}^{t+k} rx^M = a + [DP_t \; MRKT_{t-63:t} \; PC1_{t-63:t}]\, b + \varepsilon_{M,t+1}.$$

Newey-West t-statistics in parentheses. The sample is from 01-Aug-1966 to 31-Dec-2013.

	DP	$MRKT$	NMV_{PC1}	R^2
3 months	0.019	0.044	-	0.011
	(1.3)	(0.6)		
6 months	0.038	0.02	-	0.016
	(1.3)	(0.2)		
9 months	0.056	-0.007	-	0.022
	(1.3)	(-0.1)		
12 months	0.071	-0.017	-	0.027
	(1.3)	(-0.2)		
3 months	0.022	-0.053	-0.26	0.058
	(1.6)	(-0.7)	(-4.7)	
6 months	0.042	-0.099	-0.32	0.051
	(1.5)	(-1.1)	(-3.3)	
9 months	0.059	-0.13	-0.34	0.048
	(1.5)	(-1.1)	(-2.4)	
12 months	0.075	-0.15	-0.37	0.05
	(1.4)	(-1.4)	(-2.3)	

(3 months). Results are robust to varying the lag length.

The top panel of Table 7 reports the coefficient estimates, t-statistics of estimated coefficients in Eq. 35, and R^2 at various horizons, k, (3, 6, 9, and 12 months) with only the market and dividend-price ratio included as predictors. There is little evidence of return predictability at horizons up to one year, as evidenced by the insignificant t-statistics and low R^2. The bottom panel shows results when including $SMB, HML,$ and UMD as additional predictors[25]. We find that all of the coefficients for each variable at 3-9 months horizon are

[25]Estimated coefficients are nearly identical if we exclude $MRKT$ and DP (unreported).

significant and negative and R^2 increases substantially. We conclude that covariances of FFC factors with expected future market returns are *economically* significant. Related evidence of predictability is documented by Liew and Vassalou (2000). They show that SMB and HML help forecast future rates of economic growth.

Finally, we repeat the forecasting exercise using the first principal component of the anomaly long-short returns (PC1). Table 8 shows that PC1 alone has similar forecasting ability[26] to the three FFC factors combined. The statistical significance for PC1 is substantially higher, likely because $SMB, HML,$ and UMD each contain substantial idiosyncratic "noise", adding uncertainty to the estimates.

6 Conclusion

Economists agree that risk premia (expected excess returns) vary significantly over time but substantially disagree as to why. We tackle this issue with a new approach, studying the cross-sectional implications of various models. In models where some traders have biased expectations, rational investors consider times of high expected returns as "good times." In contrast, in models with stochastic preferences (risk-aversion) or technology (habits or long-run risks), investors dislike these states of the world. In both types of models, the logic of Merton (1973) implies that shocks to aggregate risk premia should enter the pricing kernel. However, the two types give opposite predictions for sign of the equilibrium price of risk for such shocks.

We estimate a general ICAPM representation of returns which obtains in both types of models. Instead of employing the typical VAR approach for estimating shocks to discount rates, we overcome the unobservability of expected market returns by using future realized returns as a proxy. This allows us to estimate factor loadings without actually observing the factor itself. Using this "model-free" empirical strategy and return data on a large set well known asset pricing anomalies, (including size, value, and momentum) we estimate a negative price of risk. Our main conclusion is that shocks to aggregate expected excess returns are perceived as "bad" by rational traders, who requires higher expected returns for holding exposure to these shocks. This evidence is consistent with "rational" explanations for time-varying risk premia and is inconsistent with canonical models of biased expectations.

[26]As measured by R^2

References

Bansal, R., D. Kiku, I. Shaliastovich, and A. Yaron (2014). Volatility, the Macroeconomy, and Asset Prices. *The Journal of Finance 69*(6), 2471–2511.

Bansal, R. and A. Yaron (2004). Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles. *The Journal of Finance 59*(4), 1481–1509.

Barberis, N., R. Greenwood, L. Jin, and A. Shleifer (2015). X-CAPM: An extrapolative capital asset pricing model. *Journal of Financial Economics 115*(1), 1–24.

Brennan, M., A. Wang, and Y. Xia (2004). Estimation and test of a simple model of intertemporal capital asset pricing. *The Journal of Finance 59*(4), 1743–1776.

Caballero, R. and A. Krishnamurthy (2008). Collective risk management in a flight to quality episode. *The Journal of Finance 63*(5), 2195–2230.

Campbell, J. and T. Vuolteenaho (2004). Bad Beta, Good Beta. *American Economic Review 94*(5), 1249–1275.

Campbell, J. Y. and J. H. Cochrane (1999). By Force of Habit: A Consumption Based Explanation of Aggregate Stock Market Behavior. *Journal of Political Economy 107*(2), pp. 205–251.

Campbell, J. Y., S. Giglio, C. Polk, and R. Turley (2015). An Intertemporal CAPM with Stochastic Volatility. Technical report.

Campbell, J. Y. and A. S. Kyle (1993). Smart money, Noise Trading and Stock Price Behaviour. *The Review of Economic Studies 60*(1), 1–34.

Campbell, J. Y. and L. M. Viceira (1999). Consumption and portfolio decisions when expected returns are time varying. *Quarterly Journal of Economics 114*(2), 433–495.

Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *Journal of Finance 52*, 57–82.

Chen, L. and X. Zhao (2009a). Return decomposition. *Review of Financial Studies 22*(12), 5213–5249.

Chen, L. and X. Zhao (2009b). Return Decomposition. *Review of Financial Studies*, hhp017.

Cochrane, J. H. (2001). *Asset Pricing, 2001.* Princeton University Press.

Cochrane, J. H. (2011). Presidential Address: Discount Rates. *The Journal of Finance 66*(4), 1047–1108.

Cochrane, J. H. and M. Piazzesi (2005). Bond Risk Premia. *American Economic Review*, 138–160.

Cochrane, J. H. and M. Piazzesi (2008). Decomposing the yield curve. *Graduate School of Business, University of Chicago, Working Paper*.

Constantinides, G. M. (1990). Habit Formation: A Resolution of the Equity Premium Puzzle. *Journal of Political Economy 98*(3), 519–543.

Dew Becker, I. (2011). A model of time-varying risk premia with habits and production.

Drechsler, I. (2013). Uncertainty, Time-Varying Fear, and Asset Prices. *The Journal of Finance 68*(5), 1843–1889.

Fama, E. and R. Bliss (1987). The information in long-maturity forward rates. *The American Economic Review*, 680–692.

Fama, E. and K. French (2008). Dissecting anomalies. *The Journal of Finance 63*(4), 1653–1678.

Fama, E. F. and K. R. French (1996). Multifactor Explanations of Asset Pricing Anomalies. *Journal of Finance 51*, 55–84.

Giovannini, A. and P. Weil (1989). Risk aversion and intertemporal substitution in the capital asset pricing model. Technical report, National Bureau of Economic Research.

Gürkaynak, R. S., B. Sack, and J. H. Wright (2006, June). The U.S. Treasury Yield Curve: 1961 to the Present.

Horowitz, J. L. (2001). *Handbook of Econometrics - Chapter 52: The Bootstrap*, Volume 5. Elsevier B.V.

Kelly, B. and S. Pruitt (2013). Market Expectations in the Cross-Section of Present Values. *The Journal of Finance 68*(5), 1721–1756.

Kelly, B. and S. Pruitt (2015). The three-pass regression filter: A new approach to forecasting using many predictors. *Journal of Econometrics 186*(2), 294–316.

Kim, T. S. and E. Omberg (1996). Dynamic nonmyopic portfolio behavior. *Review of financial studies 9*(1), 141–161.

Koijen, R., H. Lustig, and S. Van Nieuwerburgh (2015). The cross-section and time-series of stock and bond returns. Technical report.

Kozak, S. (2015). *Dynamics of Bond and Stock Returns*. Ph. D. thesis, University of Michigan.

Kozak, S., S. Nagel, and S. Santosh (2015). Interpreting factor models.

Lewellen, J., S. Nagel, and J. Shanken (2010). A skeptical appraisal of asset pricing tests. *Journal of Financial Economics 96*, 175–194.

Liew, J. and M. Vassalou (2000). Can book-to-market, size and momentum be risk factors that predict economic growth? *Journal of Financial Economics 57*(2), 221–245.

Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica: Journal of the Econometric Society*, 867–887.

Novy Marx, R. (2013). The Other Side of Value: The Gross Profitability Premium. *Journal of Financial Economics 108*(1), 1–28.

Novy-Marx, R. and M. Velikov (2014). A taxonomy of anomalies and their trading costs. Technical report, National Bureau of Economic Research.

Roll, R. (1977). A critique of the asset pricing theory's tests Part I: On past and potential testability of the theory. *Journal of financial economics 4*(2), 129–176.

Sabbatucci, R. (2015). Are dividends and stock returns predictable? New evidence using M&A cash flows. *New Evidence Using M&A Cash Flows (October 24, 2015).*

Appendix

A Unconditional Pricing Model: Non-Market-Neutral Assets

We provide extensions of our arguments in Section 2.2 to the case when test assets have non-zero covariances with the aggregate market return.

A.1 Time-varying Prices of Risk

Consider the cross-section of assets with non-zero exposure to the market:

$$\begin{aligned} \operatorname{cov}\left(r_t^i, \lambda_t\right) &= C^{i,\lambda} + \operatorname{cov}\left(\delta_t^m C^{i,m} + \delta_t^\lambda C^{i,\lambda}, \phi\lambda_t\right) \\ &= \left[1 + \phi \operatorname{cov}\left(\delta_t^\lambda, \lambda_t\right)\right] \times C^{i,\lambda} + \phi \operatorname{cov}\left(\delta_t^m, \lambda_t\right) C^{i,m} \end{aligned} \tag{36}$$

Finally, by substituting this expression into Eq. 6 and taking unconditional expectations we obtain the unconditional relation:

$$\mathbb{E}\left[r_t^i\right] + \frac{V^i}{2} \quad = \quad \hat{\delta}^m C^{i,m} + \hat{\delta}^\lambda \operatorname{cov}\left(r_t^i, \lambda_t\right) \tag{37}$$

where

$$\begin{aligned} \hat{\delta}^m &= \bar{\delta}_t^m - \phi\hat{\delta}^\lambda \operatorname{cov}\left(\delta_t^m, \lambda_t\right) \\ \hat{\delta}^\lambda &= \bar{\delta}^\lambda \left[1 + \phi \operatorname{cov}\left(\delta_t^\lambda, \lambda_t\right)\right]^{-1} \end{aligned}$$

and $\bar{\delta}^m = \mathbb{E}\left[\delta_t^m\right]$, $\bar{\delta}^\lambda = \mathbb{E}\left[\delta_t^\lambda\right]$.

A.2 Time-varying Quantities of Risk

Consider the cross-section of assets with non-zero exposure to the market:

$$\begin{aligned} \operatorname{cov}\left(r_t^i, \lambda_t\right) &= \mathbb{E}\left[\operatorname{cov}_t\left(r_t^i,\ \phi_{m,\sigma} u_t^\sigma\right)\right] + \operatorname{cov}\left[\mathbb{E}_{t-1}\left(r_t^i\right),\ \mathbb{E}_{t-1}\left(\phi_{m,\sigma}\sigma_t^2\right)\right] & (38) \\ &= \phi_{m,\sigma}\mathbb{E}\left[C_t^{i,\sigma}\right] + \operatorname{cov}\left[\delta^m C_t^{i,m} + \delta^\sigma C_t^{i,\sigma},\ \phi_{m,\sigma}\phi_\sigma\sigma_t^2\right] & (39) \\ &= \mathbb{E}\left[C_t^{i,\lambda}\right]\left(1 + \delta^\sigma\phi_\sigma\frac{\mathbb{V}\left[\sigma_t^2\right]}{\mathbb{E}\left[\sigma_t^2\right]}\right) + \delta^m\phi_{m,\sigma}\phi_\sigma\frac{\mathbb{V}\left[\sigma_t^2\right]}{\mathbb{E}\left[\sigma_t^2\right]}\mathbb{E}\left[C_t^{i,m}\right] & (40) \end{aligned}$$

where $\mathbb{V}\left[\sigma_t^2\right] \equiv \text{var}\left(\sigma_t^2\right)$. Taking unconditional expectations of the conditional model gives:

$$\mathbb{E}\left[r_t^i\right] + \frac{V^i}{2} = \hat{\delta}^m \mathbb{E}\left[C_t^{i,m}\right] + \hat{\delta}^\lambda \times \text{cov}\left(r_t^i, \lambda_t\right) \tag{41}$$

where

$$\hat{\delta}^m = \delta^m \left(1 - \hat{\delta}^\lambda \phi_{m,\sigma}\phi_\sigma \frac{\mathbb{V}\left[\sigma_t^2\right]}{\mathbb{E}\left[\sigma_t^2\right]}\right) \tag{42}$$

$$\hat{\delta}^\lambda = \delta^\lambda \left(1 + \delta^\sigma \phi_\sigma \frac{\mathbb{V}\left[\sigma_t^2\right]}{\mathbb{E}\left[\sigma_t^2\right]}\right)^{-1}. \tag{43}$$

A.3 Empirical Specification

Both in the case of time-varying prices of risk and in the case of time-varying quantities of risk we thus estimate the model

$$\mathbb{E}\left[r_t^i\right] + \frac{V^i}{2} = \hat{\delta}^m \times \mathbb{E}\left[C_t^{i,m}\right] + \hat{\delta}^\lambda \times \text{cov}\left(r_t^i, \lambda_t\right). \tag{44}$$

The price of market risk $\hat{\delta}^m$ is distorted, but the price of discount-rate risk $\hat{\delta}^\lambda$ is the same as in our derivations in Section 2.2. Therefore all the theorems about the sign of the discount-rate risk (Theorem 1, Theorem 2, Theorem 3) go through.

B Bootstrap

We construct standard errors for risk prices using the moving block bootstrap procedure as follows. There are N test assets, k factors, and T periodic observations. All moments are sample moments taken as expectations across T. The general model is $r_t = C'\delta + \varepsilon_t$. C is an $N \times k$ matrix of univariate covariances, $\text{cov}\left(r_t,\, f_t\right)$, where f_t are the k factors. Notice the model is homoskedastic. δ is the vector of risk prices, and ε_t is the vector of pricing errors. The null hypothesis is that $\delta = 0$ and $E\left[\varepsilon_t\right] = 0$. The alternative is $\delta \neq 0$.

Bootstrap procedure:

1. Estimate $\widehat{C}$ and $\widehat{\delta}$ via GMM

2. Construct $\widetilde{r}_t = r_t - \mathbb{E}\left[r_t\right]$

   (a) $\widetilde{r}_t$ is satisfies the null hypothesis of risk-neutrality and maintains all other properties of the true DGP which are shared with the null. In particular, $\text{cov}\left(\widetilde{r}_t,\, f_t\right) = \text{cov}\left(r_t,\, f_t\right)$

3. Let L be the bootstrap window width. Let $X = \begin{bmatrix} \widetilde{r}_1' & f_1' \\ \vdots & \vdots \\ \widetilde{r}_T' & f_T' \\ \widetilde{r}_1' & f_1' \\ \vdots & \vdots \\ \widetilde{r}_L' & f_L' \end{bmatrix}$. To generate bootstrap sample i, randomly draw j from $U[1, T]$. Let $s_j = X(i : i+L, :)$ in Matlab's indexing convention. Append s_j to X_i, which is initialized as $[\emptyset]$. Repeat until X_i is of length T. Unless T/L is an integer, the process yields a bootstrap sample of incorrect length. Build X_i to be at least length T then trim.

4. Estimate the two-stage regression on sample X_i and save the estimate $\widehat{\delta}_i$

5. Repeat B times (we use 100,000 replications). The estimated $\widehat{\delta}_i$ should be approximately mean zero, and std $\left(\widehat{\delta}_i\right) \approx \text{SE}\left(\widehat{\delta}\right)$

6. Perform usual asymptotic tests

C Model Calibration

We present a model (and calibration) featuring time-varying risk-aversion. Consider the dynamic problem of an investor with Epstein-Zin preferences and unitary elasticity of intertemporal substitution ($\psi = 1$) and time varying coefficient of risk aversion (γ_t). Denote his wealth at time t as W_t, his vector of additional state variables at t as $X_t = \{\alpha_t\}$, and let $\alpha_t \equiv 1 - \gamma_t$, $\rho \equiv 1 - \frac{1}{\psi}$. Investor's value function $J(W_t, X_t)$ is given by

$$\begin{aligned} J(W_t, X_t) &= \max_{\{C,\Theta_t\}} \lim_{\rho \to 0} \left\{ (1-\delta) C_t^{\rho} + \delta \left(\mathbb{E}_t \left[J(W_{t+1}, X_{t+1})^{\alpha_t} \right] \right)^{\frac{\rho}{\alpha_t}} \right\}^{\frac{1}{\rho}} \\ &= \max_{\{C,\Theta_t\}} \left\{ C_t^{1-\delta} \left(\mathbb{E}_t \left[J(W_{t+1}, X_{t+1})^{\alpha_t} \right] \right)^{\frac{\delta}{\alpha_t}} \right\}, \end{aligned} \tag{45}$$

where C_t is investor's consumption at time t and θ_t is a vector of weights allocated to each asset in his portfolio.

Let R_{t+1} denote a vector of returns on n assets available to an investor and let $r_t \equiv \ln(R_{t+1})$. His budget constraint is given by

$$W_{t+1} = (W_t - C_t)\, \theta_t' R_{t+1}. \tag{46}$$

We assume log consumption growth $\Delta c_{t+1} \equiv \ln\left(\frac{C_{t+1}}{C_t}\right)$ and risk aversion parameter follow AR(1) processes and that their shocks are uncorrelated,

$$\begin{aligned}
\Delta c_{t+1} &= \mu_c + \sigma_c \varepsilon^c_{t+1} && (47)\\
\gamma_{t+1} &= \mu_\gamma + \rho\gamma_t + \sigma_\gamma \varepsilon^\gamma_{t+1} && (48)
\end{aligned}$$

C.1 Solution

Market clearing requires $C_t = D_t$, $\theta_t = 1_n$, and $W_t = P_t$.

Since preferences are homogeneous of degree one in wealth, we define

$$J(W_t, X_t) = \phi(X_t) W_t \equiv \phi_t W_t \tag{49}$$

Taking the log of Eq. 45 and evaluating first-order conditions for consumption yields:

$$C_t = (1-\delta) W_t.$$

A unit elasticity of substitution therefore implies that consumption is proportional to wealth, i.e. that agents possess a form of (rational) myopia in consumption and savings decisions. The optimal portfolio choice is fully dynamic though unless risk aversion is also unity (Giovannini and Weil, 1989).

Substitute Eq. 49 into Eq. 45:

$$\phi_t = (1-\delta)^{1-\delta} \delta^\delta \left(\mathbb{E}_t \left[\phi^{\alpha_t}_{t+1} \left(R^M_{t+1}\right)^{\alpha_t}\right]\right)^{\frac{\delta}{\alpha_t}} \tag{50}$$

or

$$\mathbb{E}_t \left[B_t \phi^{\alpha_t}_{t+1} \left(R^M_{t+1}\right)^{\alpha_t}\right] = 1, \tag{51}$$

where $B_t = \left(\frac{(1-\delta)^{1-\delta}\delta^\delta}{\phi_t}\right)^{\frac{\alpha_t}{\delta}}$.

Take logs and guess that $\ln \phi_t = a_0 + a_1 \gamma_t$:

$$\begin{aligned}
a_0 + a_1\gamma_t &= (1-\delta)\ln(1-\delta) + \delta\left(a_0 + a_1\left[\mu_\gamma + \rho\gamma_t\right]\right) && (52)\\
&\quad + \delta\mu_c + \frac{1}{2}\delta(1-\gamma_t)\left(a_1^2\sigma_\gamma^2 + \sigma_M^2\right),
\end{aligned}$$

where we used the fact that $\mathbb{E}_t r^M_{t+1} = \mu_c - \ln\delta$ (follows from Eq. 46). It follows that

$$a_1 = -\frac{1}{2}\delta \frac{a_1^2\sigma_\gamma^2 + \sigma_M^2}{1-\delta\rho} < 0. \tag{53}$$

Portfolio problem of an investor is given by

$$\max_{\theta_t} \mathbb{E}_t \left[\phi^{\alpha_t}_{t+1} \left(\theta_t' R_{t+1}\right)^{\alpha_t}\right]$$

subject to $\theta' 1_n = 1$. First-order conditions:

$$\mathbb{E}_t \left[\phi_{t+1}^{\alpha_t} \left(R_{t+1}^M \right)^{\alpha_t - 1} \left(R_{t+1}^i - R_{t+1}^1 \right) \right] = 0,$$

where R_{t+1}^1 denotes the return on any asset withing a portfolio (for instance, risk-free rate). Combining this equation with Eq. 51, we obtain the Euler equation for any asset:

$$\mathbb{E}_t \left[B_t \phi_{t+1}^{\alpha_t} \left(R_{t+1}^M \right)^{\alpha_t - 1} R_{t+1}^i \right] = 1 \tag{54}$$

Taking logs and computing conditional expectations of a log-normally distributed variable, it follows that the risk premium on any asset is given by:

$$E_t R_{t+1}^i - r_t^f = \gamma_t \text{cov}_t \left(r_{t+1}^i, r_{t+1}^M \right) + \underbrace{a_1 \left(\gamma_t - 1 \right)}_{<0 \text{ if } \gamma_t > 1} \text{cov}_t \left(r_{t+1}^i, \gamma_{t+1} \right) \tag{55}$$

Specializing this to the market return we get the expression for the market risk premium

$$\lambda_t \equiv E_t R_{t+1}^M - r_t^f = \gamma_t \sigma_M^2$$

and can therefore express Eq. 55 in the form of Eq. 6:

$$E_t R_{t+1}^i - r_t^f = \gamma_t \text{cov}_t \left(r_{t+1}^i, r_{t+1}^M \right) + \frac{a_1}{\sigma_M^2} \left(\gamma_t - 1 \right) \text{cov}_t \left(r_{t+1}^i, \lambda_{t+1} \right). \tag{56}$$

C.2 Calibration

The pricing relation Eq. 55 is of the same form as Eq. 6 and hence conditioning down according to Theorem 2 and Theorem 1 apply. We calibrate the model at daily frequency assess plausible magnitudes of unconditional risk prices predicted by the model. We use the following parameter values (annualized):

- time preference parameter (from Bansal and Yaron, 2004) $\delta = 0.976$;
- persistence of risk aversion (estimated $\frac{D}{P}$ persistence from Sabbatucci (2015)) $\rho = 0.7$;
- volatility of market returns (from data) $\sigma_M = 16\%$
- length of the cumulative sum of market returns $T = 126$ trading days
- average risk aversion of $\bar{\gamma} = 3.5$
- unconditional volatility of risk aversion $\frac{\sigma_\gamma}{\sqrt{1-\rho^2}} = 1.25$;

Based on these parameters, expected conditional risk prices are $\mathbb{E}\delta_t^M = \bar{\gamma} = 3.5$ and $\mathbb{E}\delta_t^\lambda = \frac{a_1}{\sigma_M^2} \left(\mathbb{E}\gamma_t - 1 \right) = -875$, where a_1 is given by Eq. 53. To find unconditional prices of risk, we rely on

Eq. 11. As a result, we obtain $\hat{\delta}_M = 3.5$ and $\tilde{\delta}^{\hat{\lambda}} = -7.5$, similar to the estimates in Section 3.1. The results are quite *insensitive* to the choice of σ_γ, and $\tilde{\delta}^{\hat{\lambda}}$ increases (in absolute value) with ρ. For example, $\rho = 0.75$ gives $\tilde{\delta}^{\hat{\lambda}} = -9.3$ and $\rho = 0.8$ yields $\tilde{\delta}^{\hat{\lambda}} = -12.3$. The calibration implies std $\left[\mathbb{E}_t\left(r^M_{t+1}\right)\right] \approx 3.3\%$ (annualized), consistent with the (implied) estimate of $\approx 3.4\%$ from Sabbatucci (2015).

D Robustness

We present additional results showing the sensitivity of our results to changes in specification (or lack thereof).

D.1 Monthly Estimation

Table 9 presents risk price estimates using monthly returns on our primary test assets. The estimated parameters and model fit are very similar to the daily results in Table 2. The 4-factor FF model fit has improved to nearly perfect, but the estimated risk prices (δ_{smb}, δ_{hml}, δ_{umd}) are half of the corresponding values in Table 2. In an i.i.d serially uncorrelated model, the SDF coefficients should be identical no matter what the frequency of observation[27]. This result suggests the 4-factor model is overfit, and hence, the estimates are not consistent across frequency.

D.2 Future Market Return Horizon

Table 11 shows estimated δ_λ and cross-sectional R^2 using alternative horizons, T, to define $\widehat{\lambda} = \sum_{j=2}^{T} rx^M_{t+j}$ ranging from six months to two years (using daily return data). All estimates restrict the zero-beta rate. δ_λ declines almost monotonically with T, which is expected since cov $\left(rx^i_{t+1}, \hat{\lambda}_{t+1:t+T}\right)$ increases with T and hence δ_λ must decline. Cross-sectional R^2 are fairly stable across horizon, with a peak at one year. Table 12 shows the results of repeating the exercise using monthly returns. The point estimates and patterns are similar, confirming that our results aren't driven by the choice of horizon for future market returns.

D.3 Fama-French 25

Our main results are presented using value-weighted quintile portfolios sorted on ME, BE/ME, and Prior2-12. Table 13 gives estimates using daily returns on the Fama-French 25 portfolios sorted on ME and BE/ME. As before, the 4-factor model has better fit than our 2-factor model but at the expense of less stable estimates across horizons and test assets.

[27] Ignoring error from log-linearization

Table 9: Risk Price Estimates (Monthly Returns)

This table shows risk prices estimated using monthly returns from August 1966 to December 2013 for the two-factor model, the CAPM, and the augmented Fama-French model. The test assets are value-weighted quintile portfolios sorted on ME, BE/ME, and Prior2-12. α is annualized and "-" indicates that the intercept is restricted to zero. MAPE is average absolute pricing error, annualized. Moving block bootstrap t-statistics are in parentheses.

	α	δ^m	δ^λ	δ^{smb}	δ^{hml}	δ^{umd}	R^2	$MAPE$
2-Factor model	-	2.94	-9.27	-	-	-	78.1	0.974
		(8.3)	(-7.2)					
	1.24	2.46	-9.25	-	-	-	79.5	0.991
	(3.5)	(6.4)	(-7.2)					
CAPM	-	2.67	-	-	-	-	31	1.53
		(7.7)						
	1.35	2.15	-	-	-	-	32.7	1.61
	(3.9)	(5.8)						
4-Factor FFC	-	3.98	-	1.33	6.49	4.2	94.5	0.45
		(9.5)		(1.3)	(7.7)	(11.1)		
	0.111	3.92	-	1.35	6.46	4.19	94.5	0.457
	(1.6)	(9.5)		(1.3)	(7.7)	(11.1)		

Table 10: Risk Price Estimates: Anomalies (Monthly Returns)

This table shows risk prices estimated using monthly returns from August 1966 to December 2013 for the two-factor model, the CAPM, and the augmented Fama-French model. The test assets are value-weighted anomaly portfolios. α is annualized and "-" indicates that the intercept is restricted to zero. MAPE is average absolute pricing error, annualized. Moving block bootstrap t-statistics are in parentheses.

	α	δ^m	δ^λ	δ^{smb}	δ^{hml}	δ^{umd}	R^2	$MAPE$
2-Factor model	-	3.53	-14.4	-	-	-	58.4	2.51
		(9.2)	(-13.8)					
	1.94	2.56	-11.5	-	-	-	63.6	2.45
	(4.0)	(6.4)	(-11.2)					
CAPM	-	2.97	-	-	-	-	-232	7.33
		(7.7)						
	7.28	-0.224	-	-	-	-	5.34	3.69
	(13.2)	(-0.5)						
4-Factor FFC	-	4.31	-	1.98	8.81	5.52	27.4	2.95
		(9.1)		(1.9)	(9.7)	(12.3)		
	4.42	1.93	-	0.186	4.48	3.81	70.1	1.84
	(8.8)	(3.7)		(0.2)	(4.5)	(8.5)		

Table 11: Alternative Horizons (Daily Returns)

Estimated risk price of discount rate factor and cross-sectional R^2 for alternative choices of moving average horizon. Data are daily returns with value-weighted quintile portfolios sorted on ME, BE/ME, and Prior2-12 as test assets. Moving block bootstrap t-statistics are in parentheses.

	6m	9m	12m	15m	18m	21m	24m
δ_λ	-8.7	-6.9	-7.5	-6.2	-6.9	-4.5	-4.1
	(-3.8)	(-3.5)	(-4.2)	(-3.8)	(-4.2)	(-3.7)	(-3.7)
R^2	84	79	90	83	84	80	82

Table 12: Alternative Horizons (Monthly Returns)

Estimated risk price of discount rate factor and cross-sectional R^2 for alternative choices of moving average horizon. Data are monthly returns with value-weighted quintile portfolios sorted on ME, BE/ME, and Prior2-12 as test assets. Moving block bootstrap t-statistics are in parentheses.

	6m	9m	12m	15m	18m	21m	24m
δ_λ	-9.3	-7.3	-7.3	-6.7	-7	-5	-5.1
	(-7.2)	(-6.5)	(-7.2)	(-7.0)	(-7.5)	(-7.0)	(-7.1)
R^2	78	70	72	70	70	72	73

Table 13: Risk Price Estimates (FF 25 portfolios sorted on ME and BE/ME)

This table shows premia estimated using monthly returns from 01-Aug-1966 to 31-Dec-2013 for the two-factor model, the CAPM, and the augmented Fama-French model. The test assets are the 25 portfolios sorted on ME and BE/ME. α is annualized and "-" indicates that the intercept is restricted to zero. MAPE is average absolute pricing error, annualized. Moving block bootstrap t-statistics are in parentheses.

	α	δ^m	δ^λ	δ^{smb}	δ^{hml}	δ^{umd}	R^2	$MAPE$
2-Factor model	-	3.28	-8.59	-	-	-	73.1	1.09
		(3.1)	(-3.4)					
	1.79	2.55	-8.13	-	-	-	75	1.1
	(2.0)	(2.6)	(-3.4)					
CAPM	-	3.34	-	-	-	-	-9.08	2.38
		(3.2)						
	5.22	1.2	-	-	-	-	9.38	2.18
	(3.3)	(1.2)						
4-Factor FFC	-	6.01	-	4.41	14.8	14.6	77.2	1.03
		(3.7)		(1.7)	(3.8)	(1.9)		
	1.09	5.39	-	4.27	14.1	13.5	77.9	1.04
	(2.8)	(3.5)		(1.6)	(3.7)	(1.8)		