

How to Discount Cashflows with Time-Varying Expected Returns*

Andrew Ang[†]

Columbia University and NBER

Jun Liu[‡]

UCLA

First Version: 9 May, 2002

This Version: 8 October, 2003

*The authors would like to thank Michael Brandt, Michael Brennan, Bob Dittmar, John Graham, Bruce Grundy, Ravi Jagannathan, and seminar participants at the Australian Graduate School of Management, Columbia University, the Board of Governors of the Federal Reserve and Melbourne Business School for comments. We thank Geert Bekaert and Zhenyu Wang for helpful suggestions and especially thank Yuhang Xing for constructing some of the data. We thank Rick Green (the editor) and we are grateful to an anonymous referee for helpful comments that greatly improved the paper. The authors acknowledge funding from an INQUIRE UK grant. This article represents the views of the authors and not of INQUIRE.

[†]Columbia Business School, 3022 Broadway 805 Uris, New York NY 10027. Email: aa610@columbia.edu, ph: (212) 854-9154, WWW: <http://www.columbia.edu/~aa610>.

[‡]C509 Anderson School, UCLA CA 90095. Email: jliu@anderson.ucla.edu, ph: (310) 825-4083, WWW: <http://www.personal.anderson.ucla.edu/jun.liu/>

Abstract

While many studies document that the market risk premium is predictable and that betas are not constant, the dividend discount model ignores time-varying risk premiums and betas. We develop a model to consistently value cashflows with changing risk-free rates, predictable risk premiums and conditional betas in the context of a conditional CAPM. Practical valuation is accomplished with an analytic term structure of discount rates, with different discount rates applied to expected cashflows at different horizons. Using constant discount rates can produce large mis-valuations, which, in portfolio data, are mostly driven at short horizons by market risk premiums and at long horizons by time-variation in risk-free rates and factor loadings.

To determine an appropriate discount rate for valuing cashflows, a manager is confronted by three major problems: the market risk premium must be estimated, an appropriate risk-free rate must be chosen and the beta of the project or company must be determined. All three of these inputs into a standard CAPM are not constant. Furthermore, cashflows may co-vary with the risk premium, betas or other predictive state variables. A standard Dividend Discount Model (DDM) cannot handle dynamic betas, risk premiums or risk-free rates because in this valuation method, future expected cashflows are valued at constant discount rates.

In this paper, we present an analytical methodology for valuing stochastic cashflows that are correlated with risk premiums, risk-free rates and time-varying betas. All these effects are important. First, the market risk premium is not constant. Fama and French (2002) argue that the risk premium moved to around 2% at the turn of the century from 7-8% twenty years earlier. Jagannathan, McGratten and Scherbina (2001) also argue that the market ex-ante risk premium is time-varying and fell during the late 1990's. Furthermore, a large literature claims that a number of predictor variables, including dividend yields (Campbell and Shiller (1988a and b)), risk-free rates (Fama and Schwert (1977)), term spreads (Campbell (1987)), default spreads (Keim and Stambaugh (1986)), and consumption-asset-labor deviations (Lettau and Ludvigson (2001)), have forecasting power for market excess returns.

Second, the CAPM assumes that the riskless rate is the appropriate 1-period, or instantaneous, riskless rate, which in practice is typically proxied by a 1-month or a 3-month T-bill return. However, it is highly unlikely that over the long horizons of many corporate capital budgeting problems that the riskless rate remains constant. Since the total expected return comprises both a risk-free rate and a risk premium, adjusted by a factor loading, time-varying risk-free rates imply that total expected returns also change through time. Note that even an investor who believes that the expected market excess return is constant, and a project's beta is constant, still faces stochastic total expected returns as short rates move over time.

Finally, as companies grow, merge or invest in new projects, their risk profiles change. It is quite feasible that a company's beta changes even in short intervals, and is very likely to change over 10 or 20 year horizons. There is substantial variation in factor loadings even for portfolios of stocks, for example industry portfolios (Fama and French (1997)) and portfolios sorted by size and book-to-market (Ferson and Harvey (1999)). The popularity of multifactor models for computing unconditional expected returns (for example, Fama and French (1993)) may reflect time-varying betas and conditional market risk premiums in a conditional CAPM (see Jagannathan and Wang (1996)).

This paper presents, to our knowledge, the first analytic, tractable method of discounting cashflows that embeds the effects of changing market risk premiums, risk-free rates and time-varying betas. Previous practice adjusts the DDM by using different regimes of cashflow growth or expected returns (see Lee, Myers and Swaminathan (1999) for a recent example). These adjustments are not made in an overall framework and so are subject to Fama (1996)'s critique of ad-hoc adjustments to cashflows with changing expected returns. In contrast, our valuation is done in an internally consistent framework.

Our valuation framework significantly extends the current set of analytic present value models developed in the affine class (see, among others, Ang and Liu (2001), Bakshi and Chen (2001) and Bekaert and Grenadier (2001)). If a security's beta is constant and the market risk premium is time-varying, then the price of the security would fall into this affine framework. Similarly, the case of a time-varying beta and a constant market risk premium can also be handled by an affine model. However, unlike our set-up, the extant class of models cannot simultaneously model time-variation in both beta and the market risk premium. This is because the expected return involves a product of two stochastic, predictable variables (beta multiplied by the market premium).

We derive our valuation formula under a very rich set of conditional expected returns. Our functional form for time-varying expected returns nests the specifications of the conditional CAPM developed by Harvey (1989), Ferson and Harvey (1991, 1993 and 1999), Cochrane (1996), Jagannathan and Wang (1996), among others. These studies use instrumental variables to model the time-variation of betas or market risk premiums. In our framework, short rates also vary through time. The set-up also incorporates correlation between stochastic cashflows, betas and risk premiums.

To adapt our valuation framework to current practice in capital budgeting, we compute a term structure of discount rates applied to random cashflows. Practical cashflow valuation separates the problem into two steps: first, estimate the expected future cashflows of a project or security, and then take their present value, usually by applying a constant discount rate. Instead of applying a constant discount rate, we compute a series of discount rates, or spot expected returns, which can be applied to a series of expected cashflows. The model incorporates the effects of changing market risk premiums, risk-free rates and time-varying betas by specifying a different discount rate for each different maturity.

Brennan (1997) also considers the problem of discounting cashflows with time-varying expected returns and proposes a term structure of discount rates. Our model significantly gen-

eralizes Brennan's formulation. In his set-up, the beta of the security is constant and only the risk premium changes. Furthermore, his discount rates can only be computed by simulation and were not applied to valuing predictable cashflows. In contrast, our discount rates are tractable, analytic functions of a few state variables known at each point in time. We use this analytic form to attribute the mis-pricing effects of time-varying discount rates.

We illustrate a practical application of our theoretical framework by working with cashflows and expected returns of portfolios sorted by book-to-market ratios and industry portfolios. First, we compute the term structure of discount rates at the end of our sample, December 2000, for each portfolio. At this point in time, the term structure of discount rates is upward sloping and much lower than a constant discount rate computed from the CAPM. Second, we compute the potential mis-pricing of ignoring the time-variation of expected returns. To focus on the effects of time-varying discount rates, we compute the value of a perpetuity of an expected cashflow of \$1 received each year using the term structure of discount rates from each portfolio. Ignoring time-varying expected returns can induce large potential mis-valuations; mis-pricings of over 50% using a traditional DDM are observed.

To determine the source of the mispricings, we use our model to decompose the variance of the spot expected returns into variation due to each of the separate components betas, risk-free rates and the risk premium. We find that most of the variation is driven by changes in beta and risk-free rates at long horizons, while it is most important to take into account of the variation of the risk premium at short horizons.

The rest of this paper is organized as follows. Section I presents a model for valuing stochastic cashflows with time-varying expected returns. In Section II we show how to compute the term structure of discount rates corresponding to our valuation model and derive variance decompositions for the discount rates. We apply the model to data, which we describe in Section III. The empirical results are discussed in Section IV. Section V concludes.

I. Valuing Cashflows with Time-Varying Expected Returns

In this section, our contribution is to develop a closed-form methodology for computing spot discount rates in a system which allows for time-varying cashflow growth rates, betas, short rates and market risk premiums. We begin with the standard definition of a security's expected return.

An asset pricing model specifies the expected return of a security, where the log expected

return μ_t is defined as:¹

$$\exp(\mu_t) = E_t \left[\frac{P_{t+1} + D_{t+1}}{P_t} \right], \quad (1)$$

where P_t is the price and D_t is the cashflow of the security. If, in addition, the cashflow process D_t is also specified, then the price P_t of the security can be written as:

$$P_t = E_t \left[\sum_{s=1}^{\infty} \left(\prod_{k=0}^{s-1} \exp(-\mu_{t+k}) \right) D_{t+s} \right]. \quad (2)$$

Equation (2) can be derived by iterating equation (1) and assuming transversality.

A traditional Gordon-model formula assumes that the expected return is constant, $\mu_t = \bar{\mu}$, and the expected rate of cashflow growth is also constant:

$$E_t[D_t \exp(g_{t+1})] = E_t[D_{t+1}] = D_t \exp(\bar{g}).$$

In this case, the cashflow effects and the discounting effects can be separated:

$$P_t = \sum_{s=1}^{\infty} \frac{E_t[D_{t+s}]}{\exp(s\bar{\mu})}. \quad (3)$$

This reduces equation (2) to:

$$\begin{aligned} \frac{P_t}{D_t} &= \sum_{j=1}^{\infty} \exp(-j \cdot (\bar{\mu} - \bar{g})) \\ &= \frac{1}{\exp(\bar{\mu} - \bar{g}) - 1}, \end{aligned}$$

which is the DDM formula, expressed with continuously compounded returns and growth rates.

However, as many empirical and theoretical studies suggest, expected returns and cashflow growth rates are time-varying and correlated. When this is the case, the simple discounting formula (3) does not hold. In particular, the effect of the cashflow growth rates cannot be separated from the effect of the time-varying discount rates. We must then evaluate equation (2) directly. In order to take this expectation, we specify a rich class of conditional expected returns.

Consider a conditional log expected return μ_t specified by a conditional CAPM:

$$\mu_t = \alpha + r_t + \beta_t \lambda_t, \quad (4)$$

where α is a constant, r_t is a risk-free rate, β_t is the time-varying beta and λ_t is the time-varying market risk premium. In the class of conditional CAPM's considered by Harvey (1989),

¹ In equation (1), expected returns are continuously compounded to make the mathematical exposition simpler.

Shanken (1990), Ferson and Harvey (1991 and 1993) and Cochrane (1996), among others, the time-varying beta or risk premium are parameterized by a set of instruments z_t in a linear fashion. For example, the conditional risk premium can be predicted by z_t :

$$\lambda_t \equiv E_t[y_{t+1}^m - r_t] = b_0 + b_1' z_t, \quad (5)$$

where $y_{t+1}^m - r_t$ is the log excess return on the market portfolio. Similarly, the conditional beta can be predicted by z_t and past betas:

$$E_t[\beta_{t+1}] = c_0 + c_1' z_t + c_2 \beta_t. \quad (6)$$

The instrumental variables z_t may be any variables which predict cashflows, betas or aggregate returns. For example, Harvey (1989) specifies expected returns of securities to be a linear function of market returns, dividend yields and interest rates. Jagannathan and Wang (1996) allow for conditional expected market returns to be a function of labor and interest rates. Ferson and Harvey (1991 and 1993) allow both time-varying betas and market risk premiums to be linearly predicted by factors like inflation, interest rates and GDP growth, while Ferson and Korajzyck (1995) allow time-varying betas in an APT model. In Cochrane (1996), betas can be considered to be a linear function of several instrumental variables, which also serve as the conditioning information set.

To take the expectation (2), we need to know the evolution of the instruments z_t , the betas β_t and the cashflows of the security g_t , where $g_{t+1} = \ln(D_{t+1}/D_t)$. Suppose we can summarize these variables by a $K \times 1$ state-vector X_t , where $X_t = (g_t \beta_t z_t')'$. The first and second elements of X_t are cashflow growth and the beta of the asset, respectively, but this ordering is solely for convenience. Suppose that X_t follows a VAR(1):

$$X_t = c + \Phi X_{t-1} + \Sigma^{\frac{1}{2}} \epsilon_t, \quad (7)$$

where $\epsilon_t \sim \text{IID } N(0, I)$. The one-order lag specification of this process is not restrictive, as additional lags may be added by re-writing the VAR into a companion form. Note that the instrumental variables z_t can predict betas, as well as market risk premiums, through the companion form Φ in (7).

The following proposition shows how to compute the price of the security (2) in closed form:

Proposition I.1 *Let $X_t = (g_t \beta_t z_t')'$, with dimensions $K \times 1$, follow the process in equation (7). Suppose the log expected return (1) takes the form:*

$$\mu_t = \alpha + \xi' X_t + X_t' \Omega X_t, \quad (8)$$

where α is a constant, ξ is a $K \times 1$ vector and Ω is a symmetric $K \times K$ matrix. Then, assuming existence, the price of the security is given by:

$$\begin{aligned} P_t &= E_t \left[\sum_{s=1}^{\infty} \left(\prod_{k=0}^{s-1} \exp(-\mu_{t+k}) \right) D_{t+s} \right] \\ \frac{P_t}{D_t} &= \sum_{n=1}^{\infty} \exp(a(n) + b(n)'X_t + X_t'H(n)'X_t), \end{aligned} \quad (9)$$

where the coefficients $a(n)$ is a scalar, $b(n)$ is a $K \times 1$ vector and $H(n)$ is a $K \times K$ symmetric matrix. The coefficients $a(n)$, $b(n)$ and $H(n)$ are given by the recursions:

$$\begin{aligned} a(n+1) &= a(n) - \alpha + (e_1 + b(n))'c + c'H(n)c - \frac{1}{2} \ln \det(I - 2\Sigma H(n)) \\ &\quad + \frac{1}{2} (e_1 + b(n) + 2H(n)c)'(\Sigma^{-1} - 2H(n))^{-1} (e_1 + b(n) + 2H(n)c) \\ b(n+1) &= -\xi + \Phi'(e_1 + b(n)) + 2\Phi'H(n)c \\ &\quad + 2\Phi'H(n)(\Sigma^{-1} - 2H(n))^{-1} (e_1 + b(n) + 2H(n)c) \\ H(n+1) &= -\Omega + \Phi'H(n)\Phi + 2\Phi'H(n)(\Sigma^{-1} - 2H(n))^{-1} H(n)\Phi \end{aligned} \quad (10)$$

where e_1 represents a vector of zero's with a 1 in the 1st place and

$$\begin{aligned} a(1) &= -\alpha + e_1'c + \frac{1}{2} e_1'\Sigma e_1 \\ b(1) &= -\xi + \Phi'e_1 \\ H(1) &= -\Omega \end{aligned} \quad (11)$$

The general formulation of the expected return in equation (8) can be applied to the following special cases:

1. First, the trivial case is that $\mu_t = \bar{\mu}$ is constant, so $\xi = \Omega = 0$, $\alpha > 0$, giving the standard DDM in equation (3).
2. Second, equation (8) nests a conditional CAPM relation with time-varying betas and short rates by specifying $z_t = r_t$, the short rate, so $X_t = (g_t \beta_t r_t)'$. The one-period expected return follows:

$$\mu_t = \alpha + r_t + \beta_t \bar{\lambda} = \alpha + (e_3 + \bar{\lambda} e_2)'X_t, \quad (12)$$

where $\bar{\lambda}$ is the constant market risk premium and e_i represents a vector of zeros with a 1 in the i th place. Hence, we can set $\xi = (e_3 + \lambda e_2)$ and $\Omega = 0$.

3. Third, if the market risk premium is predictable but the security or project's beta is constant ($\beta_t = \bar{\beta}$), then we can specify $X_t = (g_t r_t z_t)'$, where z_t are predictive instruments forecasting the market risk premium:

$$\lambda_t \equiv E_t[y_{t+1}^m - r_t] = b_0 + b_1' z_t.$$

The expected return then becomes:

$$\mu_t = \alpha + r_t + \bar{\beta} \lambda_t = \alpha + (e_2 + \bar{\beta} b_1)' X_t,$$

so we can set $\xi = (e_2 + \bar{\beta} b_1)$ and $\Omega = 0$.

4. Finally, we can accommodate both time-varying betas and risk premiums. If the market risk premium $\lambda_t = b_0 + b_1' z_t$ and X_t is given by our full specification $X_t = (g_t \beta_t z_t)'$, then the conditional expected return can be written as:

$$\mu_t = \alpha + r_t + \lambda_t \beta_t = \alpha + r_t + b_0 \beta_t + \beta_t (b_1' z_t). \quad (13)$$

If r_t is included in the instrument set z_t , then equation (13) takes the form of equation (8) for appropriate choices of ξ and Ω . The quadratic term Ω is now non-zero to reflect the interaction term of $\beta_t (b_1' z_t)$.

The quadratic Gaussian structure of the discount rate μ_t in equation (8) results from modelling the interaction of stochastic betas and stochastic risk premiums. Quadratic Gaussian models have been used in the finance literature in other applications. For example, Constantinides (1992) and Ahn, Dittmar and Gallant (2002) develop quadratic Gaussian term structure models. Kim and Omberg (1996), Campbell and Viceira (1999) and Liu (1999), among others, apply quadratic Gaussian structures in portfolio allocation.

The pricing formula in equation (9) is analytic because the coefficients $a(n)$, $b(n)$ and $H(n)$ are known functions and stay constant through time. Prices move because cashflow growth or state variables affecting expected returns change in X_t . The class of affine present value models in Ang and Liu (2001), Bakshi and Chen (2001) and Bekaert and Grenadier (2001) only have the scalar and linear recursions $a(n)$ and $b(n)$. Our model has an additional recursion for a quadratic term $H(n)$. The extant class of present value models is unable to handle the interaction between betas and risk premiums. Note that the quadratic $H(n)$ term also affects the recursions of $a(n)$ and $b(n)$.²

² An alternative approach is taken by Menzly, Santos and Veronesi (2003), who price stocks in a habit economy by specifying the fraction each asset contributes to total consumption. In contrast, we specify exogenous cashflows in a way that is easily adaptable to current valuation practice.

In our analysis, we consider only a CAPM formulation with time-varying betas and time-varying market risk premiums, but Proposition I.1 is general enough to model time-varying betas for multiple factors, and time-varying risk premiums for multiple factors. This generalized setting would include linear multi-factor models, like the Fama and French (1993) three-factor model. In this case, X_t would now include time-varying betas with respect to each of the factors, and the instrumental variables z_t could predict each of the factor premiums.

In Proposition I.1, we assume that beta is an exogenous process and solve endogenously for the price of the security. Using the exogenously specified expected returns and cashflows, we can construct return series for individual assets and if the number of shares outstanding of each asset is specified, we can construct the return series of the market portfolio. We can compute the covariance of an individual stock return and the aggregate market portfolio, and hence compute the implied beta of the stock from returns. Therefore, beta is both an input to the model and an output of the model. The beta specified as an input into the VAR in equation (7) and the resulting beta from the implied returns from Proposition I.1 are not necessarily the same. To see this, our model assumes that the market return takes the following form:

$$y_{t+1}^m - r_t = \lambda_t(X_t) + \sigma_t^m(X_t)v_{t+1}^m, \quad (14)$$

where λ_t is the same market risk premium in equation (4). The continuously compounded returns of security i implied by the prices from Proposition I.1 satisfy:

$$y_{t+1}^i - r_t + \frac{1}{2}(\sigma_t^i(X_t))^2 = \beta_t^i(y_{t+1}^m - r_t) + \sigma_t^i(X_t)u_{t+1}^i, \quad (15)$$

where $\frac{1}{2}(\sigma_t^i(X_t))^2$ is the Jensen's term from working in continuously compounded returns, $y_{t+1}^i - r_t$ is the excess return for asset i , and $\sigma_t^i(X_t)$ is the idiosyncratic volatility of asset i that depends on state variables.³

We obtain returns in equation (15) using the relation $y_{t+1} = (1 + P_{t+1}/D_{t+1})/(P_t/D_t) \times \exp(g_{t+1})$. Heteroskedasticity in returns arises from the non-linear form of equation (9), even though the driving process for X_t in equation (7) is homoskedastic. The beta β_t^i specified in the VAR in equation (7) is not the same as $\text{cov}_t(y_{t+1}^i, y_{t+1}^m)/(\sigma_t^m)^2$ in equation (15). If we also aggregate the returns of individual stocks by multiplying equation (15) by the market weights ω_i

³ Equations (14) and (15) represent an arbitrage-free specification, since there is a strictly positive pricing kernel m_{t+1} that supports these returns:

$$m_{t+1} = R_t^{-1} \exp\left(-\frac{1}{2} \frac{\lambda_t^2}{(\sigma_t^m)^2} - \frac{\lambda_t}{\sigma_t^m} v_{t+1}^m\right),$$

where R_t is the gross risk-free rate $R_t = \exp(r_t)$.

of each asset i , we do not obtain equation (14). This is because of the heteroskedastic Jensen's term $\frac{1}{2}(\sigma_t^i(X_t))^2$ introduced by the stock valuation equation (9). However, we would expect the discrepancy to be small, because $\sigma_t^i(X_t)^2$ in (15) is small.

The model's implied beta from returns can be made the same as the model's beta in the VAR in three ways. First, we can simply ignore the small Jensen's term in equation (15). Second, we can perform a Campbell and Shiller (1988b) log-linearization on the returns implied from Proposition I.1, equation (9), and then re-write equation (15) using log-linearized returns. Both of these approximations imply that an asset's returns satisfy a conditional version of an APT model, where

$$\sum_i \omega_i \beta_t^i = 1 \quad \text{and} \quad \sum_i \omega_i \sigma^i u_{t+1}^i = 0.$$

The second relation is the standard assumption of a factor or APT model. That is, as the number of assets becomes large, diversification causes idiosyncratic risk to tend to zero.

Finally, we can change the model specification. Proposition I.1 specifies the log discount rate to be a quadratic Gaussian process. This ensures that the discount rate is always positive. Instead, we could work in simple returns, following the conditional CAPM specified by Ferson and Harvey (1993 and 1999). If we specify the simple discount rate to be a quadratic Gaussian process, then equation (9) would become the sum of quadratic Gaussian multiplied by exponential quadratic Gaussian terms, extending Ang and Liu (2001). Then, the implied simple returns would satisfy equation (15) without the Jensen's term and the model's beta used as an input into the VAR would be consistent with the implied model beta from returns. However, this has the disadvantage of allowing negative discount rates and does not allow a term structure of discount rates for valuation to be easily computed (below).

A final comment is that, like any present value or term structure model, Proposition I.1 has an implied stochastic singularity. By exogenously specifying a beta, risk premium and risk-free rate, we specify an expected return. Combined with the cashflow process, this implies a market valuation that may not equal the observed market price of the stock.

II. The Term Structure of Expected Returns

Current practical capital budgeting is a two-step procedure. First, managers compute expected future cashflows $E_t[D_{t+s}]$ from projections, analysts' forecasts, or from extrapolation of historical data. A constant discount rate is computed, usually using the CAPM (see Graham and Harvey (2001)). The second step is to discount expected cashflows using this discount rate. The

DDM allows this separation of cashflows and discount rates only because expected returns are assumed to be constant.

Although Proposition I.1 allows us to value stochastic cashflows with time-varying returns, it is hard to directly apply the proposition to practical situations where the expected cashflow stream is separately estimated. To adapt current practice to allow for time-varying expected returns, we maintain the separation of the problem of estimating future cashflows and discounting the cashflows. However, we change the second part of the DDM valuation method. In particular, instead of a constant discount rate, we apply a series of discount rates to the expected future cashflows, where each expected future cashflow is discounted at the discount rate appropriate to the maturity of the cashflow.

This series of discount rates is computed to specifically take into account the time-variation of expected returns. That is, we specify a series of discount rates $\mu_t(n)$ for horizon n where:

$$P_t = E_t \left[\sum_{s=1}^{\infty} \left(\prod_{k=0}^{s-1} \exp(-\mu_{t+k}) \right) D_{t+s} \right] = \sum_{s=1}^{\infty} \frac{E_t[D_{t+s}]}{\exp(s \cdot \mu_t(s))}. \quad (16)$$

Each different expected cashflow at time $t+n$, $E_t(D_{t+n})$, is discounted back at its own expected return $\mu_t(n)$, as illustrated in Figure 1.

To show how the term structure of discount rates $\mu_t(s)$ can incorporate the effects of time-varying conditional expected returns, we introduce the following definition:

Definition II.1 A “spot expected return” or “spot discount rate” $\mu_t(n)$ is a discount rate which applies between time t and $t+n$ and is determined at time t . The spot expected return is the value $\mu_t(n)$ which solves:

$$E_t \left[\left(\prod_{k=0}^{n-1} \exp(-\mu_{t+k}) \right) D_{t+n} \right] = \frac{E_t[D_{t+n}]}{\exp(n \cdot \mu_t(n))}. \quad (17)$$

The series $\{\mu_t(n)\}$ varying maturity n is the term structure of expected returns or discount rates.

In equation (17), the LHS of the equation is a single term in the pricing equation (2). Using this definition enables equation (2) to be re-written as (16).

The definition in equation (17) is a generalization of the term structure of discount rates in Brennan (1997). Brennan restricts the time-variation in expected returns to come only from risk-free rates and market risk premiums, but ignores other sources of predictability (like time-varying betas and cashflows). The spot expected returns $\mu_t(n)$ depend on the information set

at time t , and as time progresses, the term structure of discount rates changes. Note that the one-period spot expected return $\mu_t(1)$ is just the one-period expected return applying between time t and $t + 1$, $\mu_t(1) \equiv \mu_t$.

To compute the spot expected returns $\mu_t(s)$, we use the following proposition:

Proposition II.1 *Let $X_t = (g_t \beta_t z_t')'$ follow the process in equation (7) and the one-period expected return μ_t follow equation (8). Then, assuming existence, the spot expected return $\mu_t(n)$ is given by:*

$$\mu_t(n) = A(n) + B(n)'X_t + X_t'G(n)X_t, \quad (18)$$

where $A(n)$ is a scalar, $B(n)$ is a $K \times 1$ vector and $G(n)$ is a $K \times K$ symmetric matrix. In the coefficients $A(n) = (\bar{a}(n) - a(n))/n$, $B(n) = (\bar{b}(n) - b(n))/n$ and $G(n) = -H(n)/n$, $a(n)$, $b(n)$ and $H(n)$ are given by equation (10) in Proposition I.1. The coefficients $\bar{a}(n)$ and $\bar{b}(n)$ are given by the recursions:

$$\begin{aligned} \bar{a}(n+1) &= \bar{a}(n) + e_1'c + \bar{b}(n)'c + \frac{1}{2}(e_1 + \bar{b}(n))'\Sigma(e_1 + \bar{b}(n)) \\ \bar{b}(n+1) &= \Phi'(e_1 + \bar{b}(n)) \end{aligned} \quad (19)$$

where e_1 represents a vector of zero's with a 1 in the 1st place and

$$\begin{aligned} \bar{a}(1) &= e_1'c + \frac{1}{2}e_1'\Sigma e_1 \\ \bar{b}(1) &= \Phi'e_1. \end{aligned} \quad (20)$$

Note that $\mu_t(n)$ is a quadratic function of X_t , the information set at time t . This is because the price of the security or asset is a function of exponential quadratic terms of X_t in equation (9). As X_t changes through time, so do the spot expected returns. This reflects the conditional nature of the expected returns, which depend on the state of the economy summarized by X_t . Like the term structure of interest rates, the term structure of discount rates can take a variety of shapes, including upward sloping, downward sloping, humped and inverted shapes.

Besides being easily applied in practical situations, there are several reasons why our model's formulation of spot expected returns is useful in the context of valuing cashflows. First, we compute the term structure of expected returns by specifying models of the conditional expected return from a rich class of conditional CAPM's, used by many previous empirical studies. We can estimate the discount curve for individual firms by looking at discount curves for industries or for other groups of firms with similar characteristics (for example, stocks with high or low book-to-market ratios).

Second, direct examination of the discount rate curve gives us a quick guide to potential mis-pricings between taking or not taking into account time-varying expected returns. The greater the magnitude of the difference between the discount rates $\mu_t(n)$ and a constant discount rate $\bar{\mu}$, the greater the mis-valuation. This difference is exacerbated at early maturities, where the time value of money is large. Since the expected cashflows are the same in the numerator of each expression in equations (3) and (16), we can compare a valuation which takes into account the effects of changing expected returns to a valuation which ignores them by looking at the difference between the discount curve $\{\mu_t(n)\}$ and the constant expected return $\bar{\mu}$ used in the standard DDM.

Third, it may be no surprise that accounting for time-varying expected returns can lead to different prices from using a constant discount rate from an unconditional CAPM. What is economically more important is quantifying the effects of time-varying expected returns by looking at its underlying sources of variation. Our analytic term structure of discount rates in Proposition II.1 allows us to attribute the effect of time-varying expected returns into its different components. For example, are time-varying risk-free rates the most important source of variation of conditional expected returns, or is it more important to account for time-variation in the risk premium?

Finally, the discount curve is analogous to the term structure of zero coupon rates. In fixed income, cashflows are known and the zero coupon rates represent the present value of \$1 to be received at different maturities in the future. In equities, cashflows are stochastic (and are correlated with the time-varying expected return) and $\mu_t(n)$ represents the expected, rather than certain, return of receiving future cashflows in the future at time $t + n$. In fixed income markets, zero coupon yields are observable while in equity markets the spot discount rates are not observable. However, one can potentially obtain the term structure of expected returns from observing the prices of stock futures contracts of different maturities. For example, if a series of derivative securities were available, each derivative security representing the claim on a stock's dividend, payable only in each separate future period, the prices of these derivative securities would represent the spot discount curve. Given the lack of suitable traded derivatives, particularly on portfolios, we directly estimate the discount curves.

If a conditional CAPM is correctly specified, the constant α in equations (4) or (8) should be zero. Since the subject of this paper is to illustrate how to discount cashflows with time-varying expected returns, rather than correctly specifying an appropriate conditional CAPM, in our empirical calibration we include an α in the stock's conditional expected return. Proposition

II.1 does not require the conditional CAPM to be exactly true. Hence, we include a constant to capture any potential mis-specifications from a true conditional CAPM.

In addition to conducting a valuation incorporating all the time-varying risk-free, risk premium and beta components, we also compute discount curves relative to two more special cases. First, if an investor correctly takes into account the time-varying market risk premium but ignores the time-varying beta, this also results in a mis-valuation. We can measure this valuation by estimating a system $X_t = (g_t r_t z_t)'$ which omits the time-varying beta, and using a constant beta in the expected return $\mu_t = \alpha + r_t + \bar{\beta}\lambda_t$. The constant beta can be estimated using an unconditional CAPM. Second, an investor can correctly measure the time-varying beta, but ignore the predictability in the market risk premium. In this second system, the investor uses an expected return $\mu_t = \alpha + r_t + \beta\bar{\lambda}$, where $\bar{\lambda}$ is the unconditional mean of the market log excess return.

A. The Time-Variation in Discount Rates

To investigate the source of the time-variation in discount rates, we can compute the variance of the discount rate $\text{var}(\mu_t(n))$ using the following corollary:

Corollary II.1 *The variance of the discount rate $\text{var}(\mu_t(n))$ is given by:*

$$\text{var}(\mu_t(n)) = B(n)' \Sigma_X B(n) + 2\text{tr}((\Sigma_X G(n))^2), \quad (21)$$

where Σ_X is the unconditional covariance matrix of X_t , given by: $\Sigma_X = \text{devec}((I - \Phi \otimes \Phi)^{-1} \text{vec}(\Sigma))$.

It is possible to perform an approximate variance decomposition on (21), given by the following corollary:⁴

Corollary II.2 *The variance of $\mu_t(n)$ can be approximated by:*

$$\text{var}(\mu_t(n)) = (B(n) + 2G(n)\bar{X})' \Sigma_X (B(n) + 2G(n)\bar{X}), \quad (22)$$

ignoring the quadratic term in equation (21), where $\bar{X} = (I - \Phi)^{-1}c$ is the unconditional mean of X_t .

⁴ The variance from the higher order terms are extremely small, for our empirical values.

We can use equation (22) to attribute the variation of $\mu_t(n)$ to variation of each of the individual state variables in X_t . However, some of the sources of variation we want to examine are transformations of X_t , rather than X_t itself. For example, a variance decomposition with respect to cashflows (g_t) or betas (β_t) can be computed using equation (22) because g_t and β_t are contained in X_t . However, a direct application of equation (22) does not allow us to attribute the variation of $\mu_t(n)$ to sources of uncertainty driving the time-variation in the market risk premium λ_t , since λ_t is not included in X_t , but is a linear transformation of X_t . To accommodate variance decompositions of linear transformations of X_t , we can rewrite equation (22) using the mapping $Z_t = L^{-1}(X_t - l)$ for L a $K \times K$ matrix and l a $K \times 1$ vector:

$$\text{var}(\mu_t(n)) = (B(n) + 2G(n)\bar{X})'L\Sigma_ZL'(B(n) + 2G(n)\bar{X}), \quad (23)$$

where $\Sigma_Z = L^{-1}\Sigma_X(L')^{-1}$.

Orthogonal variance decompositions can be computed using a Cholesky, or similar, orthogonalizing transformation for Σ_X or Σ_Z . However, in our work our variance decompositions do not sum to 1. For a single variable, we count all the contributions in the variance of that variable, together with all the covariances with each of the other variables. Hence, our variance decompositions double-count the covariances, but are not subject to an arbitrary orthogonalizing transformation.

III. Empirical Specification and Data

The model presented in Section II is very general, only needing cashflows and betas to be included in a vector of state variables X_t . To illustrate the implementation of the methodology, we specify the vector X_t that we use in our empirical application in Section A. Section B describes the data and the calibration.

A. Empirical Specification

We specify X_t as $X_t = (g_t \beta_t \Delta po_t r_t cay_t \pi_t)'$, where g_t is cashflow growth, β_t is the time-varying beta, Δpo_t is the change in the payout ratio, r_t is the nominal short rate, cay_t is Lettau and Ludvigson (2001)'s deviation from trend of consumption-asset-labor fluctuations and π_t is ex-post inflation. We motivate the inclusion of these variables as follows.

First, to predict the risk premium, we use nominal short rates r_t and cay_t . To be specific, we

parameterize the market risk premium as:

$$\lambda_t = b_0 + b_r r_t + b_{cay} cay_t. \quad (24)$$

While many studies use dividend yields to predict market excess returns (see Campbell and Shiller (1988a)), we choose not to use dividend yields because this predictive relation has grown very weak during the 1990's (see Ang and Bekaert (2002) and Goyal and Welch (2003)). In contrast, Ang and Bekaert (2002) and Campbell and Yogo (2002) find that the nominal short rate has strong predictive power, at high frequencies, for excess aggregate returns. Lettau and Ludvigson (2001) demonstrate that cay_t is a significant forecaster of excess returns, at a quarterly frequency, both in-sample and out-of-sample. Both of these predictive instruments have stronger forecasting ability than the dividend yield for aggregate excess returns.

Second, to help forecast dividend cashflows g_t , we use the change in the payout ratio, which can be considered to be a measure of earnings growth in X_t . Vuolteenaho (2002) shows that variation in firm-level earnings growth accounts for a large fraction of the variation of firm-level stock returns. However, earnings growth is difficult to compute for stock portfolios with high turnover. Instead, we use the change in the payout ratio, the ratio of dividends to earnings. This is equivalent to including earnings growth, since the change in the payout ratio, together with g_t , contains equivalent information. To show this, if we denote earnings at time t as $Earn_t$, then gross earnings growth $Earn_t/Earn_{t-1}$ can be expressed as:

$$\frac{Earn_t}{Earn_{t-1}} = \left(\frac{1/po_t}{1/po_{t-1}} \right) \exp(g_t)$$

where $po_t = D_t/Earn_t$ represents the payout ratio.

Finally, since movements in nominal short rates must be due either to movements in real rates or inflation, we also include the ex-post inflation rate π_t in X_t . This has the advantage of allowing us to separately examine the effects of the nominal short rate or the real interest rate.

To map the notation of Propositions I.1 and II.1 into this set-up, we can specify the formulation of the one-period expected return in equation (8) as follows:

$$\begin{aligned} \mu_t &= \alpha + r_t + \lambda_t \beta_t \\ &= \alpha + e_4' X_t + (b_0 + b_r r_t + b_{cay} cay_t) \beta_t \\ &= \alpha + \xi' X_t + X_t' \Omega X_t, \end{aligned} \quad (25)$$

where $\xi = (e_4 + b_0 e_2)$ and Ω is given by:

$$\Omega = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & b_r/2 & b_{cay}/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & b_r/2 & 0 & 0 & 0 & 0 \\ 0 & b_{cay}/2 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

By applying Corollary II.2. we can attribute the variation of $\mu_t(n)$ to linear transformations of X_t . For example, to compute the variance decomposition of $\mu_t(n)$ to the risk premium λ_t , we can transform $X_t = (g_t \beta_t \Delta p_{o_t} r_t cay_t \pi_t)'$ to $Z_t = (g_t \beta_t \Delta p_{o_t} r_t \lambda_t \pi_t)'$ using the mapping:

$$X_t = l + LZ_t,$$

where l is a constant vector and L is a 6×6 matrix given by:

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -\frac{b_r}{b_{cay}} & \frac{1}{b_{cay}} & -\frac{b_r}{b_{cay}} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

B. Data Description and Estimation

To illustrate the effect of time-varying expected returns on valuation, we work with ten book-to-market sorted portfolios and the Fama and French (1997) definitions of industry portfolios.⁵ We focus on these portfolios because of the well-known value effect and because industry portfolios have varying exposure to various economic factors (see Ferson and Harvey (1991)). For the book-to-market portfolios, we focus on the deciles 1, 6 and 10, which we label “growth,” “neutral” and “value,” respectively. We use data from July 1965 - Dec 2000 for the book-to-market decile portfolios and from Jan 1964 - Dec 2000 for the industry portfolios. All portfolios are value-weighted.

To estimate dividend cashflow growth rates of the portfolios, we compute monthly dividends as the difference between the portfolio value-weighted returns with dividends and capital gains,

⁵ We exclude the industry portfolios Health, Miscellaneous and Utilities because of missing data.

and the value-weighted returns excluding dividends:

$$\frac{P_{t+1/12} + \bar{D}_{t+1/12}}{P_t} - \frac{P_{t+1/12}}{P_t} = \frac{\bar{D}_{t+1/12}}{P_t},$$

where the frequency 1/12 refers to monthly data. The bar superscript in the variable $\bar{D}_{t+1/12}$ denotes a monthly, as opposed to annual, dividend. To compute annual dividend growth, we sum up the dividends over the past 12 months, as is standard practice to remove seasonality (see Hodrick (1992)):

$$D_t = \sum_{i=0}^{11} \bar{D}_{t-i/12}.$$

Growth rates of cashflows are constructed taking logs: $g_t = \log(D_t/D_{t-1})$. These cashflow growth rates represent annual increases of cashflows but are measured at a monthly frequency.

To estimate time-varying betas on each portfolio, we employ the following standard procedure, dating back to at least Fama and MacBeth (1973). We run rolling 60-month regressions of the excess total return of the portfolio on a constant and the excess market risk return:

$$\bar{y}_{\tau/12} - \bar{r}_{(\tau-1)/12} = \alpha_t + \beta_t(\bar{y}_{\tau/12}^m - \bar{r}_{(\tau-1)/12}) + u_{\tau} \quad (26)$$

where all returns are continuously compounded, $\bar{y}_{\tau/12}$ is the portfolio's log total return over month τ , $\bar{r}_{(\tau-1)/12}$ is the continuously-compounded 1-month risk-free rate (the 1-month T-bill rate) from $(\tau - 1)/12$ to $\tau/12$ and $\bar{y}_{\tau/12}^m$ is the market's log total return over month τ . The regression is run at a monthly frequency from $\tau = t - 60/12$ to $\tau = t$. The time series of the estimated linear coefficients in the regression (26) is the observable time-series of the portfolio betas β_t . We compute an α in equation (4) so that the average portfolio excess return in the data is matched by this series of betas.

While this estimation procedure is standard and has been used by several authors to document time-varying betas, including recently Fama and French (1997), it is not the optimal method to estimate betas. If the VAR is correctly specified, then we should be able to infer the true, unobservable betas from the data of realized returns, and the other observable variables in X_t , in a more efficient fashion. For example, Adrian and Franzoni (2002) use a Kalman filter to estimate time-varying betas, while Ang and Chen (2002) and Jostova and Philipov (2002) employ a Gibbs sampler. However, these estimations are complex and it is not the aim of this paper to use sophisticated econometric methods to estimate betas. Rather, we focus on discounting cashflows under time-varying betas, using a simple, standard procedure for estimating betas as an illustration.

To predict the market risk premium, we estimate the coefficients in the regression implied from equation (24):

$$y_{t+1}^m - r_t = b_0 + b_r r_t + b_{cay} cay_t + \epsilon_{t+1}, \quad (27)$$

where $y_{t+1}^m - r_t$ is an annual market excess return, using a one-year ZCB risk-free rate. To form annual monthly returns, we first compute monthly log total returns on the market portfolio from month $t/12$ to $(t + 1)/12$ and then aggregate over 12 months to form annual log returns:

$$y_{t+1}^m = \sum_{i=1}^{12} \bar{y}_{t+i/12}^m.$$

We use the monthly data in Lettau and Ludvigson (2002) to construct a series of cay_t which uses data only up to time t to estimate a cointegrating vector to estimate the consumption-wealth-labor deviation from trend at time t . This avoids any look-ahead bias in the construction of cay_t (see Brennan and Xia (2002) and Hahn and Lee (2002)). All returns are continuously compounded and the regression is run at a monthly frequency, but with an annual horizon.

We estimate our VAR in equation (7) and the predictability regression of aggregate excess returns in equation (27) at an annual horizon. That is, t to $t + 1$ represents one year. Hence, we use one-year ZCB risk-free rates r_t , year-on-year log CPI inflation π_t and an annual change in the payout ratio, Δpo_t in the VAR. We define the payout ratio of year t to be the ratio of the sum of annual dividends to summed annual earnings per share, excluding extraordinary items, of the companies in the portfolio. To compute this, we use the COMPUSTAT annual file, and extract dividends and earnings of companies in the portfolio in December of year t . We exclude any companies with negative earnings.

To gain efficiency in estimating the VAR and the predictability regression, we use monthly data. Since we have annual horizons but monthly data, the residuals from each regression in the VAR and in the predictability regression have an MA(11) form induced by the use of overlapping observations. While all parameter estimates are consistent even with the overlap, the standard errors of the parameters are affected by the MA(11) terms. To account for this, we report standard errors computed using 12 Newey-West (1987) lags.

Panel A of Table 1 presents some selected summary statistics of the representative book-to-market portfolios and the average industry. The numbers in the average industry row are averages of the statistics over all industries. Dividend growth is quite volatile: 28% (19%) for growth (value) stocks, and 21% for the average industry. Payout ratios, as expected, are highest for neutral and value stocks, at approximately 42% and lowest for growth stocks, at 26%. The average change in the payout ratios is close to zero for all portfolios. The annualized portfolio

alpha we report is estimated using a monthly regression of the portfolio excess returns onto a constant α and the excess market return over the whole sample. The alphas for the book-to-market portfolios reflect the well known value spread, increasing from -2% for growth stocks to 4% for value stocks.

The betas of the portfolios display significant time-variation. The betas of growth (value) stocks have an annual volatility of 10% (17%), and the average industry beta volatility is 19%. These betas are also quite persistent, over 75% at an annual horizon. We plot the time-varying betas at a monthly frequency in Figure 2. The betas for growth stocks and value stocks have generally diverged across the sample, with the betas for growth stocks increasing and the betas for value stocks decreasing. For example, at the beginning of the 1970's value stocks have a beta of around 1.2, which decreases to just above 0.7 by the year 2000. The betas of industry portfolios (not shown), while exhibiting time-variation, appear more stationary.

The upward trend in the growth beta and downward trend in the value beta post-1965, has been emphasized by, among others, Adrian and Franzoni (2002), Ang and Chen (2002), Campbell and Vuolteenaho (2002) and Franzoni (2002). Campbell and Vuolteenaho (2000) discuss some reasons for the trends in growth and value stocks, related to changing discount rate and cashflow sensitivities. Our VAR requires stationarity of all variables, including beta, to make econometric inferences, particularly for computing variance decompositions in Corollary II.1. The stationary assumption for beta may appear to be violated from Figure 2. However, Adrian and Franzoni (2002) and Ang and Chen (2002) show that because betas are very persistent series, it is hard to differentiate a highly persistent beta series from a beta process with a unit root in small samples. This is analogous to interest rates, where unit root tests fail to reject the null of a unit root in small samples because of low power, but term structure models require the short rate to be a stationary process.

We list the estimates of the regression (27) in Panel B of Table 1. The coefficient on the interest rate is negative, so higher interest rates cause decreases in market risk premiums. This is the same sign found by many studies since Fama and Schwert (1977). However, while Ang and Bekaert (2002) and Campbell and Yogo (2002) document strong predictive power of the short rate at monthly horizons, the significance is greatly reduced at an annual horizon. Lettau and Ludvigson (2001) find that, in-sample, cay_t significantly predicts market risk premiums with a positive sign. However, without look-ahead bias at an annual horizon, the predictive power of cay_t is reduced. Nevertheless, it is the same sign found by Lettau and Ludvigson (2001).

Since the risk premium is a function of instrumental variables, it is possible to infer the variation of the risk premium from the regression coefficients b_r and b_{cay} in (27) using:

$$\sigma_\lambda = \sqrt{\zeta' \Sigma_X \zeta}, \quad (28)$$

where $\zeta = (0 \ 0 \ 0 \ b_r \ b_{cay} \ 0)'$ and Σ_X is the unconditional covariance matrix of X_t . From the estimated parameters in Panel B of Table 1, the unconditional volatility of the risk premium is 2.66% and the risk premium has an autocorrelation of 0.54.

IV. The Calibrated Term Structure of Expected Returns

In this section, we concentrate on presenting the term structure of discount rates for the growth, neutral and value portfolios. The term structure of discount rates from these portfolios are representative of the general picture of the spot expected returns from other portfolios. However, we look at mis-pricings from valuations incorporating time-varying expected returns from both book-to-market and industry portfolios.

A. VAR Estimation Results

We report some selected VAR estimation results in Table 2 for growth, neutral and value stocks. The average industry refers to a pooled estimation of the VAR across all industry portfolios. Table 2 shows that there are some significant feed-back effects from the instruments r_t , cay_t and Δpo_t to growth rates and time-varying betas. For example, for growth (value) stocks, lagged interest rates (cay_t) predict future cashflows, and for neutral stocks, interest rates and Δpo_t predict growth rates and betas. For the average industry, r_t , cay_t and π_t significantly predict dividend growth and betas.

In Table 2, while cashflows g_t are predictable, particularly by short rates and cay_t for industry portfolios, cashflows have weak forecasting ability for the variables driving conditional expected returns, β_t , r_t , cay_t . The VAR results for the ‘‘Average Industry’’ pools across all 45 industry portfolios and does not find any evidence of predictability for cashflows. Hence, we might expect the feed-back effect of cashflows on time-varying expected returns to be weak.

B. Discount Curves

Figure 3 plots the term structure of discount rates $\mu_t(n)$ for growth, neutral and value stocks. The discount curve for the full model is shown in circles. At the end of December 2000, the

term structure of discount rates is upward sloping. At December 2000, the risk-free rate and $caiy_t$ both predict low conditional expected returns for the market. This markedly lowers the short-end of the discount curve. Since the risk premium is mean-reverting, the discount rates increase with maturity and asymptote to a constant.⁶

In Figure 3, the spot discount curve for growth stocks lies below the discount curve for value stocks. However, in Figure 2, the betas of growth stocks are higher than value stocks. The discrepancy is due to two reasons. First, the constant α term in equation (4) is negative (positive) for growth (value) stocks. This reflects the well-known value effect (see, for example, Fama and French, 1993) and brings down the spot discount curve for growth stocks relative to value stocks. Second, the discount curves also incorporate the effect of cashflows on time-varying expected returns in the VAR in equation (7).

Figure 3 also super-imposes the discount curves for the three special cases. First, the term structure of discount rates for an unconditional CAPM is a horizontal line since it is constant across horizon. Second, the shape of the term structure of discount rates ignoring the time-variation in beta is similar to the shape of the full model, particularly for growth and neutral stocks. There is a faster gradient for value stocks, but the similarities may result in a relatively small degree of mis-valuation if we ignore the time-variation in beta. However, there is a large change in the shape of the term structure when we ignore time-variation in the risk premium. In this case, the discount curves are much higher because when we ignore time-variation of the risk premium, we cannot capture the low conditional expected returns of the market portfolio at December 2000. For growth and value stocks, the term structure of discount rates ignoring the time-varying risk premium take on inverse humped shapes, illustrating some of the variety of the different shapes the discount curves may assume.

C. Mis-Pricing of Cashflow Perpetuities

We use the term structure of discount rates in Figure 3 to value a perpetuity of an expected cashflow of \$1 received at the end of each year at the end of December 2000 in Table 3 for portfolios sorted by book-to-market ratios and selected industry portfolios. Table 3 also illustrates the large mis-valuations that may result by (counter-factually) assuming expected returns are constant, ignoring the fact that betas vary over time, or ignoring the time-variation in the market risk premium.

To compute the perpetuity values we set $E_t[D_{t+s}] = 1$ for each horizon s in equation (16).

⁶ As $n \rightarrow \infty$, $\mu(n) \rightarrow \bar{\mu}$, where $\bar{\mu}$ is a constant. This is proved in the Appendix.

We value this perpetuity within each book-to-market decile or industry, under our model with time-varying conditional expected returns. These perpetuities do not represent the prices of any real firm or project because they are not actual forecasted cashflows. By keeping expected cashflows constant across the portfolios, we directly illustrate the role that time-varying expected returns play without having to control for cashflow effects across industries in the numerator. However, in the denominator, the discount rates still incorporate the effects of cashflows on time-varying expected returns in the VAR.

After computing perpetuity values from our model, we compute perpetuity values from three mis-pricings relative to the true model: (i) using a constant discount rate from an unconditional CAPM, which is a traditional DDM valuation, (ii) ignoring the time-variation in β but recognizing the market risk premium is predictable and (iii) ignoring the predictability of the market risk premium but taking into account time-varying β . We report the mis-pricings as percentage errors:

$$\text{mis-pricing error} = \frac{\text{wrong} - \text{correct}}{\text{correct}}, \quad (29)$$

where “correct” is the perpetuity value from the full valuation and “wrong” is the perpetuity value from each special case.

We turn first to the results in Table 3 for the book-to-market portfolios. The perpetuity values are from the the baseline case of time-varying short rates, betas and risk premiums. There is a general pattern of high perpetuity values for growth stocks to low perpetuity values for value stocks, but the pattern is not strictly monotonic. This follows from the low (high) discount rates for growth (value) stocks in Figure 3. The perpetuity values are almost monotonic, except for the 2nd book-to-market decile. This is mostly due to the more negative alpha for the 2nd decile (-0.03) than the 1st decile (-0.02). In addition, the growth firms (decile 1) have low payout ratios. This may understate the potential predictability of discount rates by cashflows.

The second column in Table 3 reports large mis-pricing errors from applying a DDM, with a mean error of -15%. The maximum mis-pricing, in absolute terms, is -32% for the 2nd book-to-market decile portfolio. The DDM produces much higher cashflow perpetuity values because at the end of December 2000, the conditional expected returns from our model are low, while the unconditional expected return implied by the CAPM is much higher.

The case presented in the column labelled “Ignoring Beta” in Table 3 allows for time-varying expected returns, but only through the risk premium and short rate. Ignoring time-varying betas results in overall smaller mis-pricings, but at this point in time the effect of time-varying betas can still be large (for example, 12% for the 3rd book-to-market decile portfolio). The largest

effect in mis-specifying the expected return at December 2000 comes from ignoring the time-varying market return, in the last column, rather than mis-specifying the time-varying beta. Like the DDM, ignoring variation in the risk premium produces consistently higher values of the cashflow perpetuity relative to the baseline case. This is because as the level of the market is very high at December 2000, the conditional risk premium is very low. When we use the average risk premium, we ignore this effect.

The same picture is repeated for the industry portfolios, except the extreme mis-pricings are even larger. At December 2000, the discount rates for individual industries take on a similar shape to the discount rates for book-to-market portfolios in Figure 3, because of the low conditional risk premium versus the relatively high unconditional expected return. Table 3 lists the two portfolios with the two largest absolute pricing errors from the unconditional CAPM, which are the ship industry (-58%) and fabricated products (-33%), respectively. The ship industry has a low beta at December 2000 (0.63), which causes it to have a very high perpetuity value. The unconditional beta is much higher (1.06), which means that using the DDM with the unconditional CAPM results in a very incorrect valuation. On average, using an unconditional CAPM for valuation produces a mis-pricing of -17% across all industry portfolios. Like the book-to-market portfolios, ignoring the risk premium at December 2000 produces larger mis-valuations on average (-13%) than ignoring the time-variation of beta (-5%). In summary, the effect of time-varying expected returns on valuation is important.

D. Variance Decompositions

That ignoring time-varying expected returns, or some component of time-varying expected returns produces different valuations than the DDM is no surprise. What is more economically interesting is to investigate what is driving the time-variation in the discount rates. We examine this by applying Corollary II.2 to compute variance decompositions of the spot expected returns.

We first illustrate the volatility of the spot expected returns, $\sqrt{\text{var}(\mu_t(n))}$, at each maturity in the left column of Figure 4. As the maturity increases, the volatility of the discount rates tends to zero. This is because as $n \rightarrow \infty$, $\mu_t(n)$ approaches a constant because of stationarity, so $\text{var}(\mu_t(n)) \rightarrow 0$. At a 30-year horizon, the $\mu_t(30)$ discount rate still has a volatility above 2.5% for growth and neutral stocks, and above 7.0% for value stocks. While the volatility curve must eventually approach zero, it need not do so monotonically. In particular, for value stocks, there is a strong hump-shape, starting from around 4.7% at a 1 year horizon, increasing to near

8.0% at 13 years before starting to decline. The strong hump in $\sqrt{\text{var}(\mu_t(n))}$ for value stocks compared to growth and neutral stocks is due to the much larger persistence of the value betas (0.84 compared to 0.68 (0.57) for growth (neutral) stocks in the VAR estimates of Table 2). Note that the current beta is known in today's conditional expected return. A shock to the beta only takes effect next period and the more persistent the beta, the larger the contribution to the variance of the discount rate.

In the right-hand column of Figure 4, we decompose the variance of the discount rates. Our first result is that the time-variation in cashflows makes only a very small contribution to the variance of the spot expected returns. We add both the variance decomposition to g_t and the variance decomposition to Δpo_t together to determine the total variance decomposition to cashflows. The small effect of cashflows on discount rates is expected, because cashflows or payouts weakly predict the variables driving time-varying expected returns: time-varying betas, short rates, and cay_t . The persistence of cashflows is also very low (see Table 1), and so shocks to cashflows have little long-term effect on the variances of the discount factors.

Second, at very short maturities Figure 4 shows that the attribution of the variance of $\mu_t(n)$ to nominal risk-free rates is large, the attribution to the market risk premium is also large, and the attribution to beta is smaller than the variance decomposition to risk-free rates or to the market risk premium. For example, for neutral stocks, approximately 65% of $\text{var}(\mu_t)$ is accounted for by risk-free rates, 72% by the market risk premium, and 20% by time-varying beta. Hence, at short horizons, it is crucial to account for time-varying short rates and risk premiums. The effect of beta is secondary.

Some intuition for this result can be gained by more closely examining the one-period expected return:

$$\begin{aligned}
\mu_t &= r_t + \beta_t \lambda_t \\
&= (r_t + \bar{r} - \bar{r}) + (\beta_t + \bar{\beta} - \bar{\beta})(\lambda_t + \bar{\lambda} - \bar{\lambda}) \\
&= \text{const} + (r_t - \bar{r}_t) + \bar{\beta}(\lambda_t - \bar{\lambda}) + \bar{\lambda}(\beta_t - \bar{\beta}) + (\beta_t - \bar{\beta})(\lambda_t - \bar{\lambda}), \tag{30}
\end{aligned}$$

where \bar{r} , $\bar{\beta}$ and $\bar{\lambda}$ represent the unconditional means of nominal interest rates, beta and risk premiums, respectively. Ignoring the covariance and other higher-order terms in (30), we have:

$$\text{var}(\mu_t) \approx \text{var}(r_t) + \bar{\beta}^2 \text{var}(\lambda_t) + \bar{\lambda}^2 \text{var}(\beta_t). \tag{31}$$

The variance of r_t enters one for one and so has a large effect, but $\text{var}(\lambda_t)$ and $\text{var}(\beta_t)$ are scaled by the effects of $\bar{\beta}$ and $\bar{\lambda}$. Since $\bar{\beta}$ is approximately 1, the variance of the risk premium also has

a large effect. However, the average log risk premium in the data is of the order of 5%, which means that $\text{var}(\beta_t)$ has a smaller effect on the variance of μ_t than risk-free rates or market risk premia. For value stocks, the variance of betas is relatively large, allowing betas to account for up to 41% of the variance of $\mu_t(1)$, but this is still smaller than the one-period variance decompositions to risk-free rates (71%) and risk premia (72%).

Third, the variance decomposition of the risk premium decreases as n increases. While the time-variation in the market is very important for the value of short-term cashflows, we can pay less attention to the predictability of the market premium for long-term cashflows. Mathematically, the risk premium is a linear function of the instrumental variables r_t and cay_t . The autocorrelation of r_t is around 0.74 at an annual horizon, and cay_t is much less autocorrelated (0.63 at an annual horizon). The risk premium is a linear function of both r_t and cay_t , and is less autocorrelated than the short rate (0.54). This means that at long horizons, shocks to the risk premium are less persistent than shocks to the short rate and other variables in the system, leading to a reduction in the variance decomposition to the risk premium as n increases.⁷

Finally, the variance decomposition of the risk-free rate can increase or decrease with horizon, and can dominate, or be dominated by the variance of time-varying beta. For growth stocks, the attribution of $\text{var}(\mu_t(n))$ to the interest rate only slightly decreases as n increase, while for value stocks the risk-free rate variance decomposition becomes much smaller at long horizons. Hence, growth stocks are more sensitive to movements in the nominal term structure than value stocks. This is in line with intuition as growth stocks have few short-term cashflows but potentially large long-term cashflows.

The mechanism by which the nominal risk-free rate or β_t can dominate the variance decomposition of $\text{var}(\mu_t(n))$ at long horizons is due to the relative persistence of the interest rate versus beta and the size of the predictive coefficients in the risk premium. Since the interest rate is very persistent, shocks to r_t tend to dominate at long horizons unless the autocorrelation of beta is large enough, relative to the autocorrelation of real rates, to off-set its effects. The autocorrelation of beta (0.86) is much larger than the autocorrelation of the beta of growth stocks (0.76), which allows the variance attribution to β to dominate at long horizons for the value portfolio.

In Figure 5, we perform a more detailed variance decomposition of $\text{var}(\mu_t(n))$ to risk-free rates. Figure 5 repeats the variance decompositions to r_t from Figure 4 and also plots the

⁷ If dividend yields are used instead of cay_t , the variance decomposition to the risk premium falls across all horizons. While the dividend yield is more persistent than both the nominal or ex-post real risk-free rate, the predictive coefficient of the dividend yield in the risk premium regression is almost zero in our sample.

variance decompositions to actual (or ex-post) real rates $r_t - \pi_t$ and expected (or ex-ante) real rates $E_t(r_{t+1} - \pi_{t+1})$. First, the variance decompositions to nominal, actual and expected real rates all follow the same patterns in absolute magnitude. In particular, at short horizons the variance decompositions to real rates, like nominal rates, is large. The same intuition for these results for the nominal rate using the approximation in equation (31) also applies to the ex-ante or ex-post real rates.

Second, the variance decomposition to ex-ante and ex-post real rates is negative, compared to the positive variance decompositions to r_t . The reason is that while r_t is unconditionally positively correlated with the other state variables, the actual and expected real rates are negatively correlated with the other state variables. For example, for value stocks, the correlation of r_t with β_t is 56%, whereas the correlation of $r_t - \pi_t$ with β_t is -25% and the correlation of $E_t(r_{t+1} - \pi_{t+1})$ with β_t is -35%.⁸ By definition, the variance decompositions to risk-free rates, ex-ante and ex-post real rates that do not count the covariances must be positive. Hence, the negative variance decompositions result solely from the unconditional negative correlations of real rates with other state variables. Finally, the variance decompositions of actual real rates are larger than than the variance decompositions of ex-ante real rates. This is expected, as the actual real rate comprises the ex-ante real rate plus unpredictable inflation noise.

V. Conclusion

Despite the strong evidence for time-variation in the market risk premium, factor loadings and risk-free rates, the main tool of valuation, the Dividend Discount Model does not take into account any of these stylized facts. We develop a valuation methodology which incorporates time-varying risk premiums, betas and risk-free rates by computing a series of discount rates which differ across maturity. The price of a security has an analytical solution, which depends only on observable instruments.

For application to practical capital budgeting problems, we develop an analytical, tractable term structure of discount rates. This series of discount rates differs across maturity and can be applied to value a series of expected cashflows. The discount curve is constructed in such a way to consistently model the dynamics of time-varying risk-free rates, betas and risk premiums.

⁸ The fact that the actual and expected real rates are negatively correlated with inflation (at -58% and -43%, respectively), while there is a positive correlation of nominal risk-free rates and inflation (70%), is the well-known Mundell (1963) and Tobin (1965) effect.

We estimate the term structure of discount rates for book-to-market and industry portfolios, and find the effect of time-variation in risk-free rates, betas and risk premiums is large. By computing a variance decomposition of the discount rates, we show that at short horizons, investors should be most concerned with the impact of time-varying interest rates and risk premiums for discounting cashflows. At long horizons, the time-variation in risk-free rates or beta is more important.

While we provide an easily applicable methodology for handling the effects of time-varying risk premiums, risk-free rates and beta, and demonstrate that all these are important for valuation, future research must deal with some practical issues. For example, parameter uncertainty in the predictability of the market risk premium and estimating betas will affect the capital budgeting problem. Time-varying risk-free rates, betas and risk premiums can only make potential mis-pricings in these situations even larger.

Appendix

A. Proof of Proposition I.1

Before proving Proposition I.1 we first prove a useful lemma:

Lemma A.1 *Let ϵ be a $K \times 1$ vector, where $\epsilon \sim N(0, \Sigma)$, A a $K \times K$ matrix and Ω a symmetric $K \times K$ matrix. If $(\Sigma^{-1} - 2\Omega)$ is strictly positive definite, then:*

$$E[\exp(A\epsilon + \epsilon'\Omega\epsilon)] = \exp\left(-\frac{1}{2} \ln \det(I - 2\Sigma\Omega) + \frac{1}{2} A'(\Sigma^{-1} - 2\Omega)^{-1}A\right).$$

Proof:

$$\begin{aligned} E[\exp(A\epsilon + \epsilon'\Omega\epsilon)] &= \int (2\pi)^{-K} |\det \Sigma|^{-1/2} \exp\left(-\frac{1}{2} \epsilon' \Sigma^{-1} \epsilon\right) \exp(A'\epsilon + \epsilon'\Omega\epsilon) d^K \epsilon \\ &= \int (2\pi)^{-K} |\det \Sigma|^{-1/2} \exp\left(-\frac{1}{2} \epsilon' (\Sigma^{-1} - 2\Omega) \epsilon + A'\epsilon\right) d^K \epsilon \end{aligned} \quad (\text{A-1})$$

If $(\Sigma^{-1} - 2\Omega)$ is strictly positive definite then there exists a $K \times K$ lower triangular matrix M , $\det M \neq 0$, such that $(\Sigma^{-1} - 2\Omega) = (MM')^{-1}$. Substituting this into the above expression we obtain:

$$E[\exp(A\epsilon + \epsilon'\Omega\epsilon)] = \int (2\pi)^{-K} |\det \Sigma|^{-1/2} \exp\left(-\frac{1}{2} \epsilon' (M')^{-1} M^{-1} \epsilon + A'\epsilon\right) d^K \epsilon \quad (\text{A-2})$$

Now substitute $u = M^{-1}\epsilon$:

$$\begin{aligned} E[\exp(A\epsilon + \epsilon'\Omega\epsilon)] &= \int (2\pi)^{-K} |\det \Sigma|^{-1/2} |\det M| \exp\left(-\frac{1}{2} u' u + A' M u\right) d^K u \\ &= |\det M| |\det \Sigma|^{-1/2} \exp\left(\frac{1}{2} A' M M' A\right) \\ &= |\det M| |\det \Sigma|^{-1/2} \exp\left(\frac{1}{2} A' (\Sigma^{-1} - 2\Omega)^{-1} A\right) \end{aligned} \quad (\text{A-3})$$

Finally, looking at the determinant terms we have:

$$\begin{aligned} |\det M| |\det \Sigma|^{-1/2} &= \left(\frac{|\det(\Sigma^{-1} - 2\Omega)^{-1}|}{|\det \Sigma|}\right)^{(1/2)} \\ &= \left(\frac{1}{\det \Sigma(\Sigma^{-1} - 2\Omega)}\right)^{(1/2)} \\ &= \exp\left(-\frac{1}{2} \ln |\det(I - 2\Sigma\Omega)|\right) \end{aligned} \quad (\text{A-4})$$

Substituting into (A-3) we obtain the required result. ■

Constantinides (1992) states a continuous-time version of Lemma A.1, except his result is only for a univariate case.

To prove Proposition I.1, consider one cashflow at time $t + n$ in the term in equation (2) divided by D_t :

$$E_t \left[\left(\prod_{k=0}^{n-1} e^{-\mu_{t+k}} \right) D_{t+n} \right] / D_t. \quad (\text{A-5})$$

We show that:

$$\mathbb{E}_t \left[\left(\prod_{k=0}^{n-1} e^{-\mu_{t+k}} \right) \frac{D_{t+n}}{D_t} \right] = \exp(a(n) + b(n)'X_t + X_t' H(n)' X_t). \quad (\text{A-6})$$

Summing up over n gives Proposition I.1.

The initial condition is given by:

$$\begin{aligned} \frac{\mathbb{E}_t[\exp(\mu_t)D_{t+1}]}{D_t} &= \mathbb{E}_t[\exp(-\alpha - \xi'X_t - X_t'\Omega X_t + e_1'X_{t+1})] \\ &= \exp\left(-\alpha - \xi'X_t - X_t'\Omega X_t + e_1'(c + \Phi X_t) + \frac{1}{2}e_1'\Sigma e_1\right) \end{aligned} \quad (\text{A-7})$$

Equating coefficients gives equation (11).

Using induction, for arbitrary time $t + n + 1$ we can evaluate $\mathbb{E}_t \left[\left(\prod_{k=0}^n e^{-\mu_{t+k}} \right) \frac{D_{t+n+1}}{D_t} \right]$ as:

$$\begin{aligned} &= \mathbb{E}_t \{ \exp(-\alpha - \xi'X_t - X_t'\Omega X_t + e_1'X_t) \\ &\quad \times \mathbb{E}_t[\exp(a(n) + b(n)'X_{t+1} + X_{t+1}'H(n)X_{t+1})] \} \\ &= \exp(a(n) - \alpha - \xi'X_t - X_t'\Omega X_t) \mathbb{E}_t[\exp((e_1 + b(n))'X_{t+1} + X_{t+1}'H(n)X_{t+1})] \\ &= \exp(a(n) - \alpha - \xi'X_t - X_t'\Omega X_t) \\ &\quad \times \exp((e_1 + b(n))'(c + \Phi X_t) + (c + \Phi X_t)'H(n)(c + \Phi X_t)) \\ &\quad \times \mathbb{E}_t[\exp((e_1 + b(n))'\epsilon_{t+1} + 2(c + \Phi X_t)'H(n)\epsilon_{t+1} + \epsilon_{t+1}'H(n)\epsilon_{t+1})], \end{aligned} \quad (\text{A-8})$$

which involves taking the expectation of a quadratic Gaussian. This can be done using Lemma A.1 to obtain:

$$\begin{aligned} &\exp(a(n) - \alpha + (e_1 + b(n))'c + c'H(n)c \\ &\quad - \xi'X_t + (e_1 + b(n))'\Phi X_t + 2c'H(n)\Phi X_t - X_t'\Omega X_t + X_t'\Phi'H(n)\Phi X_t) \\ &\quad \times \exp\left(-\frac{1}{2} \ln \det(I - 2\Sigma H(n)) + \frac{1}{2}D'(\Sigma^{-1} - 2H(n))^{-1}D\right) \end{aligned} \quad (\text{A-9})$$

where $D = (e_1 + b(n) + 2H(n)(c + \Phi X_t))$. Expanding the expression gives equation (10). ■

Our methodology can easily be extended to allow for heteroskedasticity of an affine form in the conditional volatility, along the lines of Duffie and Kan (1996) in continuous-time or the discrete-time set-up in Ang and Liu (2001). To implement this, specify the now-constant conditional volatility Σ of the system as time-varying:

$$X_t = c + \Phi X_{t-1} + u_t, \quad (\text{A-10})$$

where $u_t \sim \text{IID } N(0, \Sigma_{t-1})$, where

$$\Sigma_t = Q_0 + Q \odot X_t \quad (\text{A-11})$$

and “ \odot ” represents a tensor product:

$$Q \odot X_t \equiv \sum_{j=1}^K X_{tj} Q^{(j)},$$

where X_{tj} refers to the j th element of the $K \times 1$ vector X_t . The $K \times K$ matrices Q_0 and $Q^{(j)}$ are symmetric. In this set-up, $\mu_t(n)$ still has the same quadratic form as equation (18), and the recursions for $a(n)$, $b(n)$ and $H(n)$ take a similar form as equation (10), except with additional terms to accommodate the heteroskedasticity.

B. Proof of Proposition II.1

To prove Proposition II.1, we determine separately the expressions of the numerator and the denominator of the fraction in the expression:

$$\mu_t(n) = \frac{1}{n} \ln \left(\frac{\mathbb{E}_t[D_{t+n}]/D_t}{\mathbb{E}_t \left[\left(\prod_{k=0}^{n-1} e^{-\mu_{t+k}} \right) D_{t+n} / D_t \right]} \right) \quad (\text{B-12})$$

in Definition II.1 of the spot expected return.

We first begin by showing that:

$$E_t[D_{t+n}]/D_t = \exp(\bar{a}(n) + \bar{b}(n)'X_t), \quad (\text{B-13})$$

where $\bar{a}(n)$ and $\bar{b}(n)$ given in equations (19) and (20). The solution method is similar to those used in discrete-time affine economies such as Ang and Liu (2001) and Bekaert and Grenadier (2001). The initial condition is given by:

$$\begin{aligned} E_t[D_{t+1}]/D_t &= E_t[\exp(g_{t+1})] = E_t[\exp(e_1'X_{t+1})] \\ &= \exp(e_1'c + \frac{1}{2}e_1'\Sigma e_1 + e_1'\Phi X_t). \end{aligned} \quad (\text{B-14})$$

Equating coefficients gives the initial conditions in equation (20). For the recursion, we use proof by induction and take iterative expectations:

$$\begin{aligned} \frac{E_t[D_{t+n+1}]}{D_t} &= E_t \left[\frac{D_{t+1}}{D_t} E_{t+1} \left[\frac{D_{t+n}}{D_t} \right] \right] \\ &= E_t[\exp(e_1'X_{t+1} + \bar{a}(n) + \bar{b}(n)'X_{t+1})] \\ &= \exp(e_1'c + e_1'\Phi X_t + \bar{a}(n) + \bar{b}(n)'c + \bar{b}(n)'\Phi X_t) E_t[\exp((e_1 + \bar{b}(n))'\Sigma^{\frac{1}{2}}\epsilon_{t+1})] \\ &= \exp \left(e_1'c + \bar{a}(n) + \bar{b}(n)'c + (e_1 + \bar{b}(n))'\Phi X_t + \frac{1}{2}(e_1 + \bar{b}(n))'\Sigma(e_1 + \bar{b}(n)) \right) \end{aligned} \quad (\text{B-15})$$

Equating coefficients gives the recursion in equation (19).

Proposition I.1 shows that the denominator of (B-12) has the form

$$E_t \left[\left(\prod_{k=0}^{n-1} e^{-\mu_{t+k}} \right) \frac{D_{t+n}}{D_t} \right] = \exp(a(n) + b(n)'X_t + X_t'H(n)'X_t), \quad (\text{B-16})$$

where $a(n)$, $b(n)$ and $H(n)$ follow the recursions in equation (10).

This allows us to write $\mu_t(n)$ in equation (B-12) as:

$$\mu_t(n) = A(n) + B(n)'X_t + X_t'G(n)X_t, \quad (\text{B-17})$$

where $A(n) = (\bar{a}(n) - a(n))/n$, $B(n) = (\bar{b}(n) - b(n))/n$ and $G(n) = -H(n)/n$. ■

C. Variance of the Discount Rate

Proof of Corollary II.1

To compute the variance of the discount rate, $\text{var}(\mu_t(n))$, in Corollary II.1 we use the following lemma:

Lemma C.1 For an arbitrary matrix M , the following relationships are true (see Harville (1997)):

$$\begin{aligned} \frac{\partial}{\partial \lambda} \ln \det M &= \text{tr} \left(M^{-1} \frac{\partial M}{\partial \lambda} \right) \\ \frac{\partial}{\partial \lambda} M^{-1} &= -M^{-1} \frac{\partial M}{\partial \lambda} M^{-1}. \end{aligned} \quad (\text{C-1})$$

Since, from Proposition II.1, $\mu_t(n) = A(n) + B(n)'X_t + X_t'G(n)X_t$ we have:

$$\text{var}(\mu_t(n)) = B(n)'\Sigma_X B(n) + \text{var}(X_t'G(n)X_t). \quad (\text{C-2})$$

We use a moment generating function to evaluate the last term. Let:

$$g(\lambda) = E[\exp(\lambda \epsilon' M \epsilon)]. \quad (\text{C-3})$$

Then:

$$\text{var}(\epsilon' M \epsilon) = \frac{\partial^2 g(\lambda)}{\partial \lambda^2} \Big|_{\lambda=0} - \left(\frac{\partial g(\lambda)}{\partial \lambda} \Big|_{\lambda=0} \right)^2. \quad (\text{C-4})$$

Using Lemma A.1, we have

$$g(\lambda) = \exp \left(-\frac{1}{2} \ln \det(I - 2\lambda \Sigma M) \right), \quad (\text{C-5})$$

where Σ is the variance matrix of ϵ . Using Lemma C.1, we have:

$$\frac{\partial}{\partial \lambda} g(\lambda) \ln \det M = g(\lambda) \text{tr} \left((I - 2\lambda \Sigma M)^{-1} \Sigma M \right), \quad (\text{C-6})$$

and

$$\begin{aligned} & \frac{\partial^2}{\partial \lambda^2} g(\lambda) \ln \det M \\ &= g(\lambda) \left\{ \left(\text{tr} (I - 2\lambda \Sigma M)^{-1} \Sigma M \right)^2 + \text{tr} \left((I - 2\lambda \Sigma M)^{-1} 2 \Sigma M (I - 2\lambda \Sigma M)^{-1} \Sigma M \right) \right\}. \end{aligned} \quad (\text{C-7})$$

Therefore, we have:

$$\text{var}(\epsilon' M \epsilon) = \text{tr} \left((\Sigma M)^2 \right), \quad (\text{C-8})$$

evaluating equation (C-5). Hence,

$$\text{var}(\mu_t(n)) = B(n)' \Sigma_X B(n) + 2 \text{tr} \left((\Sigma_X G(n))^2 \right) \quad (\text{C-9})$$

given in equation (21). ■

Proof of Corollary II.2

From the definition of $\mu_t(n) = A(n) + B(n)' X_t + X_t' G(n) X_t$ we can write:

$$\begin{aligned} \mu_t(n) &= (A(n) - B(n)' \bar{X}) + B(n)' (X_t - \bar{X}) + (X_t - \bar{X} + \bar{X})' G(n) (X_t - \bar{X} + \bar{X}) \\ &= (A(n) - B(n)' \bar{X} + \bar{X}' G(n) \bar{X}) + (B(n) + 2G(n) \bar{X})' (X_t - \bar{X}) \\ &\quad + (X_t - \bar{X})' G(n) (X_t - \bar{X}). \end{aligned} \quad (\text{C-10})$$

Ignoring the quadratic term, we have:

$$\text{var}(\mu_t(n)) = (B(n) + 2G(n) \bar{X})' \Sigma_X (B(n) + 2G(n) \bar{X}). \quad (\text{C-11})$$

Since $\Sigma_X = L \Sigma_Z L'$, we can rewrite this expression as:

$$\text{var}(\mu_t(n)) = (B(n) + 2G(n) \bar{X})' L \Sigma_Z L' (B(n) + 2G(n) \bar{X}), \quad (\text{C-12})$$

giving us equation (22). ■

D. Long-Term Discount Rates

It is possible to compute the long-term discount rate $\mu(\infty)$ using the following proposition:

Proposition D.1 *Suppose the following limits exist:*

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\bar{a}(n)}{n} &= \bar{a}_\infty, \\ \lim_{n \rightarrow \infty} \frac{a(n)}{n} &= a_\infty, \\ \lim_{n \rightarrow \infty} \bar{b}(n) &= \bar{b}_\infty, \\ \lim_{n \rightarrow \infty} b(n) &= b_\infty, \\ \lim_{n \rightarrow \infty} H(n) &= H_\infty, \end{aligned} \quad (\text{D-1})$$

then the long-term discount rate $\mu(\infty)$ is given by:

$$\mu(\infty) = (\bar{a}_\infty - a_\infty), \quad (\text{D-2})$$

where:

$$\begin{aligned} \bar{a}_\infty &= e_1'c + \bar{b}'_\infty c + \frac{1}{2}(e_1 + \bar{b}_\infty)'\Sigma(e_1 + \bar{b}_\infty) \\ a_\infty &= -\alpha + (e_1 + b_\infty)'c + c'H_\infty c - \frac{1}{2} \ln \det(I - 2\Sigma H_\infty). \end{aligned} \quad (\text{D-3})$$

Given H_∞ , \bar{b}_∞ and b_∞ are given by:

$$\begin{aligned} \bar{b}_\infty &= (I - \Phi')^{-1} \Phi' e_1 \\ b_\infty &= (I - \Phi' - 2\Phi' H_\infty (\Sigma^{-1} - 2H_\infty)^{-1})^{-1} \\ &\quad (-\xi + \Phi' e_1 + 2\Phi' H_\infty c + 2\Phi' H_\infty (\Sigma^{-1} - 2H_\infty)^{-1} (e_1 + 2H_\infty c)). \end{aligned} \quad (\text{D-4})$$

The matrix H_∞ solves the relation:

$$H_\infty = -\Omega + \Phi' H_\infty \Phi + 2\Phi' H_\infty (\Sigma^{-1} - 2H_\infty)^{-1} H_\infty \Phi. \quad (\text{D-5})$$

In practice, it is easy to solve for H_∞ by iterating the recursion for $H(n)$ in equation (10) to a very large number. After $n > 100$ years, $H(n)$ is constant. Once H_∞ is found, \bar{a}_∞ and a_∞ are closed-form.

References

- Adrian, Tobias, and Francesco Franzoni, 2002, Learning about beta: An explanation of the value premium, Working paper, MIT.
- Ahn, Dong Hyun, Robert F. Dittmar, and A. Ronald Gallant, 2002, Quadratic term structure models: Theory and evidence, *Review of Financial Studies* 15, 243-288.
- Ang, Andrew, and Geert Bekaert, 2002, Stock return predictability: Is it there? Working paper, Columbia Business School.
- Ang, Andrew, and Joseph Chen, 2002, CAPM over the long-run: 1926-2001, Working paper, Columbia Business School.
- Ang, Andrew, and Jun Liu, 2001, A general affine earnings valuation model, *Review of Accounting Studies* 6, 397-425.
- Bakshi, Gurdip, and Zhiwu Chen, 2001, Stock valuation in dynamic economies, Working paper, Yale University.
- Bekaert, Geert, and Steve Grenadier, 2001, Stock and bond pricing in an affine economy, Working paper, Columbia Business School.
- Brennan, Michael J., 1997, The term structure of discount rates, *Financial Management* 26, 81-90.
- Brennan, Michael J., and Yihong Xia, 2001, TAY's as good as CAY, Working paper, Wharton.
- Campbell, John Y., 1987, Stock returns and the term structure, *Journal of Financial Economics* 18, 373-399.
- Campbell, John Y., and Robert J. Shiller, 1988a, Stock prices, earnings and expected dividends, *Journal of Finance* 43, 661-676.
- Campbell, John Y., and Robert J. Shiller, 1988b, The dividend-price ratio and expectations of future dividends and discount factors, *Review of Financial Studies* 1, 195-228.
- Campbell, John Y., and Luis M. Viceira, 1999, Consumption and portfolio decisions when expected returns are time varying, *Quarterly Journal of Economics* 114, 433-495.
- Campbell, John Y., and Tuomo Vuolteenaho, 2002, Bad beta, good beta, Working paper, Harvard University.
- Campbell, John Y. and Motohiro Yogo, 2002, Efficient tests of stock return predictability, Working paper, Harvard University.
- Cochrane, John H., 1996, A cross-sectional test of an investment-based asset pricing model, *Journal of Political Economy* 104, 572-621.
- Constantinides, George M., 1992, A theory of the nominal term structure of interest rates, *Review of Financial Studies* 5, 531-552.
- Duffie, Darrel and Rui Kan, 1996, A yield-factor model of interest rates, *Mathematical Finance* 6, 379-406.
- Fama, Eugene F., 1996, Discounting under uncertainty, *Journal of Business* 69, 415-428.
- Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 25, 23-49.
- Fama, Eugene F., and Kenneth R. French, 1997, Industry costs of equity, *Journal of Financial Economics* 43, 153-193.
- Fama, Eugene F., and Kenneth R. French, 2002, The equity premium, *Journal of Finance* 57, 637-659.
- Fama, Eugene F., and James MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607-636.
- Fama, Eugene F., and G. William Schwert, 1977, Asset returns and inflation, *Journal of Financial Economics* 5, 115-146.
- Ferson, Wayne E., and Campbell R. Harvey, 1991, The variation of economic risk premiums, *Journal of Political Economy* 99, 385-415.
- Ferson, Wayne E., and Campbell R. Harvey, 1993, The risk and predictability of international equity returns, *Review of Financial Studies* 6, 527-566.

- Ferson, Wayne E., and Campbell R. Harvey, 1999, Conditioning variables and the cross section of stock returns, *Journal of Finance* 54, 1325-1360.
- Ferson, Wayne E., and Robert A. Korajczyk, 1995, Do arbitrage pricing models explain the predictability of stock returns? *Journal of Business* 68, 309-349.
- Franzoni, Francesco, 2002, Where is beta going? The riskiness of value and small stocks, Working paper, MIT.
- Goyal, Amit, and Ivo Welch, 2003, The myth of predictability: Does the dividend yield forecast the equity premium? *Management Science* 49, 639-654.
- Graham, John R., and Campbell R. Harvey, 2001, The theory and practice of corporate finance: Evidence from the field, *Journal of Financial Economics* 61, 1-28.
- Hahn, Jaehoon, and Hayong Lee, 2002, On the estimation of the consumption-wealth ratio: Cointegrating parameter instability and its implications for stock return forecasting, Working paper, University of Washington.
- Harvey, Campbell R., 1989, Time-varying conditional covariances in tests of asset pricing models, *Journal of Financial Economics* 24, 289-317.
- Harville, David A., 1997, *Matrix algebra from a statistician's perspective* (Springer-Verlag, New York).
- Hodrick, Robert J., 1992, Dividend yields and expected stock returns: Alternative procedures for inference and measurement, *Review of Financial Studies* 5, 357-386.
- Jagannathan, Ravi, Ellen R. McGrattan, and Anna Scherbina, 2001, The declining US equity premium, *Quarterly Review, Federal Reserve Bank of Minneapolis* 24, Fall, 3-19.
- Jagannathan, Ravi, and Zhenyu Wang, 1996, The conditional CAPM and the cross-section of expected returns, *Journal of Finance* 51, 3-53.
- Jostova, Gergana, and Alex Philipov, 2002, Bayesian analysis of stochastic betas, Working paper, Boston College.
- Keim, Donald B., and Robert F. Stambaugh, 1986, Predicting returns in the stock and bond markets, *Journal of Financial Economics* 17, 357-390.
- Kim, Tong S., and Edward Omberg, 1996, Dynamic nonmyopic portfolio behavior, *Review of Financial Studies* 9, 141-161.
- Lee, Charles M. C., James N. Myers, and Bhaskaran Swaminathan, 1999, What is the intrinsic value of the Dow? *Journal of Finance* 64, 1693-1741.
- Lettau, Martin, and Sydney Ludvigson, 2001, Consumption, aggregate wealth, and expected returns, *Journal of Finance* 56, 815-849.
- Lettau, Martin, and Sydney Ludvigson, 2002, Measuring and modeling variation in the risk-return tradeoff, Working paper, NYU.
- Liu, Jun, 1999, Portfolio selection in stochastic environments, Working paper, UCLA.
- Menzly, Lior, Tano Santos, and Pietro Veronesi, 2003, The time series of the cross section of asset prices, Working paper, University of Chicago.
- Mundell, Robert, 1963, Inflation and real interest, *Journal of Political Economy* 71, 280-283.
- Newey, Whitney, and Kenneth West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703-708.
- Shanken, Jay, 1990, Intertemporal asset pricing – An empirical investigation, *Journal of Econometrics* 45, 99-120.
- Tobin, James, 1965, Money and economic growth, *Econometrica* 33, 671-684.
- Vuolteenaho, Tuomo, 2002, What drives firm-level stock returns?, *Journal of Finance* 57, 233-264.

Table 1: Sample Moments**Panel A:** Selected Summary Statistics

	Returns			Dividend Growth g_t			Beta β_t			Payout Ratio po_t		
	mean	stdev	α	mean	stdev	auto	mean	stdev	auto	mean	stdev	auto
Growth	0.10	0.22	-0.02	0.05	0.28	-0.27	1.18	0.10	0.76	0.26	0.11	0.69
Neutral	0.13	0.15	0.02	0.07	0.13	-0.11	0.96	0.07	0.76	0.42	0.08	0.64
Value	0.16	0.18	0.04	0.09	0.19	0.06	0.99	0.17	0.86	0.41	0.12	0.62
Average Industry	0.13	0.21	-0.01	0.05	0.21	0.04	1.07	0.19	0.76	0.36	0.13	0.38

Panel B: Risk Premium Regression

	estim	std err	p-val
const	0.08	0.05	0.13
r	-0.71	0.90	0.43
cay	1.97	1.66	0.24

Panel A reports summary statistics mean, standard deviation (stdev) and annual autocorrelation (auto) for total returns, cashflow growth g_t and betas β_t of book-to-market decile portfolios 1 (growth), 6 (neutral) and 10 (value) and the average mean, average standard deviation and average autocorrelation across 46 industry portfolios. All growth rates and returns are continuously compounded and have an annual horizon but are sampled at a monthly frequency. The column labeled α denotes the CAPM alpha, from running a regression of monthly excess portfolio returns onto a constant (α) and the excess market return. The alpha is reported as an annualized number. The sample period is July 1965 to December 2000 for the book-to-market portfolios and January 1965 to December 2000 for the industry portfolios. Panel B reports the result of a predictive regression of $y_{t+1}^m - r_t = \alpha + \beta_r r_t + \beta_{cay} cay_t$, where y_t^m is the annual market return, r_t is a 1 year zero coupon bond rate, and cay is Lettau-Ludvigson (2002)'s consumption-asset-labor deviations, estimated recursively. The sample period is June 1965 to December 2000 and the regression is run at a monthly frequency.

Table 2: Companion Form Φ Parameter Estimates

		g_t	β_t	Δpo_t	r_t	cay_t	π_t
Growth Stocks B/M Decile = 1	g_t	-0.35 (0.17)	0.45 (0.32)	0.37 (0.32)	-4.06 (1.43)	1.86 (2.40)	1.69 (1.23)
	β_t	-0.00 (0.03)	0.68 (0.11)	-0.08 (0.08)	0.48 (0.48)	0.85 (0.54)	-0.74 (0.44)
	Δpo_t	0.04 (0.02)	0.11 (0.15)	-0.37 (0.18)	0.71 (0.58)	-0.19 (1.42)	-0.31 (0.68)
	r_t	-0.00 (0.00)	-0.02 (0.03)	-0.04 (0.03)	0.60 (0.12)	0.21 (0.18)	0.14 (0.14)
	cay_t	0.00 (0.01)	0.03 (0.01)	-0.01 (0.02)	0.09 (0.08)	0.54 (0.09)	0.07 (0.05)
	π_t	0.01 (0.00)	-0.04 (0.04)	-0.03 (0.03)	-0.09 (0.16)	0.07 (0.16)	0.73 (0.15)
Neutral Stocks B/M Decile = 6	g_t	-0.13 (0.18)	0.03 (0.27)	0.61 (0.24)	-1.60 (0.91)	-0.06 (1.58)	1.22 (1.12)
	β_t	-0.00 (0.05)	0.57 (0.12)	-0.11 (0.09)	1.20 (0.38)	-0.23 (0.50)	-0.10 (0.34)
	Δpo_t	0.12 (0.07)	-0.02 (0.10)	-0.32 (0.13)	0.83 (0.51)	0.11 (0.57)	-0.38 (0.34)
	r_t	0.02 (0.01)	0.02 (0.03)	-0.01 (0.04)	0.58 (0.13)	0.12 (0.15)	0.14 (0.12)
	cay_t	0.00 (0.01)	0.01 (0.02)	-0.00 (0.01)	0.06 (0.08)	0.65 (0.08)	0.02 (0.05)
	π_t	0.00 (0.02)	0.03 (0.05)	-0.01 (0.03)	-0.16 (0.18)	-0.09 (0.20)	0.81 (0.15)
Value Stocks B/M Decile = 10	g_t	-0.06 (0.12)	0.20 (0.20)	-0.16 (0.13)	1.37 (1.19)	5.83 (1.50)	-1.26 (1.48)
	β_t	-0.04 (0.04)	0.84 (0.07)	-0.12 (0.07)	-0.12 (0.42)	0.40 (0.82)	0.74 (0.44)
	Δpo_t	0.16 (0.05)	-0.15 (0.14)	-0.43 (0.20)	1.16 (0.42)	0.72 (1.09)	0.44 (0.46)
	r_t	0.01 (0.01)	0.02 (0.01)	-0.04 (0.01)	0.57 (0.11)	0.16 (0.14)	0.16 (0.14)
	cay_t	0.00 (0.00)	0.00 (0.01)	0.01 (0.01)	0.06 (0.06)	0.63 (0.09)	0.01 (0.05)
	π_t	0.03 (0.01)	0.05 (0.02)	-0.04 (0.01)	-0.17 (0.14)	-0.08 (0.15)	0.65 (0.14)
Average Industry	g_t	-0.16 (0.24)	-0.04 (0.19)	0.19 (0.22)	-0.75 (0.00)	1.41 (0.00)	1.02 (0.00)
	β_t	0.00 (0.01)	0.91 (0.13)	-0.02 (0.26)	-0.03 (0.02)	0.10 (0.00)	0.10 (0.00)
	Δpo_t	-0.01 (0.01)	0.02 (0.02)	-0.45 (0.14)	0.40 (0.02)	0.40 (0.01)	0.14 (0.00)
	r_t	0.00 (0.01)	0.00 (0.03)	-0.00 (0.00)	0.58 (0.02)	0.11 (0.01)	0.18 (0.02)
	cay_t	0.00 (0.01)	0.00 (0.01)	-0.00 (0.00)	0.07 (0.00)	0.64 (0.01)	0.02 (0.03)
	π_t	0.01 (0.11)	0.01 (0.08)	-0.00 (0.00)	-0.11 (0.00)	-0.11 (0.00)	0.80 (0.02)

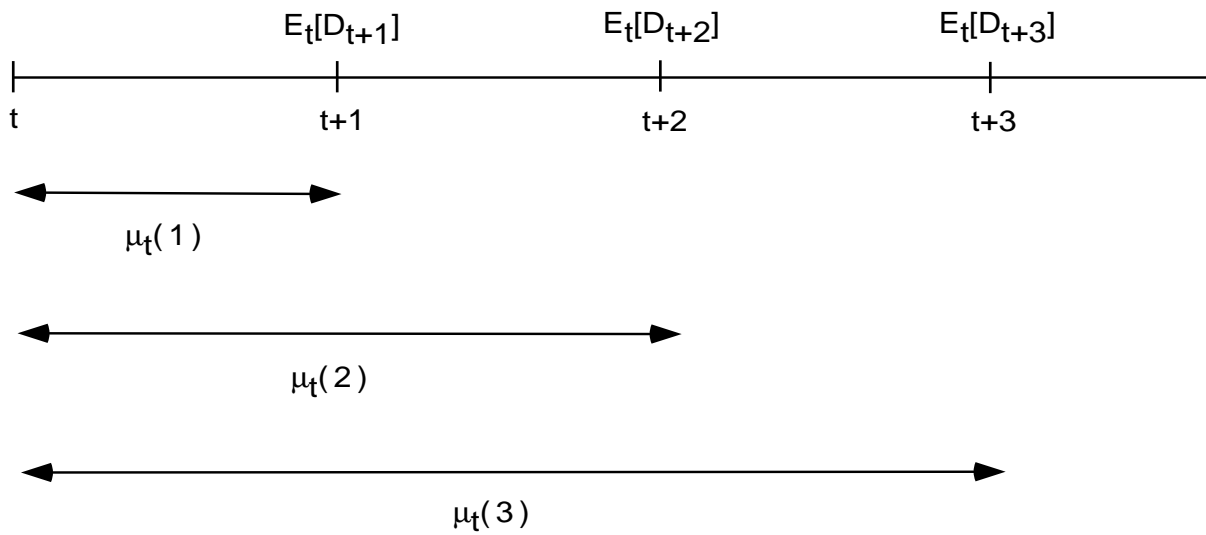
The table reports estimates of the companion form Φ of the VAR in (7). The estimation is done at an annual horizon, using monthly (overlapping) data. For the average industry results, we pool data across all industries. Standard errors are computed using Newey-West (1987) 12 lags. Parameters significant at the 95% level are denoted in bold. The sample period is July 1970 to December 2000 for the book-to-market sorted portfolios and from January 1970 to December 2000 for the industry portfolios.

Table 3: Mis-Pricing of Portfolios

	Perpetuity Value	Mis-Pricing Errors %		
		Unconditional CAPM	Ignoring Beta	Ignoring Risk Premium
Book-to-Market Sorted Portfolios				
1 Growth	11.17	-13.41	3.72	-8.33
2	16.39	-31.98	-9.69	-28.74
3	10.81	-7.65	12.39	-9.29
4	10.90	-15.09	1.18	-16.31
5	10.97	-15.48	0.84	-18.45
6	8.93	-13.96	-2.34	-14.42
7	9.09	-9.89	-0.22	-13.21
8	7.78	-13.30	-6.85	-14.09
9	7.25	-18.83	-9.44	-8.63
10 Value	7.02	-13.54	-4.04	-7.56
Average Mis-pricings Across Book-to-Market Sorted Portfolios				
Mean Error		-15.31	-1.45	-13.39
Stdev Error		6.59	6.67	6.39
Selected Industries				
FabPr	17.65	-32.85	-7.04	-16.77
Ships	16.10	-57.87	-51.84	4.84
Average Mis-pricings Across All Industry Portfolios				
Mean Error		-16.89	-4.81	-12.74
Stdev Error		9.94	9.05	6.82

We value a perpetuity of an expected cashflow of \$1 received at the end of each year using the time-varying expected returns for each book-to-market portfolio at the end of December 2000. We report percentage mis-pricing errors (wrong-correct)/correct for valuation using a wrong model versus the full model valuation. Three wrong models are considered: using a constant discount rate, ignoring the time-varying betas and ignoring the time-varying market risk premium.

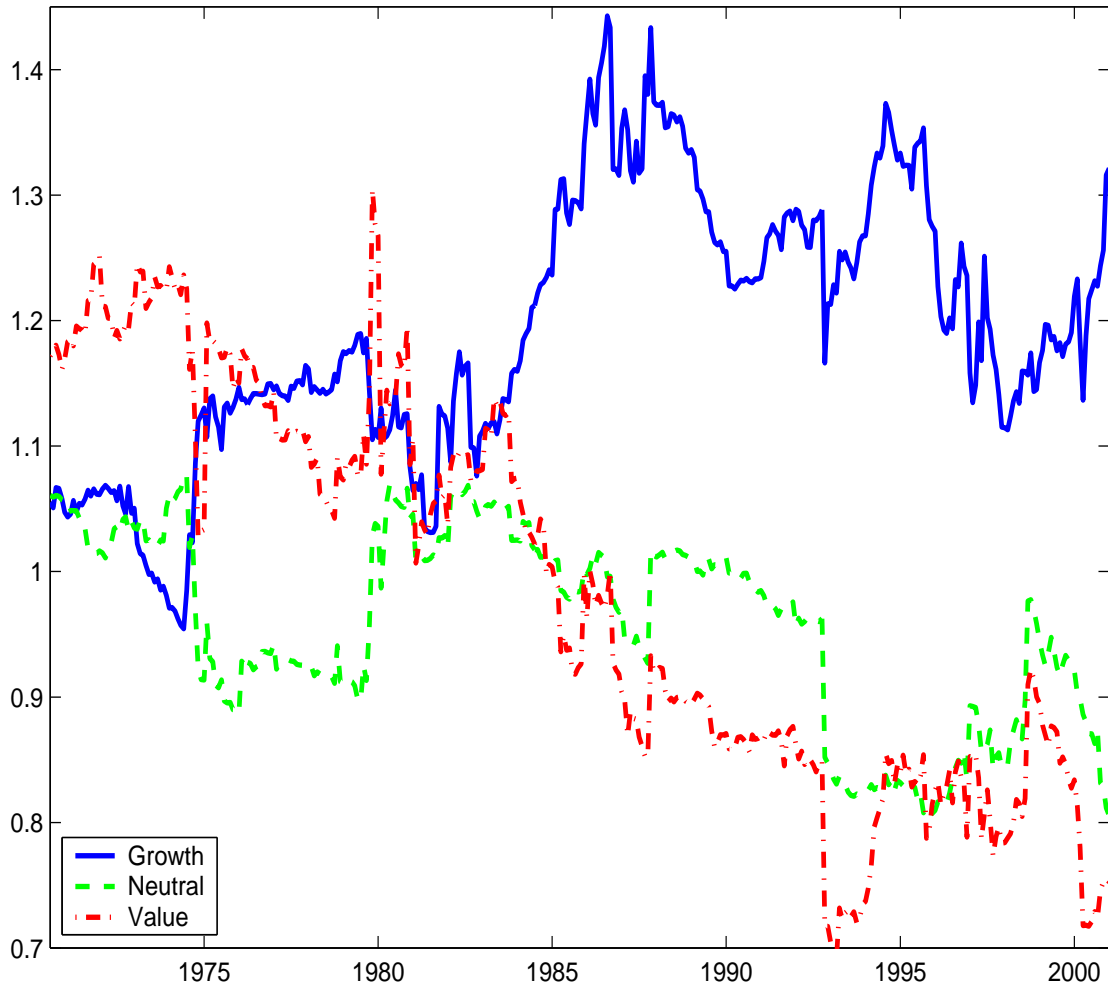
Figure 1: The Spot Discount Curve $\mu_t(n)$



The spot expected return $\mu_t(n)$ is used to discount an expected risky cashflow $E_t[D_{t+n}]$ of a security at time $t + s$ back to time t . The spot expected return $\mu_t(n)$ solves:

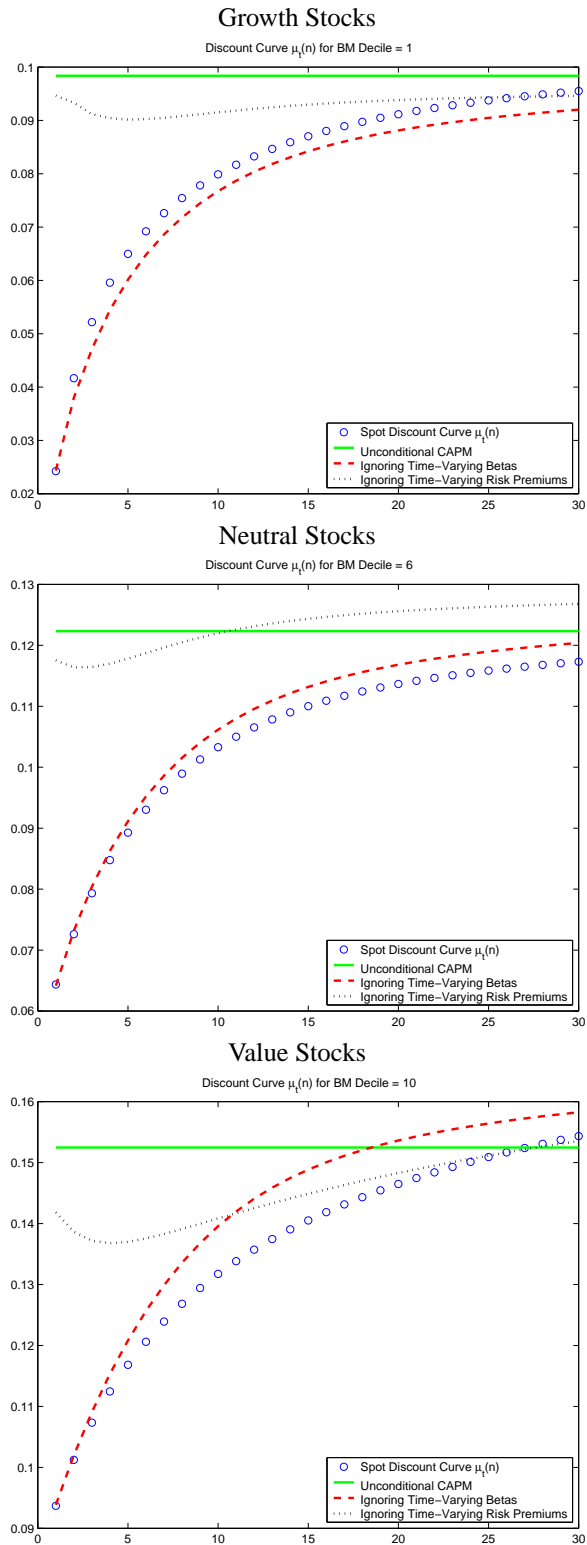
$$E_t \left[\left(\sum_{k=0}^{n-1} \exp(-\mu_{t+k}) \right) D_{t+n} \right] = \frac{E_t[D_{t+n}]}{\exp(n \cdot \mu_t(n))},$$

where μ_t is the one period expected return from t to $t + 1$.



The figure shows time-varying betas of growth, neutral and value stocks, computed using rolling 60-month regressions of excess portfolio returns on market excess returns.

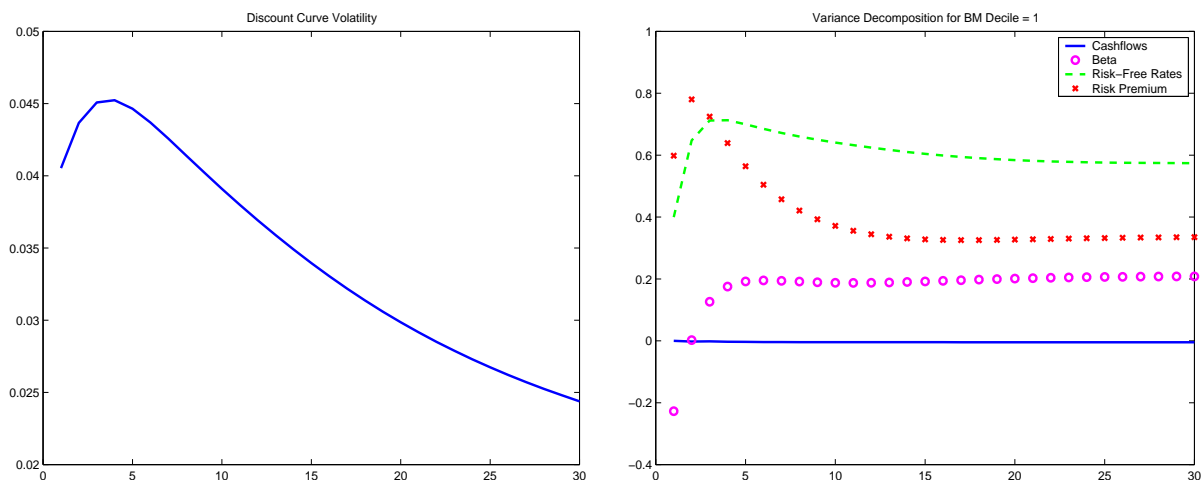
Figure 2: Time-Varying Beta of Book-to-Market Portfolios



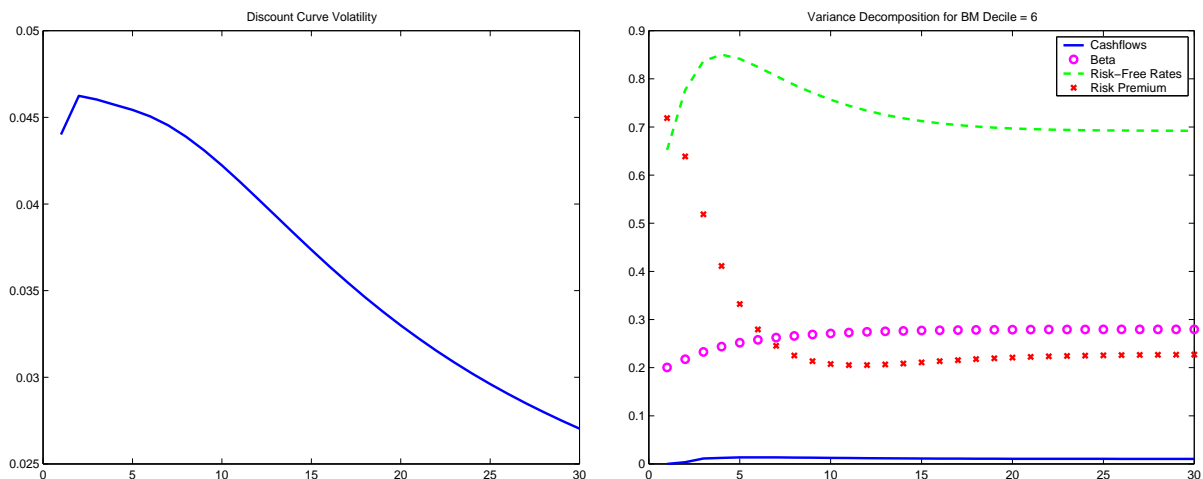
The figure shows discount curves $\mu_t(n)$, with n in years on the x -axis, computed at the end of December 2000 for various book-to-market portfolios.

Figure 3: Discount Curves

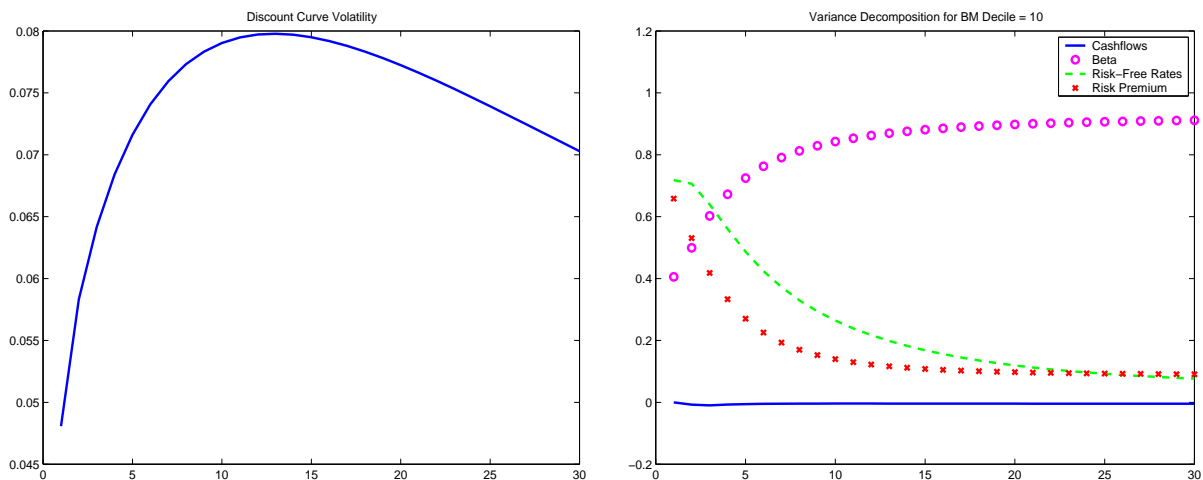
Growth Stocks



Neutral Stocks

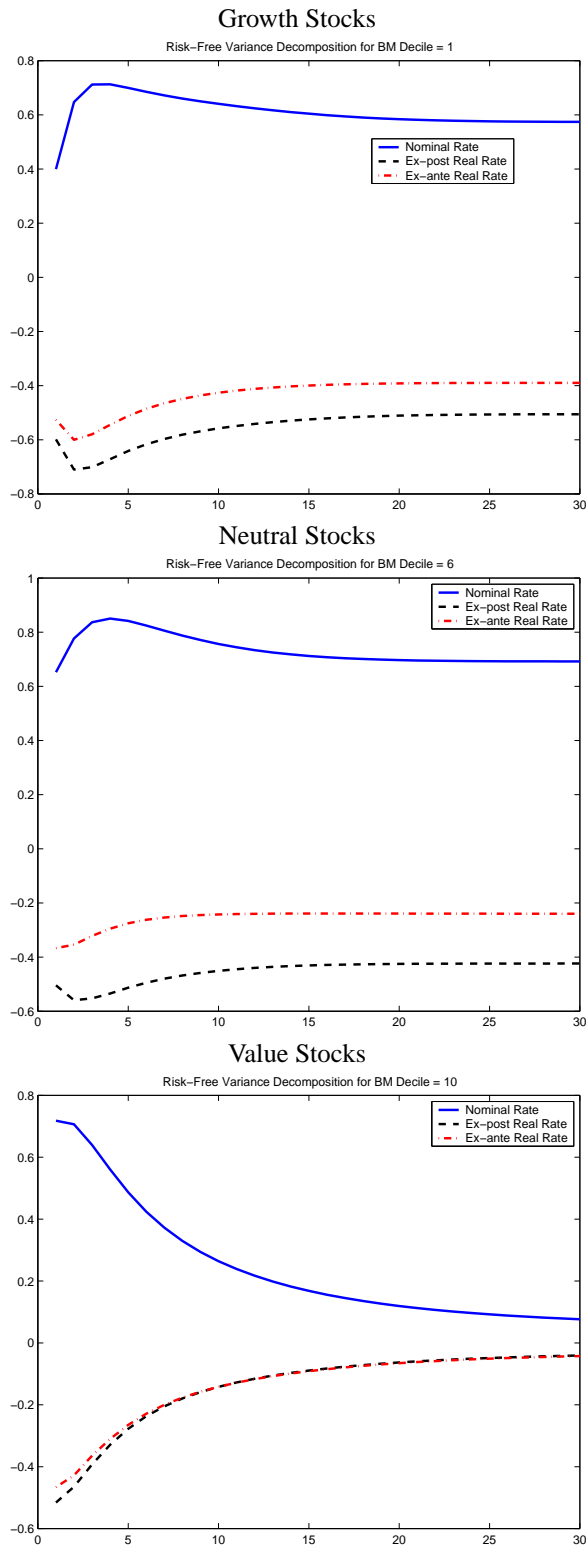


Value Stocks



The left-hand column plots $\sqrt{\text{var}(\mu_t(n))}$, for each n on the x -axis. The right-hand column attributes the $\text{var}(\mu_t(n))$ into proportions due to dividend growth, beta, the risk-free rate and the risk premium. The proportions double-count the covariances and so do not sum to 1.

Figure 4: Variance Decomposition for the Term Structure of Discount Rates



The figure shows risk-free variance decompositions for growth, neutral and value stocks for nominal risk-free rates (r_t), ex-post real rates ($r_t - \pi_t$) and ex-ante real rates $E_t(r_{t+1} - \pi_{t+1})$. The proportions double-count the covariances and so do not sum to 1.

Figure 5: Risk-Free Rate Variance Decompositions