Consumption Strikes Back?*

Lars Peter Hansen[†]
John C. Heaton[‡]
Nan Li[§]

July 14, 2004

*Conversations with Fernando Alvarez, Ravi Bansal, John Cochrane, Chris Sims and Pietro Veronesi were valuable in completing this paper. Hansen gratefully acknowledges support from the National Science Foundation, Heaton from the Center for Research in Securities Prices, and Li from the Olin Foundation.

[†]Department of Economics, University of Chicago and NBER
[‡]Graduate School of Business, University of Chicago and NBER
[§]Department of Economics, University of Chicago

1 Introduction

Economically grounded models of asset pricing feature a role for information and risk as a device for explaining return heterogeneity. While much research in the past has questioned the empirical relevance of risk identified from economic aggregates such as consumption, recent investigations have documented a qualitative role for consumption and other macroeconomic aggregates in explaining return heterogeneity. Much of this evidence features long-run risk, risk that is related to how either consumption, dividends or returns respond to shocks far into the future. Our paper aims to explore the resulting empirical challenges. How sensitive are risk-measures to details in the specification of the time series evolution? How accurately can we measure these components? When should we expect these components to play a fundamental role in valuation? To address these questions we are compelled to address formally the role of specification, measurement and pricing theory in quantifying risk.

Many models of investor preferences and decision-making emphasize the importance of the intertemporal timing of risk. Examples include recursive utility theory, habit persistence,[1] and staggered decision making.[2] In this paper we consider the preference specification developed by Epstein and Zin (1989b) and Weil (1990), together with a log-linear model of consumption and cash flows, to explore the challenges in measuring and modeling long-run risk. The resulting statistical depiction of the time series allows us to evaluate easily the importance of alternative macroeconomic and financial variables in predicting future consumption. Different specifications of the time series evolution of consumption give rise to different configurations of shocks pertinent to risk assessment. Moreover, specification of the long-run component of this evolution alters the implied riskiness of returns and cash flows.

Even under time and state separable specifications of investor preferences, the intertemporal composition of cash-flow riskiness is important to investors and hence to model-builders. This composition reflects the potential uncertainty that investors confront and how this uncertainty is encoded in asset values. In addition to exploring challenges for statistical measurement, we provide a theoretical characterization of when the riskiness of cash flows far into the future is an important ingredient in current values.

In section 2 we study a familiar model of asset prices to show why the intertemporal composition of risk might matter to an investor. Following Hansen and Singleton (1983) the model is deliberately stylized in order to yield restrictions on linear time series models. In section 3 we examine whether the model's prediction about intertemporal risk and returns are consistent with time series evidence. The predicted relationship between average returns and the covariance between returns and future consumption is examined. Section 4 develops a different notion of risk based on the low frequency properties of cash flows and consumption. Starting with the assumption of cointegration between cash flows and consumption we construct a decomposition of security prices that displays the contribution of long-run risk to returns and prices. In section 5 we provide statistical evidence for cointegration between consumption and the dividends of portfolios of stocks. We examine whether measured dif-

[1] Examples include Constantinides (1990), Heaton (1995), and Sundaresan (1989)

[2] Examples include Gabaix and Laibson (2002) and Lynch (1996)

ferences in long-run risk exposure are consistent with differences in average returns across portfolios. Section 6 concludes.

2 Asset Pricing

Models of asset pricing link investor preferences and opportunities to deduce equilibrium relations for returns and prices. These models aim to explain return heterogeneity by the existence of risk premia. Investors require larger expected returns as compensation for holding risky portfolios. Alternative asset pricing models imply alternative risk-return tradeoffs. Equivalently [*e.g*, see Hansen and Richard (1987)] they imply an explicit model of a stochastic discount factors, the variables $S_{t+1,t}$ used by investors to value one-period and hence multiple period assets.

There remains considerable controversy within the asset pricing literature about the feasibility of constructing an economically meaningful model of stochastic discount factors and hence risk premia. Nevertheless in this section we find it useful to consider one such model that, by design, leads to tractable restrictions on economic time series. This model is rich enough to help us examine return heterogeneity as it relates to risk and to understand better the intertemporal values of equity.

2.1 Preferences

We follow Epstein and Zin (1989b) by depicting preferences recursively. As we show below this model of preferences provides a simple justification for examining a long-run relationship between consumption and returns. In addition this model of preferences provides a separation between risk aversion and the elasticity of intertemporal substitution. This feature is convenient because it allows us to examine the effects of changing risk exposure while not dramatically altering the risk-free rate.

In our specification of these preferences we use a CES recursion:

$$V_t = \left[(1-\beta)\left(C_t\right)^{1-\rho} + \beta\mathcal{R}_t(V_{t+1})^{1-\rho}\right]^{\frac{1}{1-\rho}} . \tag{1}$$

Under a Cobb-Douglas specification ($\rho = 1$), this recursion becomes:

$$V_t = \left(C_t\right)^{(1-\beta)} \mathcal{R}_t(V_{t+1})^{\beta}.$$

In what follows, the $\rho = 1$ will receive special attention. In these recursions, the function $\mathcal{R}_t$ adjusts the continuation value for risk via:

$$\mathcal{R}_t(V_{t+1}) = \left[E\left(V_{t+1}\right)^{1-\theta} |\mathcal{F}_t\right]^{\frac{1}{1-\theta}}$$

where $\mathcal{F}_t$ is the current period information set.

We will need to include stochastic growth in consumption. This leads us to study an alternative recursion:

$$\frac{V_t}{C_t} = \left[(1-\beta) + \beta \mathcal{R}_t \left(\frac{V_{t+1}}{C_{t+1}}\frac{C_{t+1}}{C_t}\right)^{1-\rho}\right]^{\frac{1}{1-\rho}}$$

Since consumption and continuation values are positive, we find it convenient to work with logarithms instead. Let v_t denote the logarithm of the continuation value relative to the logarithm of consumption, and let c_t denote the logarithm of consumption. We rewrite recursion (1) as

$$v_t = \frac{1}{1-\rho}\log\left((1-\beta) + \beta \exp\left[(1-\rho)\,\mathcal{Q}_t(v_{t+1}+c_{t+1}-c_t)\right]\right), \tag{2}$$

where $\mathcal{Q}_t$ is the so-called risk-sensitive recursion:

$$\mathcal{Q}_t(v_{t+1}) = \frac{1}{1-\theta}\log E\left(\exp\left[(1-\theta)v_{t+1}\right]|\mathcal{F}_t\right).$$

(See Hansen and Sargent (1995) and Tallarini (1998).)

2.2 Shadow Valuation

Consider the shadow valuation associated with a given consumption process. The utility recursion gives rise to a corresponding valuation recursion and implies a stochastic discount factor used to represent this valuation.

The first utility recursion (1) is homogeneous of degree one in consumption and the future continuation utility. Use Euler's Theorem to write:

$$V_t = (MC_t)C_t + E\left[(MV_{t+1})V_{t+1}|\mathcal{F}_t\right] \tag{3}$$

where

$$\begin{aligned} MC_t &= (1-\beta)(V_t)^{\rho}(C_t)^{-\rho} \\ MV_{t+1} &= \beta(V_t)^{\rho}\left[\mathcal{R}_t(V_{t+1})\right]^{\theta-\rho}(V_{t+1})^{-\theta} \end{aligned}$$

The right-hand side of (3) measures the shadow value of consumption today and the continuation value of utility tomorrow.

Let consumption be numeraire, and suppose for the moment that we value claims to the future continuation value V_{t+1} as a substitute for future consumption processes. Divide both sides of (3) by MC_t and use marginal rates of substitution to compute shadow values. The shadow value of a claim to a continuation value is priced using $\frac{MV_{t+1}}{MC_t}$ as a stochastic discount factor. Moreover, equilibrium wealth is given by $W_t = \frac{V_t}{MC_t}$. A claim to next period's consumption is valued using

$$S_{t+1,t} = \frac{MV_{t+1}MC_{t+1}}{MC_t} = \beta\left(\frac{C_{t+1}}{C_t}\right)^{-\rho}\left(\frac{V_{t+1}}{\mathcal{R}_t(V_{t+1})}\right)^{\rho-\theta}$$

as a stochastic discount factor. There are two contribution (typically highly correlated) contributions to the stochastic discount factor. One is the direct consumption growth contribution familiar from the Lucas (1978) model of asset pricing. The other is the continuation value relative to its risk adjustment. The contribution is forward-looking and is present provided that ρ and θ differ.

A challenge in using this model empirically is to model V_{t+1} which is linked to future consumption via the recursion (1). One approach is to use the relationship between wealth and the continuation value, $W_t = V_t/MC_t$, to construct a representation of the stochastic discount factor based on consumption growth and the return to a claim on future consumption. In general this return is unobservable. An aggregate stock market return is sometimes used to proxy for this return as in Epstein and Zin (1989a), for example. In this investigation we are interested in maintaining the direct link between the continuation value and future consumption. In the case of logarithmic intertemporal preferences the link between future consumption and the continuation value easily can be calculated as we demonstrate in the next section.

2.3 The special case in which $\rho = 1$

The special case in which $\rho = 1$ gives the recursive counterpart to logarithmic preferences. This can be seen by taking the $\rho = 1$ limit in recursion (2):

$$v_t = \beta \mathcal{Q}_t(v_{t+1} + c_{t+1} - c_t). \tag{4}$$

The stochastic discount factor in this special case is:

$$S_{t+1,t} \equiv \beta \left(\frac{C_t}{C_{t+1}} \right) \left[\frac{(V_{t+1})^{1-\theta}}{\mathcal{R}_t(V_{t+1})^{1-\theta}} \right].$$

Notice that the term associated with the risk-adjustment satisfies

$$E \left[\frac{(V_{t+1})^{1-\theta}}{\mathcal{R}_t(V_{t+1})^{1-\theta}} | \mathcal{F}_t \right] = 1$$

and can thus be thought of as distorting the probability distribution. As is familiar for logarithmic preferences, consumption and wealth are proportional.[3]

Recursion (4) was used by Tallarini (1998) in his study of *risk sensitive* business cycles. An important limiting case occurs when $\theta = 1$. In this case preferences are logarithmic and separable over time and states of the world with discount factor β. As in the asset pricing work of Campbell and others, we use this $\rho = 1$ specification as a convenient benchmark. Campbell (1996) argues for less intertemporal substitution and Bansal and Yaron (2004) assume more, but both use comparable utility recursions.

[3] The constancy of the consumption-wealth when $\rho = 1$ prevents our use of this model to interpret directly the findings of Lettau and Ludvigson (2001b) and Lettau and Ludvigson (2001a).

To make our formula for the marginal rate of substitution operational, we need a formula for V_{t+1} computed using the equilibrium consumption process. Suppose that the first-difference of the logarithm of equilibrium consumption has a moving-average representation:

$$c_t - c_{t-1} = \gamma(L)w_t + \mu_c$$

where $\{w_t\}$ is a vector, *iid* standard normal process and

$$\gamma(z) = \sum_{j=0}^{\infty} \gamma_j z^j$$

where γ_j is a row vector and

$$\sum_{j=0}^{\infty} |\gamma_j|^2 < \infty.$$

This linear times series representation is adopted to help us interpret some of the time series evidence that we will discuss subsequently. Log-linear approximations are often used in macroeconomic modelling, although in what follows we will take the log-linear specification as being correct.

Guess a solution:

$$v_t = v(L)w_t + \mu_v.$$

Rewrite recursion (4) as:

$$v_t = \frac{\beta}{1-\theta} \log E\left(\exp\left[(1-\theta)(v_{t+1} + c_{t+1} - c_t)\right] | \mathcal{F}_t\right).$$

Thus v must solve:

$$zv(z) = \beta[v(z) - v(0) + \gamma(z) - \gamma(0)],$$

which in particular implies that

$$v(0) + \gamma(0) = \gamma(\beta).$$

Solving for v and μ_v:

$$\begin{aligned} v(z) &= \beta \frac{\gamma(z) - \gamma(\beta)}{z - \beta} \\ \mu_v &= \frac{\beta}{1-\beta}[\mu_c + \frac{(1-\theta)}{2}\gamma(\beta) \cdot \gamma(\beta)]. \end{aligned}$$

The formula $v(z)$ is the solution to the forecasting problem:

$$v(L)w_t = \sum_{j=1}^{\infty} \beta^j E\left(c_{t+j} - c_{t+j-1} - \mu_c | \mathcal{F}_t\right)$$

familiar from the rational expectations literature on the permanent income model of consumption. The risk parameter θ enters only the constant term of continuation value process. The term $\gamma(\beta)$, which enters the formulas for v and μ_v is the *discounted impulse response* of consumption growth rate to a shock.

The logarithm of the stochastic discount factor can now be depicted as:

$$s_{t+1,t} \equiv \log S_{t+1,t} = -\delta - \gamma(L)w_{t+1} - \mu_c + (1-\theta)\gamma(\beta)w_{t+1} - \frac{(1-\theta)^2\gamma(\beta)\cdot\gamma(\beta)}{2}$$

where $\beta = \exp(-\delta)$. The term $\gamma(\beta)w_{t+1}$ is the solution to

$$(1-\beta)\sum_{j=0}^{\infty}\beta^j\left[E(c_{t+j}|\mathcal{F}_{t+1}) - E(c_{t+j}|\mathcal{F}_t)\right].$$

It is a geometric average of current and future consumption responses to a shock at a fixed date (say date $t+1$). The discount factor dictates the importance of future responses in this weighted average. As the subjective discount factor β tends to unity, $\gamma(\beta)$ converges to $\gamma(1)$ which is cumulative growth rate response or equivalently the limiting consumption response in the infinite future.

The stochastic discount factor includes both the familiar contribution from contemporaneous consumption plus a forward-looking term that discounts the *impulse responses* for consumption growth. The innovation to the logarithm $s_{t+1,t}$ of the stochastic discount factor is:

$$[-\gamma(0) + (1-\theta)\gamma(\beta)]w_{t+1},$$

which shows how a shock at date $t+1$ alters the stochastic discount factor. This term determines the magnitude of the risk premium.

Example 2.1. *Suppose that consumption evolves according to:*

$$c_{t+1} - c_t = \mu_c + U_c \cdot x_t + \gamma_0 w_{t+1}$$

where z_t evolves according to first-order vector autoregression:

$$x_{t+1} = Gx_t + Hw_{t+1}.$$

The matrix G has strictly stable eigenvalues (eigenvalues with absolute values that are strictly less than one), and $\{w_{t+1} : t = 0, 1, ...\}$ is iid normal with mean zero and covariance matrix I. Then for $j > 0$,

$$\gamma_j = U_c'(G^{j-1})H,$$

and

$$v_t = U_v \cdot x_t + \mu_v$$

where

$$U_v \doteq \beta(I - G'\beta)^{-1}U_c,$$

$$\mu_v \doteq \frac{\beta}{1-\beta}\left[\mu_c + \frac{(1-\theta)}{2}\gamma(\beta)\cdot\gamma(\beta)\right],$$

and

$$\gamma(\beta) = \gamma_0 + \beta U_c'(I - G\beta)^{-1}H.$$

The logarithm of the stochastic discount factor is:

$$s_{t+1,t} = -\delta - \mu_c - U_c \cdot x_t - \gamma_0 w_{t+1} + (1-\theta)\gamma(\beta)w_{t+1} - \frac{(1-\theta)^2\gamma(\beta)\cdot\gamma(\beta)}{2}$$

One goal of our analysis is to interpret differences in observed average returns. To map returns into our model, write the logarithm of the gross return to security j as r_t^j. This return is assumed to have a moving-average representation of the form:

$$r_t^j = \mu_r^j + \kappa^j(L)w_t$$

When $\rho = 1$ the model implies that:

$$E[r_{t+1}^j|\mathcal{F}_t] - r_{t+1}^f = \mu_r^j - r_{t+1}^f = -\frac{|\kappa^j(0)|^2}{2} + [\gamma(0) + (\theta-1)\gamma(\beta)]\cdot\kappa^j(0) \tag{5}$$

where r_{t+1}^f is the logarithm of the (conditionally) risk-free return. The impact of risk is reflected in the familiar term $\gamma(0)\cdot\kappa^j(0)$ which measures the conditional covariance between the return and consumption growth as in Hansen and Singleton (1983). The second component measures the conditional covariance between the return and discounted future consumption. To the extent that shocks to returns have different ability to forecast future consumption, these returns will different means. Large values of the risk parameter θ enhance the importance of this component. This latter effect is featured in the analysis of Bansal and Yaron (2004).[4]

This model presents a measurement challenge for an econometrician. More than just the one-period response of consumption to underlying economic shocks matters. In addition the discounted response of consumption in underlying economic shocks is what is required to quantify the risk that matters to investors. For discount factors close to unity, this challenge is known to be more accute.

While this model has a simple and usable characterization of how temporal dependence in consumption growth alters risk premia, it has the counterfactual implication of risk premia that are time invariant. Other authors, including Campbell and Cochrane (1999) argue that risk premia vary over the business cycle. Time varying risk premia could be added to the model by allowing for stochastic variation in volatility as in Bansal and Yaron (2004). We

[4]Anderson, Hansen, and Sargent (2003) suggest a different interpretation for the parameter θ. Instead of risk, this parameter may reflect model misspecification that investors confront by not knowing the precise riskiness that they must confront in the marketplace. As argued by Anderson, Hansen, and Sargent (2003), under this alternative interpretation, $|(1-\theta)\gamma(\beta)|$ is measure of model misspecification that investors have trouble disentangling because this misspecification is disguised by the underlying shocks that impinge on investment opportunities.

see no reason why this complexity, however, simplifies the measurement or approximation problem.

Because of the logarithmic nature of preferences, wealth in this economy is proportional to consumption

$$W_t = \frac{C_t}{1-\beta}.$$

As noted by Gibbons and Ferson (1985), we may use the return on the wealth portfolio as a proxy for the consumption growth rate. In particular, the return on a claim to wealth is:

$$R^w_{t+1} = \frac{W_{t+1}}{\frac{\beta C_t}{1-\beta}} = \frac{C_{t+1}}{\beta C_t}.$$

Thus

$$r^w_{t+1} = c_{t+1} - c_t - \log \beta$$

This leads Campbell and Vuolteenaho (2003) and Campbell, Polk, and Vuolteenaho (2003) to use a market wealth return as a proxy for consumption growth. With this proxy, these papers take $\gamma(0)$ to be the familiar (conditional) CAPM risk adjustment and $(1-\theta)\gamma(\beta)$ as an additional adjustment where γ is now measured using a market return.[5]

2.4 An Expansion

While $\rho = 1$ is a convenient benchmark, we are also interested in departures from this specification. To assess these departures, we consider an expansion for the continuation value around the point $\rho = 1$. Our aim is to compute a derivative Dv_t to use in a first-order approximation:

$$v_t \approx v_t^1 + (\rho - 1) Dv_t^1$$

where v_t^1 is the continuation value for the case in which $\rho = 1$. In appendix 6, we derive the following recursion for the derivative:

$$Dv_t^1 = -\frac{(v_t^1)^2}{2\beta} + \beta \tilde{E}(Dv_{t+1}^1 | \mathcal{F}_t)$$

where $\tilde{E}$ is the distorted expectation operator associated with the density

$$\frac{\left(V_{t+1}^1\right)^{1-\theta}}{E\left[\left(V_{t+1}^1\right)^{1-\theta} | \mathcal{F}_t\right]}.$$

For the log-normal model of consumption, this distorted expectation appends a mean to the shock vector w_{t+1}. The distorted distribution of w_{t+1} remains normal, but instead of mean zero, it has the typically more pessimistic mean of $(1-\theta)\gamma(\beta)$.

[5] Campbell and Vuolteenaho (2003) refer to this second term as the *bad* β term.

The corresponding expansion for the logarithm of the stochastic discount factor is:

$$s_{t+1,t} \approx s^1_{t+1,t} + (\rho - 1) D s^1_{t+1,t}$$

where[6]

$$D s^1_{t+1,t} = v^1_{t+1} - \frac{1}{\beta} v^1_t + (1-\theta) \left[D v^1_{t+1} - \tilde{E} \left(D v^1_{t+1} | \mathcal{F}_t \right) \right].$$

Example 2.1 implies convenient formulas for this expansion: Dv^1_t includes a constant, linear and quadratic terms in the Markov state x_t. The logarithm of the approximating stochastic discount factor will include quadratic terms as well. See appendix 6 for details.

3 Shocks and Benchmark Returns

We use vector autoregressive (VAR) to both identify interesting aggregate shocks and to estimate $\gamma(L)$. As we discuss below we also use these methods to identify important long-run risks. We consider a specification that is rich enough to allow experimentation with different long-run assumptions and different variables that may be important in identifying shocks.

In all of our VAR specifications we let consumption be the first element of an n-dimensional vector y_t. The least restrictive specification we consider is:

$$A_0 y_t + A_1 y_{t-1} + A_2 y_{t-2} + ... + A_\ell y_{t-\ell} + B_0 + B_1 t = w_t , \tag{6}$$

The vectors B_0 and B_1 are n-dimensional, and similarly the square matrices $A_j, j = 1, 2, ..., \ell$ are n by n. The shock vector w_t has mean zero and covariance matrix I. We normalize A_0 to be lower triangular with positive entries on the diagonals. Form:

$$A(z) \doteq A_0 + A_1 z + A_2 z^2 + ... + A_\ell z^\ell.$$

We are interested in a specification in which $A(z)$ is nonsingular for $|z| < 1$. Given this model, the discounted response of consumption to shocks is given by:

$$\gamma(\beta) \quad = \quad (1-\beta) e_1' A(\beta)^{-1} \tag{7}$$

where e_1 is a column vector with a one in the first position and zeros elsewhere.

[6]There are two ways we could use this formula to approximate one period pricing. We use $\exp\left[s^1_{t+1,t} + (\rho-1) D s^1_{t+1,t}\right]$, as an approximate discount factor, but instead we could have used

$$\exp(s^1_{t+1,t}) \left[1 + (\rho - 1) D s^1_{t+1,t}\right].$$

The second approximation is not necessarily positive, but it produces the first-order expansion of the one-period pricing operator.

3.1 Consumption and Earnings

For our measure of aggregate consumption we use aggregate consumption of nondurables and services taken from the National Income and Product Accounts. This measure is quarterly from 1947 Q1 to 2002 Q4, is in real terms and is seasonally adjusted.

Motivated by the work of Lettau and Ludvigson (2001b) and Santos and Veronesi (2001), in several of our specifications we allow for a second source of aggregate risk that captures aggregate exposure to stock market cash flows. This is measured as the ratio of corporate earnings to aggregate consumption. Corporate earnings are taken from NIPA. In all of the specifications reported below the VAR models were fit using five quarters of lags.

To begin our analysis consider a model in which:

$$y_t = \begin{bmatrix} c_t \\ e_t \end{bmatrix}$$

where e_t is the logarithm of corporate earnings at time t. We consider two specifications of the evolution. In both cases, $B_1 = 0$, so time trends are excluded from the analysis. In one case that is the only restriction, and in the other we restrict the matrix $A(1)$ to have rank one:

$$A(1) = \alpha \begin{bmatrix} 1 & -1 \end{bmatrix}.$$

where the column vector α is freely estimated. This parameterization imposes two restrictions on the $A(1)$ matrix. We refer to the first specification as the *without cointegration* model and second as the *with cointegration* model. The second system imposes a unit root in consumption and earnings, but restricts these series to grow together. Since the cointegration relation we consider is prespecfied, the *with cointegration* model can be estimated as a vector autoregression in the first-difference of the log consumption and the difference between the log earnings and log consumption.

Corporate earnings relative to consumption appears to identify an long-run important shock as demonstrated by figures 1 and 2. First consider figure 1 which plots the spectral densities for consumption and earnings growth with and without cointegration. Under the plausible cointegration restriction, there is an important low frequency component to the consumption growth rate. Because discounted future consumption enters the pricing model this low-frequency component of consumption is potentially important. For example, Bansal and Yaron (2004) argue that this component of consumption is important for understanding the equity premium.

Figure 2 displays the response of consumption and earnings to each shock in the model. A shock to earnings in the model with cointegration has important effects both in the long-run and the short-run.

As can be seen by the figures, the imposition of the cointegration restriction is critical to locating an important low frequency component in consumption. Moreover, this restriction is important in finding feedback from earnings shocks to consumption. Notice from the impulse responses in figure 2, that the earnings shocks has little impact on consumption for the *no cointegration* specification. On the other hand, the spectral densities in figure 1 are

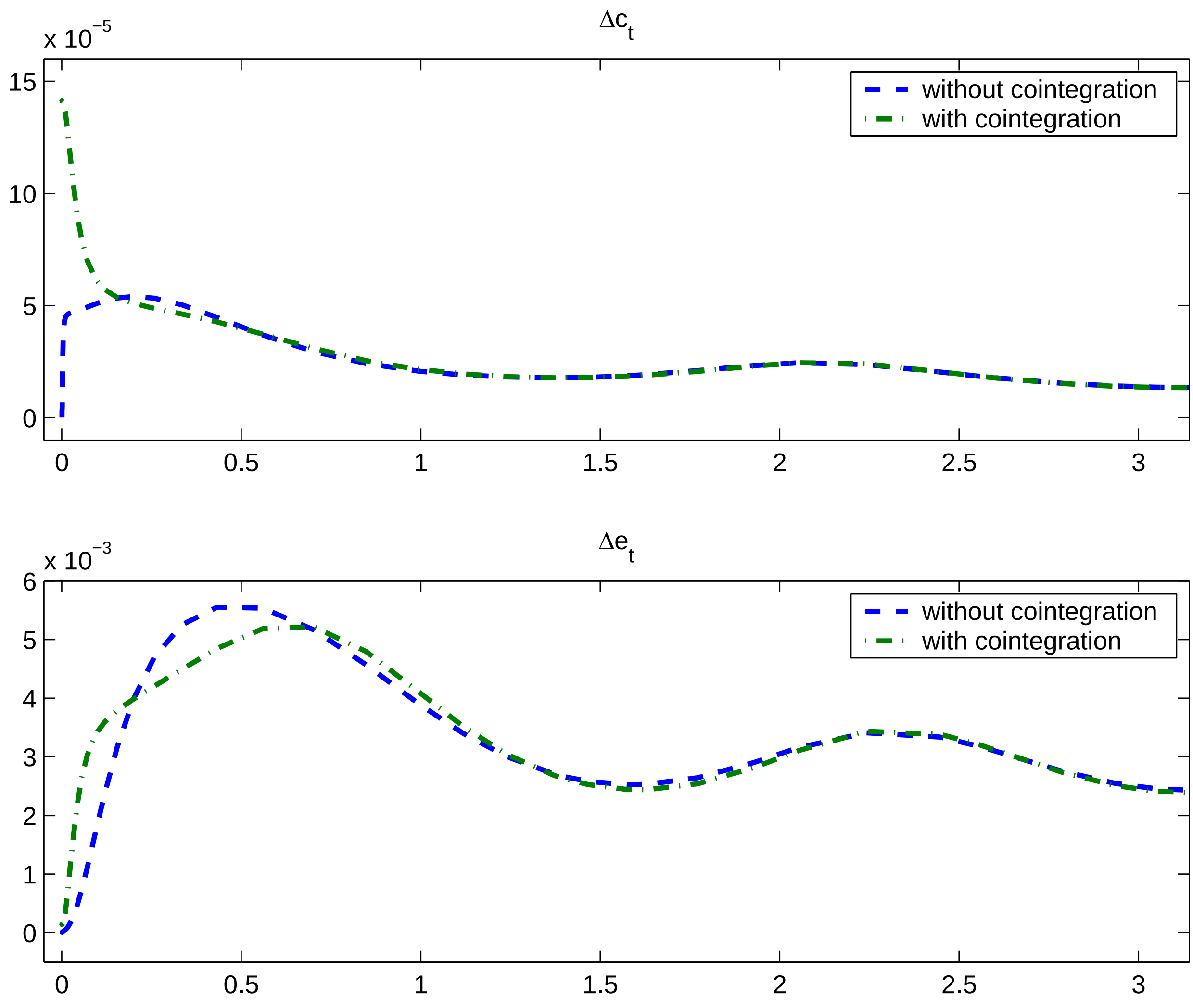

Figure 1: *Spectral Densities of consumption and earnings growth*

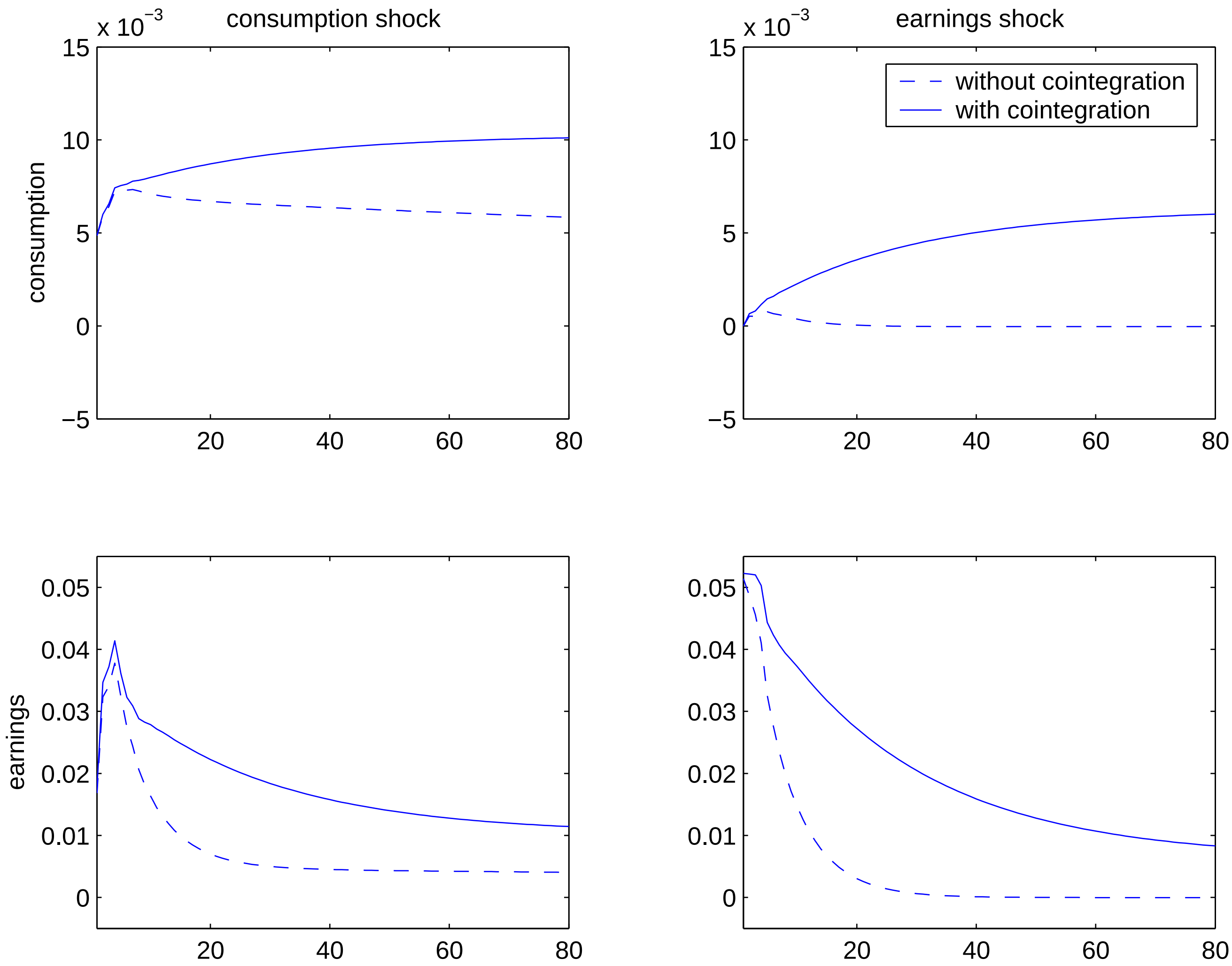

Figure 2: *Impulse Responses of Consumption and Earnings to a Consumption Shock and an Earnings Shock.* The impulse responses without imposing cointegration were constructed from a bivariate VAR with entries c_t, e_t. These responses are given by the dashed lines $---$. Solid lines —— are used to depict the impulse responses estimated from a cointegrated system. The impulse response functions are computed from a VAR with $c_t - c_{t-1}$ and $c_t - e_t$ as time series components.

very close except at a fairly narrow frequency range, suggesting that on purely these two models are hard to distinguish on statistical grounds.[7]

3.2 Book to Market Portfolios

To evaluate the importance of consumption risk in explaining returns, we begin by following Fama and French (1993) and construct portfolios of returns by sorting stocks according to their book-to-market values. We use a coarser sort into 5 portfolios to make our analysis tractable. Summary statistics for the basic portfolios are reported in table 1. Notice that the portfolios are ordered by average book to market values where portfolio 1 has the lowest book-to-market value and portfolio 5 has the highest. Average returns also follow this sort. Portfolio 1 has the lowest average return and portfolio 5 has the highest return. It is well known that the differences in average returns are not well explained by exposure to contemporaneous covariance with consumption. This is reflected in the last row of 1 which reports the correlation between consumption growth and each return. Notice that there is little variation in this measure of risk.

To measure risk in the returns of the five portfolios, combine the cointegrated model for consumption and earnings with the price-dividend ratio for each portfolio and the dividend growth for each portfolio:

$$y_t \doteq \begin{bmatrix} c_t - c_{t-1} \\ e_t - c_t \\ p_t^j - d_t^j \\ d_t^j - d_{t-1}^j \end{bmatrix} \tag{8}$$

where d_t^j and p_t^j are (in logs) the time t dividend paid to portfolio j and the time t price of portfolio j. The construction of portfolio dividends is detailed in Hansen, Heaton, and Li (2004) and follows the work of Bansal, Dittmar, and Lundblad (2002), Menzly, Santos, and Veronesi (2004). We examine each portfolio separately to avoid dramatic parameter proliferation.

Notice that in (8) we do not model returns directly but instead model dividend growth and the price-dividend ratio. To derive an implication for returns, write the one-period gross return to security j as:

$$R_{t+1}^j = \frac{P_{t+1}^j + D_{t+1}^j}{P_t^j} = \frac{(1 + P_{t+1}^j / D_{t+1}^j) D_{t+1}^j / D_t^j}{P_t^j / D_t^j}$$

where P_t^j is the price and D_t^j is the dividend both at time t. Take logarithms and write

$$r_{t+1}^j = \log(1 + P_{t+1}^j / D_{t+1}^j) + (d_{t+1}^j - d_t^j) - (p_t^j - d_t^j)$$

[7]The model with cointegration imposes two restrictions on the matrix $A(1)$. The likelihood ratio for the two model is 5.9. The Bayesian information or Schwarz criterion would select the unrestricted model. From a classical perspective, marginal significance level is just above .05.

Table 1: Properties of Portfolios Sorted by Book-to-Market

	Portfolio					
	1	2	3	4	5	Market
Avg. Return (%)	6.48	6.88	8.90	9.32	11.02	7.23
Std. Return (%)	18.8	16.4	14.8	15.8	17.8	16.5
Avg. B/M	0.32	0.62	0.84	1.12	2.00	0.79
Sharpe Ratio	0.18	0.20	0.28	0.28	0.30	0.21
Correlation with Consumption	0.20	0.18	0.20	0.20	0.21	0.20

Portfolios formed by sorting stocks into 5 portfolios using NYSE breakpoints from Fama and French (1993). Portfolios are ordered from lowest to highest average book-to-market value. Data are quarterly from 1947 Q1 to 2002 Q4 for returns and annual from 1947 to 2001 for B/M ratios. Returns are converted to real units using the implicit price deflator for nondurable and services consumption. Average returns are converted to annual units using the natural logarithm of quarterly gross returns multiplied by 4. The standard deviation of returns is also put in annual units by multiplying the standard deviation of quarterly log gross returns by 2. This assumes that returns are independently distributed over time. "Avg. B/M" for each portfolio is the average portfolio book-to-market over the period computed from COMPUSTAT. The Sharpe Ratio is based on quarterly observations. Correlation with consumption is measured as the contemporaneous correlation between log returns and log consumption growth.

where once again lower case letters denote the corresponding logarithms. Next approximate:

$$\log(1 + P^j_{t+1}/D^j_{t+1}) \approx \log[1 + \exp(\overline{p^j - d^j})] + \frac{1}{1 + \exp(\overline{d^j - p^j})}(p^j_{t+1} - d^j_{t+1} - \overline{p^j - d^j})$$

where $\overline{p^j - d^j}$ is the average logarithm of the price dividend ratio. Use this approximation to write:

$$r^j_{t+1} - (d^j_{t+1} - d^j_t) = \chi^j + \rho^j(p^j_{t+1} - d^j_{t+1}) - (p^j_t - d^j_t) \tag{9}$$

where

$$\rho^j \equiv \frac{1}{1 + \exp(\overline{d^j - p^j})}.$$

As shown by Campbell and Shiller (1988), this approximation is reasonably accurate in practice.

A representation for returns is deduced by considering the implied moving-average representation for the price-dividend ratio and for dividend growth:

$$p^j_t - d^j_t = \pi^j(L)w_t + \mu^j_p \ ,$$

$$d^j_t - d^j_{t-1} = \eta^j(L)w_t + \mu^j_d \ .$$

Then $\kappa^j(z)$ is given by:

$$\kappa^j(z) = (\rho^j - z)\pi^j(z) + \eta^j(z).$$

To examine the model's predictions for average returns consider first the implied conditional covariance between each return and consumption growth from period t to $t + \tau$ as a measure of risk. In other words consider:

$$\kappa^j(0) \cdot \sum_{s=0}^{\tau} \gamma_s$$

This is the measure of risk considered by Parker and Julliard (2003). It can be viewed as an approximation of the term $\gamma(\beta)\cdot\kappa^j(0)$ in (5). Estimates of this quantity for each portfolio and different values of τ are reported in figure 3. The "on impact" covariance between returns and consumption growth occurs when $\tau = 0$. This is an appropriate measure of risk when $\theta = 1$ and preferences are given by time-additive log utility. When $\tau = 0$ there is relatively little difference in the measured covariances and hence little difference in predicted risk. As τ increases, however, we measure substantial changes in the conditional covariances. Parker and Julliard (2003) show that these differences are statistically important in explaining cross-sectional differences in average returns.

The model with recursive utility indicates that the impact of shocks on future consumption are important, but these effects should be discounted by the parameter β. In figure 4 we compute $\kappa^j(0) \cdot \gamma(\beta)$ for each portfolio and for different values of β on the x-axis. For large values of θ this term will dominate the calculation in (5). Discounting narrows the

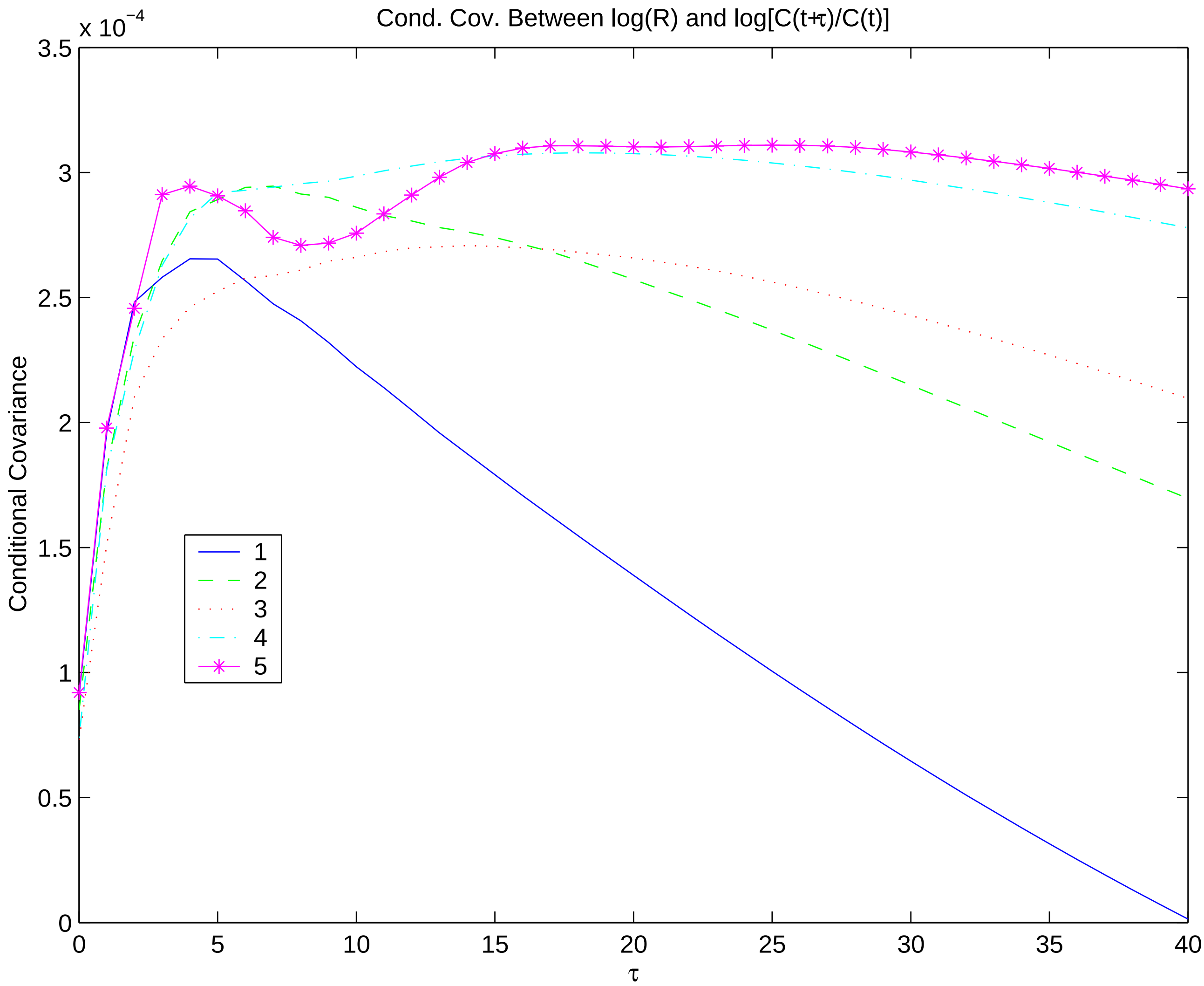

Figure 3: *Conditional covariance between returns and future consumption growth*

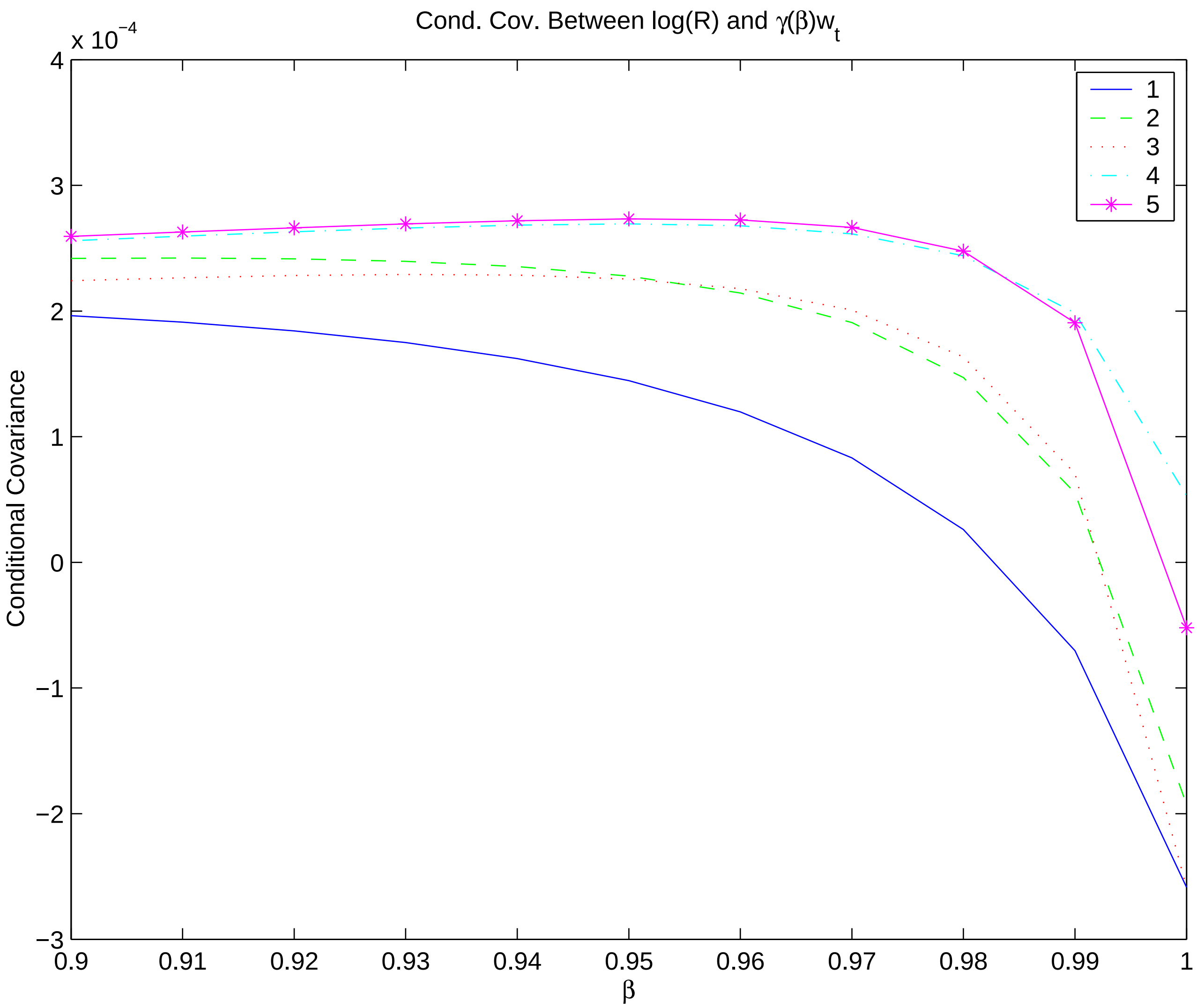

Figure 4: *Conditional covariance between shocks to returns and* $\gamma(\beta)w_t$

differences in the implied risk across portfolios which is not surprising given the patterns in figure 3.

As previously noted, when $\rho = 1$ consumption is proportional to wealth, r^w. We make the additional assumption that the return on the wealth portfolio is reasonably well approximated by the CRSP value weighted return, r_t^M. Under this assumption we consider the following model for y_t:

$$y_t \doteq \begin{bmatrix} r_t^M \\ p_t^M - d_t^M \\ p_t^j - d_t^j \\ d_t^j - d_{t-1}^j \end{bmatrix}$$

and set the first element equal to consumption growth. The implied covariance between portfolio returns and discounted consumption growth are plotted in figure 5. As in figure 4 there are substantial differences in these covariances when we consider portfolios 1 and 5. When the market portfolio is used as a proxy for consumption, however, the ordering of risk across portfolios changes as the discount rate increases.

4 Long Run Cash Flow Risk

Up to this point, we have studied how the long-run composition of consumption risk influences the riskiness of single period returns on equity. We now investigate how long-run risk is encoded in dividend-price ratios. Specifically, we consider when riskiness about long-run cash flow growth can have an important contribution to current value.

Assessing long-run cash flow risk requires both a statistical model and a valuation model. It requires a specification of probabilities governing cash flow growth and a specification of how that growth is valued. We suggest methods for characterizing this risk and the duration of this risk. We aim to characterize when cash flows far into the future contribute to current values and how these contributions are related to risk.

4.1 Dividend growth behavior

The dividends from financial portfolios do not appear to grow one-to-one with consumption. This has been documented in a variety of different places and is evident in figure 6, where we report the logarithms of portfolio dividends relative to aggregate consumption.[8]

When we use book-to-market portfolios, we are using constructed dividend processes in our investigation, it is of interest to consider the role of this construction in generating heterogeneity in dividend growth. For instance, suppose dividend processes on primitive securities display common long run growth but have persistence short run differences. If the primitive assets are sorted based on the expected dividend growth, then this sorting can

[8]In an attempt to construct consumption-dividend ratios that are stationary, Menzly, Santos, and Veronesi (2004) divide consumption by population but not dividends. While population is not a simple time trend, its time series trajectory is much smoother than either consumption or dividends.

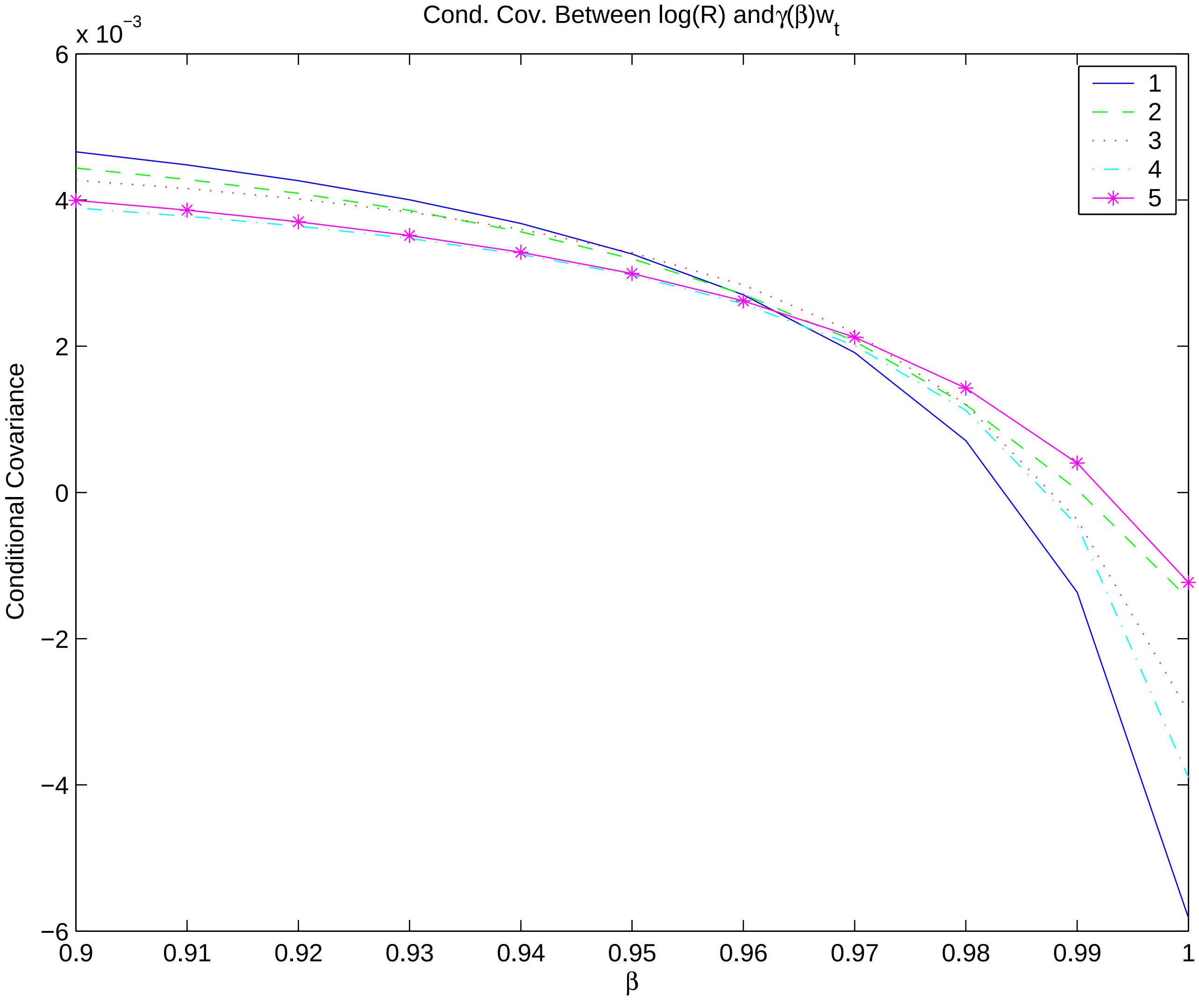

Figure 5: *Conditional Covariance Between Shocks to Returns and* $\gamma(\beta)w_t$ Market return used as proxy for consumption growth

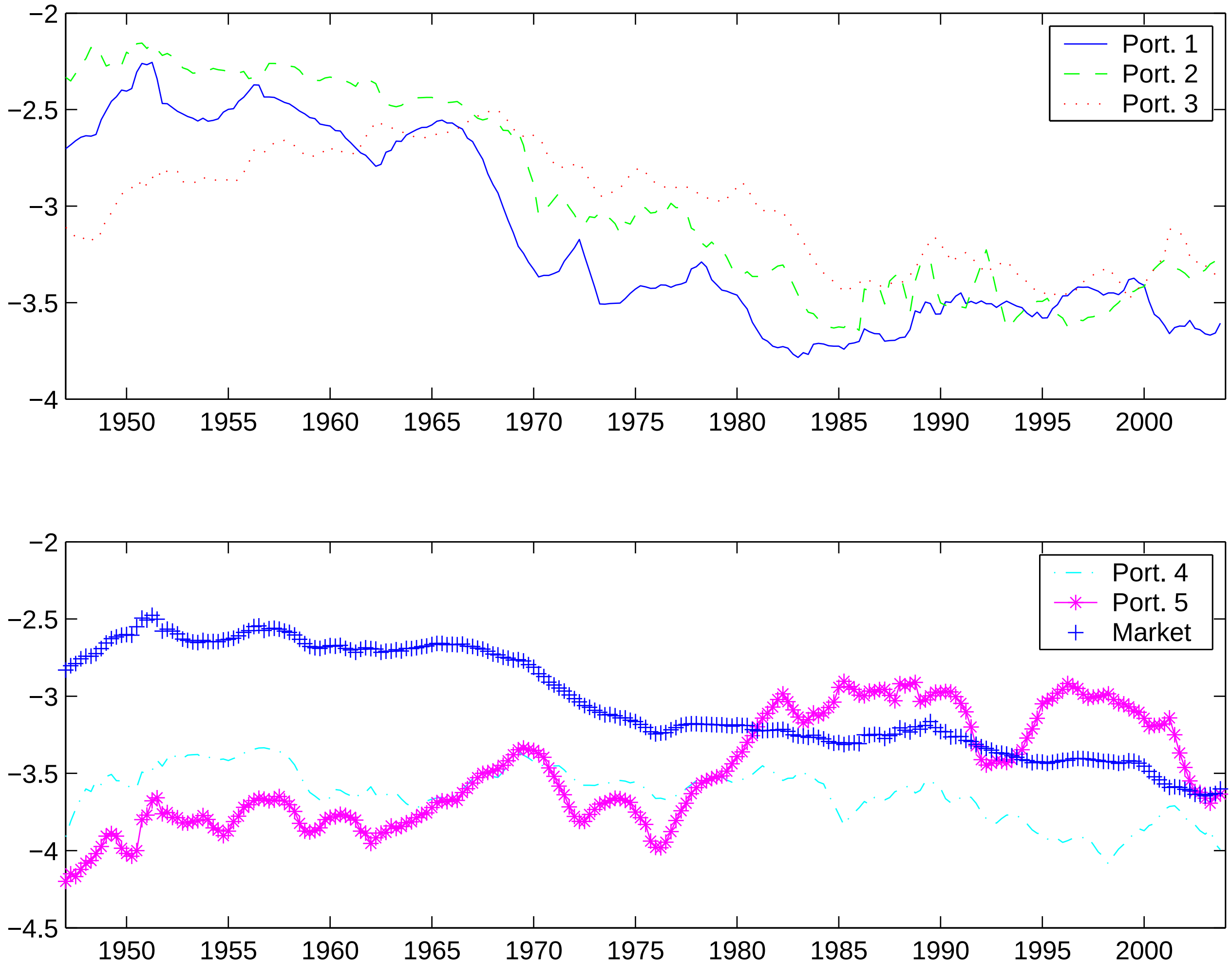

Figure 6: *Portfolio Dividends Relative to Consumption*

induce permanent differences. A correct version of this argument requires consideration of the relative price adjustments needed to accommodate the changing composition over time. This relative price adjustment can offset the expected differences in the dividend growth rates as we illustrate in the following example.

Example 4.1. *Consider an endowment economy with two deterministic endowment processes, (i) and (ii). These endowments oscillate between high and low growth rates. When endowment process (i) grows at a high rate g_h endowment process (ii) grows at a low rate g_ℓ and conversely. The long run growths of both endowments are the same and given by the average of the high and low growth rates. There is no aggregate uncertainty and the equilibrium discount factor* $\exp(s)$ *is constant. For a well defined equilibrium, we require that* $g_h + g_\ell + 2s < 0$.

As in Lucas (1978), let each of the endowment sequences be dividend streams. In this example it suffices to compute two different dividend price ratios. If the endowment sequence is in the low growth state, the price-dividend ratio is:

$$\begin{aligned}(P/D)_\ell &= \exp(s+g_\ell)\sum_{j=0}^{\infty}\exp[j(2s+g_\ell+g_h)] + \sum_{j=1}^{\infty}\exp[j(2s+g_\ell+g_h)] \\ &= \frac{[\exp(s+g_h)+1]\exp(s+g_\ell)}{1-\exp(2s+g_\ell+g_h)}.\end{aligned}$$

Similarly, the price-dividend ratio in the high growth state is:

$$(P/D)_h = \frac{[\exp(s+g_\ell)+1]\exp(s+g_h)}{1-\exp(2s+g_\ell+g_h)}$$

Price dividend ratios oscillate, and returns are equated across the two securities.

Construct two portfolios. One portfolio continually picks the security with the the low growth endowment sequence and the other with the high endowment sequence. Capital gains are reinvested in the portfolios. The price appreciation for the primitive security with the low dividend growth rate between date t and $t+1$:

$$\frac{P_{t+1}}{P_t} = \frac{(P/D)_h}{(P/D)_\ell}\exp(g_\ell) = \frac{1+\exp(s+g_\ell)}{1+\exp(s+g_h)}\exp(g_h).$$

The price appreciation for the high growth security is:

$$\frac{P_{t+1}}{P_t} = \frac{(P/D)_\ell}{(P/D)_h}\exp(g_h) = \frac{1+\exp(s+g_h)}{1+\exp(s+g_\ell)}\exp(g_\ell).$$

While the first portfolio selects the asset with low dividend growth, the growth in the implied portfolio dividends will be determined by the price appreciation. The asset with low growth between today and tomorrow is known to have high growth in the subsequent time period, and its price will appreciate accordingly. The capital gains will be reinvested in a security with a low dividend growth rate.

This example is extreme because of the known high frequency changes in growth rates. As a consequence, the growth rate in the constructed dividends for the first portfolio will actually be higher than that of the second portfolio. If instead the growth rate differences across the two endowments is permanent, there would be no change in the composition of the portfolios. In this case, the price-dividend ratios are constant and constructed dividends grow at the same rate as the dividends on the primitive securities. Persistent but transitory differences in the growth rates will presumably result in intermediate cases in which the difference in the portfolio dividend growth rates will be more muted than the one-period growth rate differences in the primitive securities.

4.2 Long Run Consumption Risk

A convenient time series statistical model that accommodates growth in a restricted way is a model with cointegration between consumption and dividends. Following Bansal, Dittmar, and Lundblad (2002) and others, we consider such a model with time trends included. Specifically, consider a stationary Markov specification for $\{x_t\}$, a process used to depict the underlying valuation. The logarithm of consumption evolves according to:

$$c_{t+1} - c_t = \mu_c(x_t) + \sigma_c(x_t) \cdot w_{t+1}.$$

This allows consumption to grow over time, but growth rates are stationary. Example 2.1 is a special case of this specification with $\mu_c(x_t) = U_c \cdot x_t$ and $\sigma_c(x_t) = \gamma_0$. The consumption process in conjunction with a specification of preferences is used to produce a valuation model as described in section 2.

We also use this consumption process a reference point for cash flow growth as in the time series literature. We study the valuation of cash flows that are co-integrated with consumption:

$$d_t = \lambda c_t + \zeta t + \phi(x_t) \tag{10}$$

where d_t is the logarithm of the cash flow. Since the Markov process is stationary, growth is governed by the parameter pair (λ, ζ). For the time being we allow for nonlinearities in the time series model, although in our computations we will revert to the log-linear specification used previously. As we will argue in the next section, the use of time trends in cointegration models of long-run risk creates both measurement and interpretation problems.

4.3 A Digression on Matrices

We study long-run behavior using eigenvalue methods. Prior to developing these methods, consider a square matrix M raised to a power j. For simplicity suppose the matrix has distinct eigenvalues, and write the eigenvalue decomposition as:

$$M = T\Delta T^{-1}$$

where Λ is a diagonal matrix of eigenvalues. Then

$$M^j = T\Delta^j T^{-1}.$$

Suppose that the largest eigenvalue in absolute value is positive, and call this eigenvalue δ. Then the sequence, $\{\delta^{-j}\Delta^j : j = 1, 2, ...\}$ converges to a matrix of zeros except in one position where there is a one. Thus

$$\delta^{-j} M^j = T(\delta^{-j}\Delta^j)T^{-1}$$

converges to a constant matrix as j gets arbitrarily large. The eigenvalue δ determines the asymptotic growth factor of the matrix sequence M^j and $\log \delta$ is the corresponding asymptotic growth rate.

We use this same approach, but applied to operators. Markov valuation and conditional expectation operators over multiple time intervals can be depicted as iterates of their single-period counterparts. Logarithms of dominant eigenvalues give us a characterization of long-run growth in values and expectations. In what follows, we construct the operators of interest and apply them to characterize long-run behavior.

4.4 Operator Valuation

One counterpart to the matrix M, is a one-period valuation operator given by:

$$\mathcal{P}_\lambda \psi(x) = E\left(\exp\left[s_{t+1,t} + \lambda\left(c_{t+1} - c_t\right)\right]\psi(x_{t+1})|x_t = x\right).$$

Formally, we view this operator as a mapping from L^2 into L^2 where L^2 is the space of (Borel measurable) functions ψ for which

$$E\psi(x_t)^2 < \infty.$$

This operator takes a payoff at date $t+1$ of the form:

$$\exp\left(\lambda c_{t+1}\right)\psi(x_{t+1})$$

and maps it into a price today scaled by $\exp(\lambda c_t)$. Since payoffs and prices are scaled, the valuation operator depends on the choice of λ used in the scaling. Consistent with formula (10), consumption provides the only source of growth in this specification.

Multi-period prices can be inferred from this one-period pricing operator through iteration. The value of a date $t+j$ cash flow:

$$\exp\left[\zeta(t+j) + \lambda c_{t+j}\right]\psi(x_{t+j}) \tag{11}$$

is:

$$\exp\left[\zeta(t+j) + \lambda c_t\right]\left(\mathcal{P}_\lambda\right)^j \psi(x_t).$$

The notation $(\mathcal{P}_\lambda)^j$ denotes the application of the one-period valuation operator j times.

If we take this cash flow to be a dividend process, the date t price-dividend ratio is:

$$\frac{P_t}{D_t} = \frac{\sum_{j=1}^{\infty} \exp(\zeta j)\left[(\mathcal{P}_\lambda)^j \psi(x_t)\right]}{\psi(x_t)}. \tag{12}$$

The term

$$\frac{\exp(\zeta j)\left[(\mathcal{P}_\lambda)^j \psi(x_t)\right]}{\psi(x_t)} \tag{13}$$

is the contribution of the date $t+j$ derivative to the price-dividend ratio, and the price-dividend ratio adds over these objects. Computing these individual terms gives a value decomposition of the price-dividend ratio by time horizon.

Since we allow for the growth rates in the cash flows to vary over time, we shall also have need to define operators that we use to measure these rates and the limiting growth behavior. Let

$$\mathcal{G}_\lambda \psi(x) = E\left(\exp\left[\lambda\left(c_{t+1} - c_t\right)\right] \psi(x_{t+1}) | x_t = x\right).$$

By iterating on this growth operator, we can study expected cash flow growth over multi-period horizons. In particular, the expected value of the cash flow (11):

$$\exp\left[\zeta(t+j) + \lambda c_t\right] \left(\mathcal{G}_\lambda\right)^j \psi(x_t).$$

4.5 Limiting Behavior

For positive cash flows we can characterize the limiting or tail contribution of cash-flow valuation by studying the limit:

$$\lim_{j\to\infty} \frac{1}{j} \log\left[\frac{\exp(\zeta j)\left[(\mathcal{P}_\lambda)^j \psi(x_t)\right]}{\psi(x_t)}\right] = \zeta + \lim_{j\to\infty} \frac{\log\left[(\mathcal{P}_\lambda)^j \psi(x_t)\right]}{j}$$

which will depend on λ but often not depend on ψ. This calculation gives us an asymptotic decay rate in cash flow valuation. The rate measures how long-run prospects about cash flows contribute to current period valuation. This decay rate depends on λ, ζ and the economic value associated with that growth.

Analogously, the asymptotic cash-flow growth is measured as:

$$\lim_{j\to\infty} \frac{1}{j} \log\left[\frac{\exp(\zeta j)\left[(\mathcal{G}_\lambda)^j \psi(x_t)\right]}{\psi(x_t)}\right] = \zeta + \lim_{j\to\infty} \frac{\log\left[(\mathcal{G}_\lambda)^j \psi(x_t)\right]}{j}$$

As in the case of matrices, the asymptotic decay rate is determined by the dominate eigenvalue of the valuation operator $\mathcal{P}_\lambda$. An eigenvalue solves the equation:

$$\mathcal{P}_\lambda \psi_\lambda = \exp(-\nu_\lambda)\psi_\lambda$$

where ψ is an eigenfunction and $\exp(-\nu)$ is an eigenvalue. Then the invariance property of an eigenfunction implies that

$$\left(\mathcal{P}_\lambda\right)^j \psi = \exp(-j\nu_\lambda)\psi_\lambda$$

for any j. The dominant eigenvalue (when it exists) is associated with a strictly positive eigenfunction. Moreover,

$$\lim_{j\to\infty} \frac{\log\left[(\mathcal{P}_\lambda)^j \psi(x_t)\right]}{j} = -\nu_\lambda$$

provided that ψ is strictly positive. This limiting rate is invariant to the specific choice of positive ψ used in defining the cash flow. It does of course depend on λ.

Similarly, we compute the limiting growth rate of the cash flow by solving the eigenfunction problem:

$$\mathcal{G}_\lambda \varphi_\lambda = \exp(\epsilon_\lambda)\varphi_\lambda.$$

Then

$$\lim_{j\to\infty} \frac{\log\left[(\mathcal{G}_\lambda)^j \psi(x_t)\right]}{j} = \epsilon_\lambda$$

for strictly positive ψ. This limit is of interest because because part of the valuation decay over long horizon is tied directly tied to cash flow growth.

4.6 Tail Returns

Cash flow valuation far into the future is dominated by the valuation of the dominant eigenfunction. Returns to the tail component of cash flows are approximated by the return to a security with a payoff given by the dominant eigenfunction. We construct the approximate by supposing, we have a security with a dividend:

$$D_{t+j} = \exp\left[(t+j)\zeta + \lambda c_{t+j}\right] \psi_\lambda(x_{t+j}).$$

This security has a constant price/dividend ratio by construction. Using the eigenvalue property and formula (12)

$$\frac{P_t}{D_t} = \frac{\exp\left(\zeta - \nu_\lambda\right)}{1 - \exp\left(\zeta - \nu_\lambda\right)}.$$

The corresponding return is:

$$R_{t+1} = \left(\frac{D_{t+1}}{D_t}\right) \exp\left(-\zeta + \nu_\lambda\right).$$

Thus a feature of cash flows constructed from the dominant eigenfunction is that the expected returns can be inferred directly from the expected dividend growth. While gross returns are proportional to dividend growth factors, the proportionality factor depends in part on riskiness of the security.

Let R^j_{t+j} denote the return compounded over j time periods. Expected return growth can be inferred directly from expected dividend growth. In particular,

$$\lim_{j\to\infty} \frac{\log E(R^j_{t+j}|x_t)}{j} = \zeta + \epsilon_\lambda - \zeta + \nu_\lambda = \epsilon_\lambda + \nu_\lambda.$$

So far we have shown how to compute the expected value of a long-horizon tail return implied by an asset pricing model. What matters is the value of λ. Not surprisingly, the deterministic component of the cash flow growth vanishes in our return calculation.

It is common in finance to produce a risk-return tradeoff by comparing returns to a risk-free counterpart. Given our interest in long-horizon risk, the natural benchmark is obtained by setting $\lambda = 0$. The dominant eigenvalue of $\mathcal{G}_0$ is one and is associated with a constant eigenfunction. Thus $\xi_0 = 0$. When $\lambda = 0$ the only mechanism for cash-flow growth is a positive value of ζ, which has no impact on returns. Thus the long-horizon counterpart to an expected (logarithm of a) risk-free return is ν_0, and thus risk-premium associated with λ:

$$\epsilon_\lambda + \nu_\lambda - \nu_0.$$

The $\lambda = 0$ return turns out to be the maximal growth return of Bansal and Lehmann (1997). This follows from the work of Alvarez and Jermann (2001) and Hansen and Scheinkman (2003). Alvarez and Jermann (2001) study the holding period returns to long-horizon discount bonds and show that in the limit these holding period returns approximate the maximal growth return of Bansal and Lehmann (1997). Hansen and Scheinkman (2003) show that this limiting return is the return on the dominant eigenfunction for the ($\lambda = 0$) pricing operator. As a consequence it is approximately the long-horizon return on any security with a terminal payoff of the form $\psi(x_{t+j})$ not just a discount bond.

We compute long-run average excess rates of return $\epsilon_\lambda + \nu_\lambda - \nu_0$ implied by the asset pricing model in example 2.1. We also report the dividend growth rate ϵ_λ implied by the model. The calculation is performed using the formulas described in the example that follows. We used the consumption evolution estimated by fitting an AR(4) model for consumption growth evolution to the times series data. For the purposes of this calculation, we presumed only a single consumption shock. The results are reported in figure 7.

For a fixed value of ρ and θ, the curves are upward sloping. This is to be expected because increasing λ increases the long-run covariation with consumption. Increasing risk aversion by increasing θ, makes the curves steeper as expected. Large assumed values of θ can produce substantial differences in predicted returns if there is substantial heterogeneity in λ.

The asymptotic discount rate for securities with growth attributes characterized by (λ, ζ) is $\nu_\lambda - \zeta$. It requires estimates of λ and ζ. We turn to this measurement problem in the next section.

Example 4.2. *Consider the first-order autoregressive specification in example 2.1:*

$$\begin{aligned} x_{t+1} &= Gx_t + Hw_{t+1} \\ c_{t+1} - c_t &= U_c \cdot x_t + \mu_c. \end{aligned}$$

Write:

$$s_{t+1,t} + \lambda\left(c_{t+1} - c_t\right) = -\frac{1}{2}(x_{t+1})'\Xi_1 x_{t+1} - \xi_1 \cdot x_{t+1} - \frac{1}{2}(x_t)'\Xi_2 x_t - \xi_2 \cdot x_t - \xi_0$$

We seek an eigenfunction that is log-quadratic:

$$\log \psi(x) = -\frac{1}{2}x'\Omega x - \omega \cdot x.$$

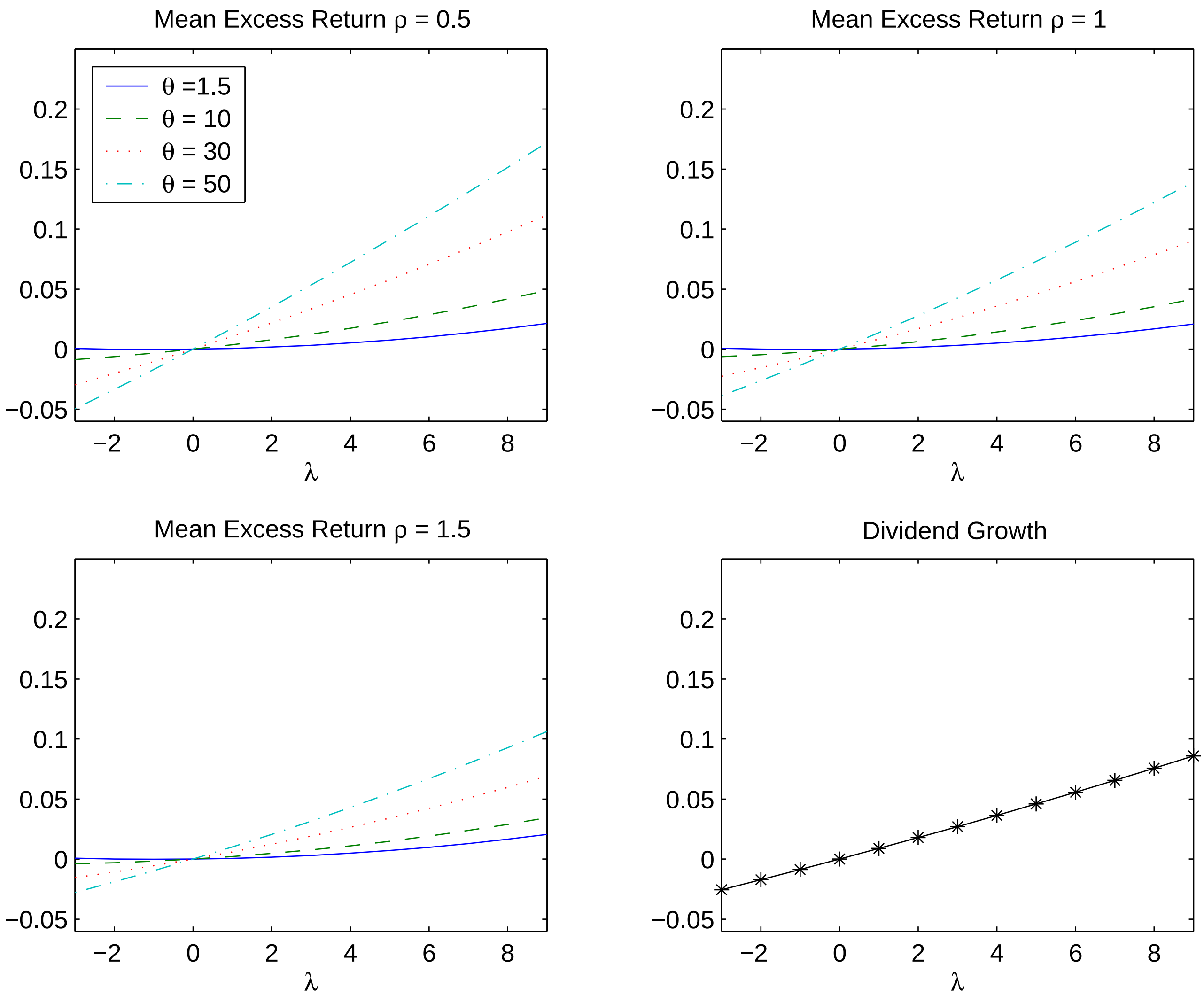

Figure 7: *Model implications for long-run return and dividend growth* The plots labelled *Mean Excess Return* contain curves of the form $\epsilon_\lambda + \nu_\lambda - \nu_0$ as a function of λ. The curves are computed using a subjective discount factor of .95 and they are indexed by the parameter pair (θ, ρ). The plot labelled *Dividend Growth* depicts the curve ϵ_λ as a function of λ.

When $\rho = 1$, the matrices Ξ_1 and Ξ_2 are all zero and the eigenfunction will be log-linear ($\Omega = 0$).

Let $\hat{\Omega} = \Omega + \Xi_1$. Then $\hat{\Omega}$ satisfies the Riccati equation:

$$\Xi_1 + \Xi_2 + G'\hat{\Omega}G - G'\hat{\Omega}H(I + G'\hat{\Omega}G)^{-1}H'\hat{\Omega}G = \hat{\Omega}$$

which is easily solved.

Consider next the equation for ω. Let $\hat{\omega} = \xi_1 + \omega$ and

$$G^* = G - H(I + H'\hat{\Omega}H)^{-1}H'\hat{\Omega}G.$$

Then

$$(G^*)'\hat{\omega} + \xi_1 + \xi_2 = \hat{\omega}$$

implying that

$$\hat{\omega} = (I - (G^*)')^{-1}(\xi_1 + \xi_2).$$

Finally the equation for eigenvalue is given by

$$\nu_\lambda = -\frac{1}{2}\hat{\omega}'H(I + H'\hat{\Omega})^{-1}H'\hat{\omega} + \xi_0 + \frac{1}{2}\log\det\left(I + H'\hat{\Omega}H\right)$$

For this example it is also straightforward to compute the decomposition of the price-dividend ratio (13) provided that the dividend process is can be expressed as:

$$d_t - \lambda c_t - \zeta t = -x_t\Phi_d x_t - \phi_d \cdot x_t - \mu_d.$$

This entails iterating on the Riccati equation for matrices in the quadratic forms, along with the coefficients of the linear and constant terms.

5 Statistical Evidence for Cointegration

In the previous section we showed how the long-run relation between dividends and consumption is reflected in the decomposition of the dividend-price ratios by horizon. We described methods for deducing when the long-run properties of cash flows have an important contribution to the current values of assets that are claims to those cash flows. To apply this apparatus requires knowledge of the joint process of consumption and cash flows or dividends. We now consider statistical evidence for a long-run link between consumption and book-to-market sorted portfolio cash flows using the VAR model (6).

5.1 Bivariate evidence

We study bivariate models in which:

$$y_t \doteq \begin{bmatrix} c_t \\ d_t^j \end{bmatrix}.$$

We explore the cointegrating restriction $A(1) = \alpha F'$ where $F' = \begin{bmatrix} -\lambda & 1 \end{bmatrix}$.

In Figure 8 we depict twice the log-likelihood ratio of the unconstrained model against various restricted models with the lag $\ell = 5$. Small likelihood ratios are indicative of little evidence against the restrictions. In all cases we explore the cointegrating restrictions. The vertical axis in the plots gives twice the difference in the log-likelihood relative to the VAR of (6) where cointegration is not imposed. The likelihood functions are all constructed conditioned on the first five observations, as is typical in the VAR literature. The horizonal axis explores changes in the cointegrating vector parameterized as $F' = \begin{bmatrix} -\lambda & 1 \end{bmatrix}$. Since the figures present twice the log likelihood ratio of the unrestricted versus restricted models, the Maximum Likelihood Estimator (MLE) of λ is found at the minimum of each line.

Two different assumptions about the deterministic time trend are depicted. *No time trends* imposes the restriction that $B_1 = 0$. *Time trends* allows for a time trend in both consumption and dividends in addition to a single unit root. The minimum value of the curves in each case can be compared to a chi-square distribution with one degree of freedom to conduct classical inference regarding the cointegration restriction. Alternative inference can be conducted using these values as well. For example, under the Bayesian Information Criterion of Schwarz (1978) the minimized value can be compared to $\log(T) = 5.4$. Under the Akaike criterion the minimized value can be compared to 2.

When time trends are allowed, the cointegration restriction seems reasonable. If the MLE is used as an estimate of long-run risk, then there is the potential for substantial differences in predicted returns based on the calculations depicted in figure 7. For example the MLE λ for portfolios 1 and 2 are negative and positive for portfolio 5. using these estimates, for large θ portfolio 1 is predicted to have a much lower long-run average return than portfolio 5.

The presence of a time trend is quite important in producing differences in long-run risk as measured by the parameter λ. For example in the absence of a trend, the parameter λ is very poorly identified for portfolios one, two and five as reflected by the flat log-likelihood function. When time trends are allowed the minimized log-likelihood ratio drops substantially for portfolios 1 and 2 but much less dramatically for the other portfolios. Under classical inference, the difference between the curves can be compared to a chi-square distribution with two degrees of freedom. Although there is substantial heterogeneity across portfolios in the MLE of λ for each portfolio, figure 8 indicates that this parameter is not well identified. For instance, there are large regions of λ for portfolios 2, 3 and 5 that satisfy the Akaike criterion. Thus there is substantial statistical uncertainty in the long-run risk exposure in each portfolio as measured via co-integration.

To better understand the importance of time trends, figure 9 plots the level of dividends for each of the portfolios along with fitted dividends from the VAR where the dividend shock (the second shock) has been set to zero. When a time trend is allowed, the consumption shock accounts for a much larger amount of the low frequency fluctuation in dividends. At low frequencies these portfolios have a negative relationship with consumption and hence are predicted to have a lower average tail returns for the implied cash flows.

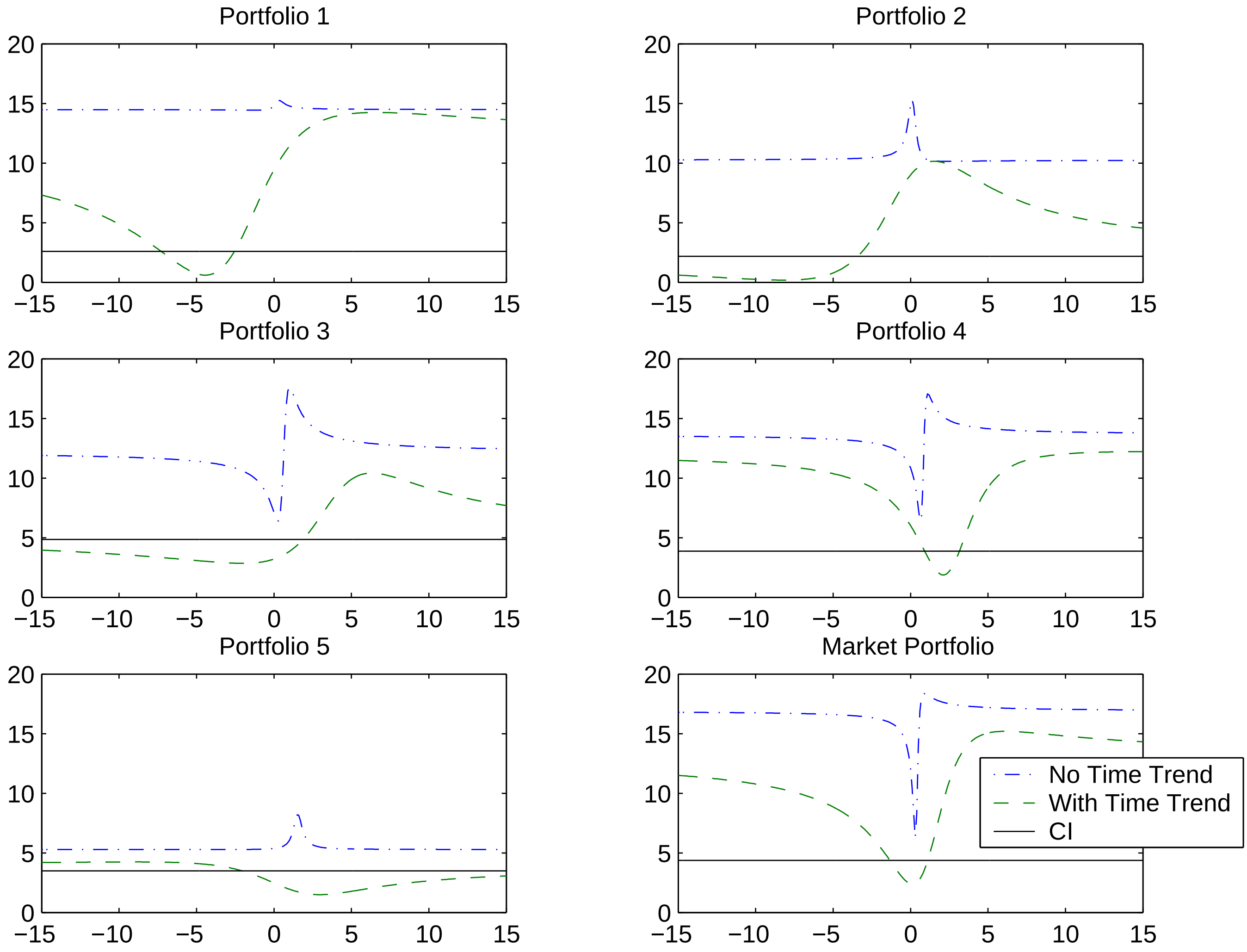

Figure 8: Twice the log-likelihood ratios relative to unconstrained model with $\ell = 5$ in the model (6). Different values of λ are plotted on the horizontal axes. The solid line "CI" gives the value of Akaike Criterion relative to the minimized value in the model with time trends.

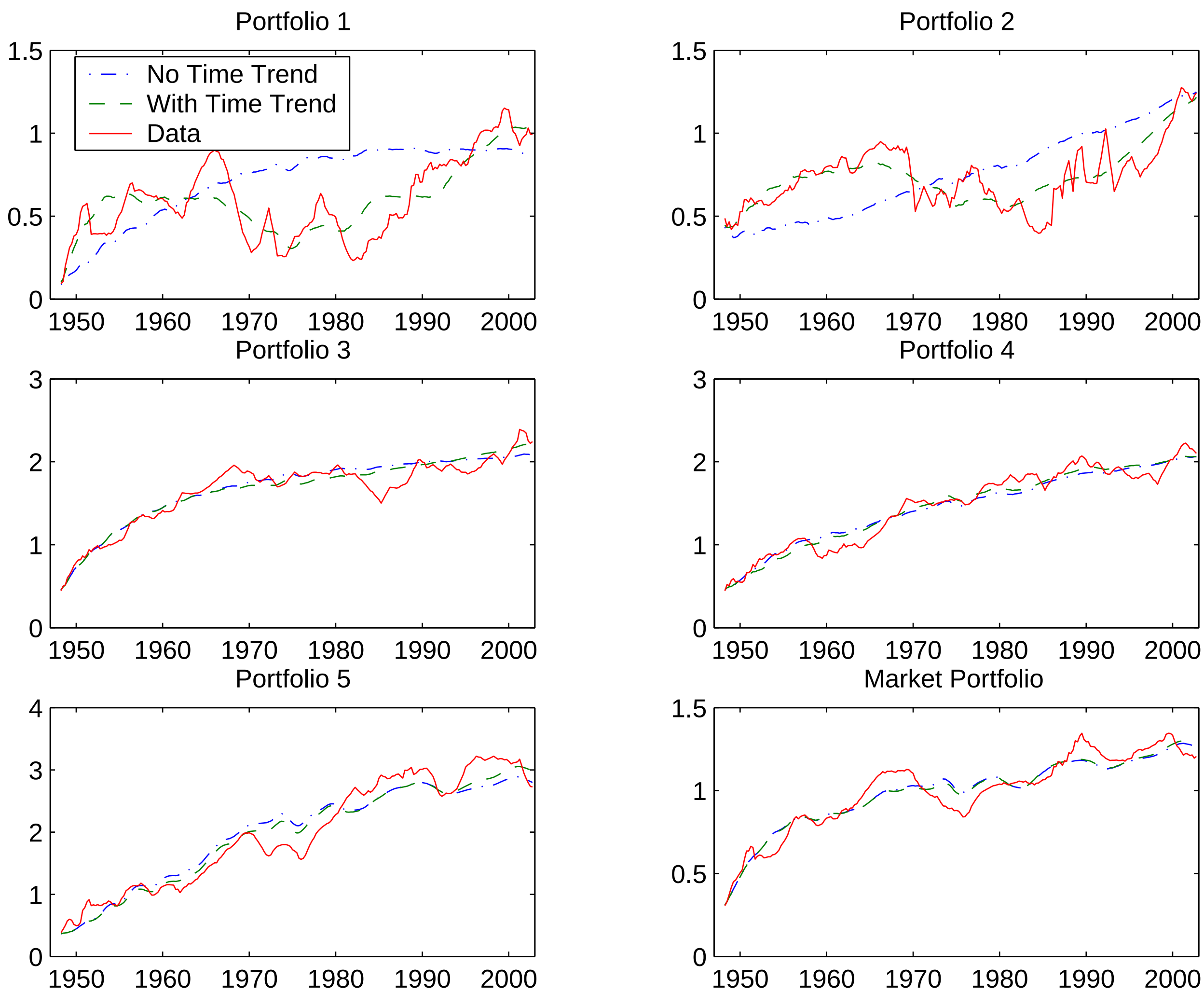

Figure 9: Portfolio dividends relative to fitted values based on consumption innovations alone.

5.2 Consumption, earnings and co-integration

We showed in section 3 that a model in which corporate earnings are assumed to be cointegrated with consumption captures an important low frequency component of consumption. Allowing for this low frequency component may affect the predicted long-run relationship between consumption and dividends. We therefore consider models where for each dividend series:

$$y_t^j \doteq \begin{bmatrix} c_t \\ e_t \\ d_t^j \end{bmatrix} .$$

We assume that consumption and corporate earnings are cointegrated and dividends are allowed to have a linear time trend. We impose the restriction that the joint dynamics for consumption and earnings are not Granger caused by dividends to prevent consumption and earnings from inheriting any dynamics from the trend in dividends which would then feedback onto dividends. In other words only a linear trend for dividends is allowed by this system. We continue to examine the cointegration restriction for each portfolio.

Figure 10 plots twice the log-likelihood ratio of an unrestricted model with a time trend relative to models where λ is restricted. The dashed-dotted line in each figure is for the case where the time trend is assumed to be zero. The dashed line depicts the results where a time trend in dividends is allowed. The horizontal solid line depicts the value of the Akaike Information Criterion of 2 which gives a probability value of 16% when there is one degree of freedom. As in figure 8 the presence of a trend is very important for portfolios 1, 2 and 4. Also consistent with the models with more flexible trends, the MLE estimates of λ are negative for portfolios 1 and 2. These portfolios appear to provide long-run consumption insurance.

Figure 11 plots both the level of dividends and the fitted values implied by the "aggregate" innovations to consumption and corporate earnings alone. As we noted before the presence of a deterministic trend allows the model to fit the low frequency movements of dividends for portfolios 1 and 2 much better. This low frequency effect is reflected in figure 12 which presents the spectral densities of dividend growth implied by the VAR model with time trends. The dashed lines give the spectral densities implied by all of the shocks and the solid lines give the spectral densities implied by the aggregate innovations alone. For comparison figure 13 presents the analogous figure for the case where the models are restricted to have no time trend. When time trends are included, the aggregate shocks do a much better job of matching the low frequency dynamics of dividend growth for portfolios 1, 2 and 4. In the presence of a time trend, the cointegrating relationship between consumption and dividends better captures the low frequency movements in dividends beyond 16 years.

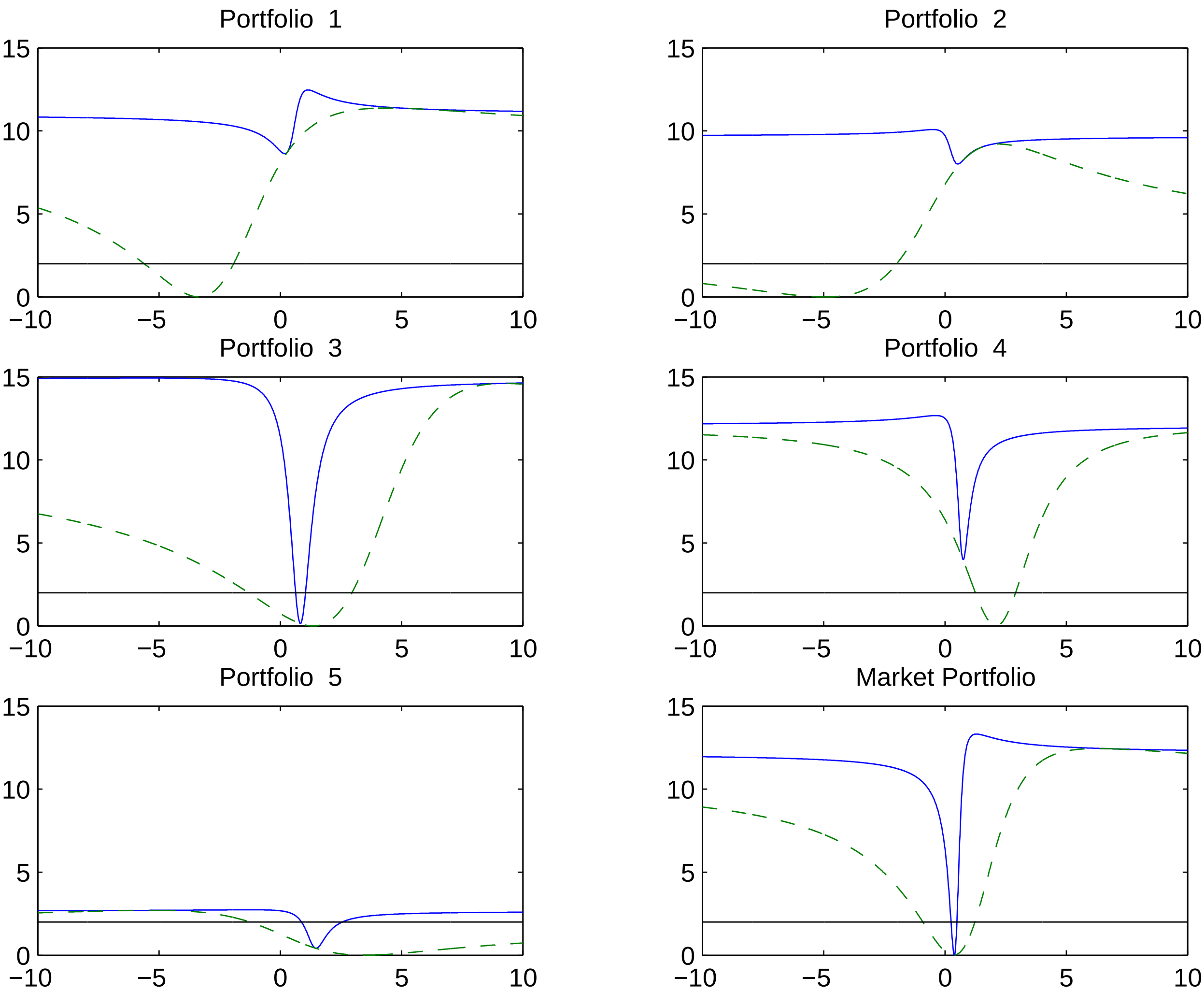

Figure 10: Twice the log-likelihood ratios relative to a model with only Granger Causality restrictions imposed. VAR with consumption, earnings and dividends. Dividends are restricted to not Granger cause consumption and earnings. Consumption and Earnings are assumed to be cointegrated.

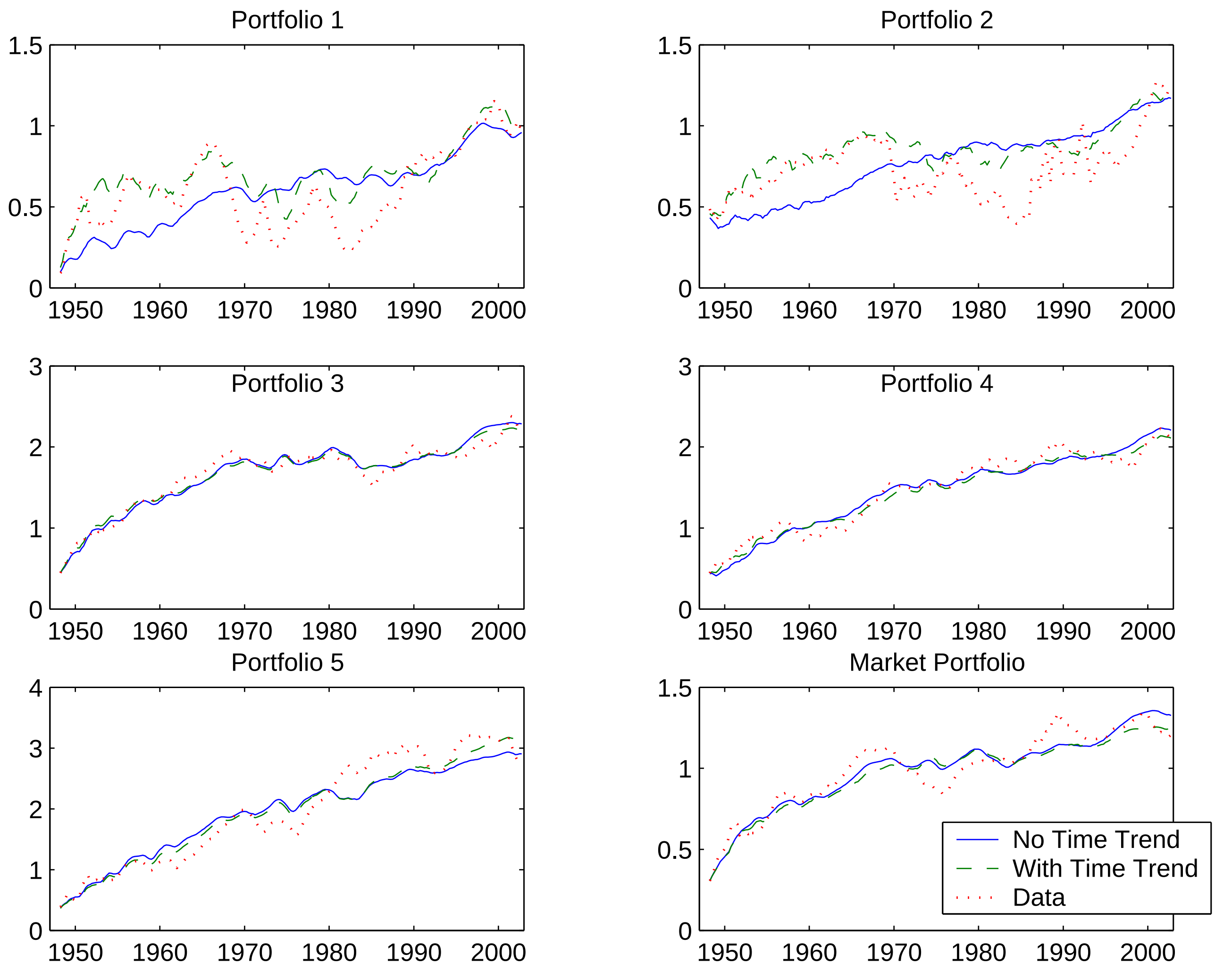

Figure 11: Portfolio dividends relative to fitted values based on aggregate innovations alone.

The point estimates of λ for portfolios 1 and 2 imply low frequency consumption insurance and hence low returns for these portfolios. In contrast the MLE estimates of λ for portfolios 4 and 5 are positive. As demonstrated by figure 7 these point estimates imply large differences in the implied tail returns of the cash flows.

As we found previously, there is substantial statistical uncertainty in the measurement of λ for several of the portfolios. To demonstrate these effects consider again the solid line in figure 10 that plots the value of the Akaike Information Criterion of 2. We build confidence sets for λ by finding those values that result in movements in twice the log-likelihood ratio below this cutoff. Confidence intervals for λ are reported in 2. Intervals are reported for confidence levels of 84% and 98% implied by the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) respectively. For the models where a time trend is allowed, the implied values of the trend coefficient ζ from the concentrated likelihood are also reported in the table. The intervals for ζ are numerically ordered but are "flipped" relative to the variation in λ. For example for portfolio 1 the confidence interval for λ is $(-5.6, -1.9)$. The corresponding value for ζ when $\lambda = -5.6$ is 0.052 and it is 0.021 when $\lambda = 1.9$. We report the resulting interval as $(0.021, 0.052)$. A model with $\lambda = 1$ would be selected by the BIC criterion for portfolios 3, 4 and 5 over a cointegrated model with an unknown coefficient. On the other hand, BIC would exclude a model with $\lambda = 1$ for portfolios 1 and 2.

5.3 What are the time trends?

Up until now, we have taken the linear cointegration model literally, As we have seen, the introduction of time trends into this model changes substantially statistical measures of long-run risk for portfolios 1 and 2. Is it realistic to think of these as deterministic time trends in studying the economic components of long-run risk? We suspect not. While there may be important components to the cash flows for portfolios 1 and 2 that are very persistent, it seems unlikely that these are literally deterministic time trends known to investors. Within the statistical model, the time trends for these portfolios in part offset the negative growth induced by the cointegration. We suspect that the substantially negative estimates of λ probably are not likely to be the true limiting measures of how dividends respond to consumption and earnings shocks. While the long-run risk associated with portfolios 1 and 2 looks very different from that of portfolio five, a literal interpretation of the resulting co-integrating relation is hard to defend.

There is a potential pitfall in using maximum likelihood methods conditioned on initial data points as we have here. Sims (1991) and Sims (1996) warn against the use of such methods because the resulting estimates might imply that

> ... the first part of the sample behavior of the data is dominated by a large "transient". That is, the estimates imply that the initial data points are very far from the deterministic trend line or steady state, in the sense that the estimated model implies that future deviations as great as the initial deviation will be extremely rare.

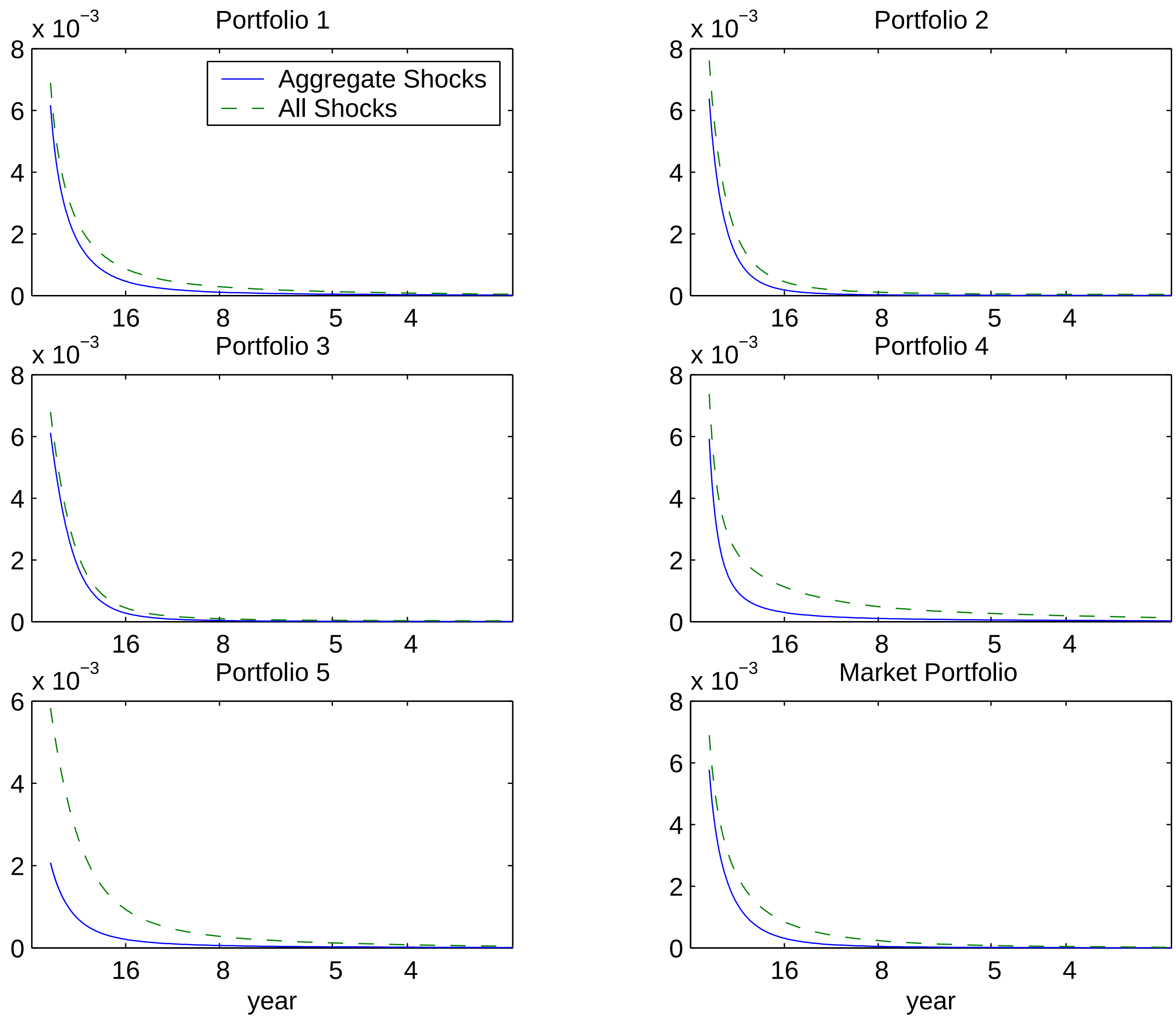

Figure 12: Implied spectral density for dividend growth from VAR's with consumption and aggregate earnings and a time trend in dividends. Densities are normalized by the standard deviation of dividend growth implied by the model. The solid lines give the spectral densities implied by the aggregate shocks to consumption and corporate earnings alone. The dashed lines give the implied spectral densities from the complete model. The x-axes give the number of years that correspond to each frequency.

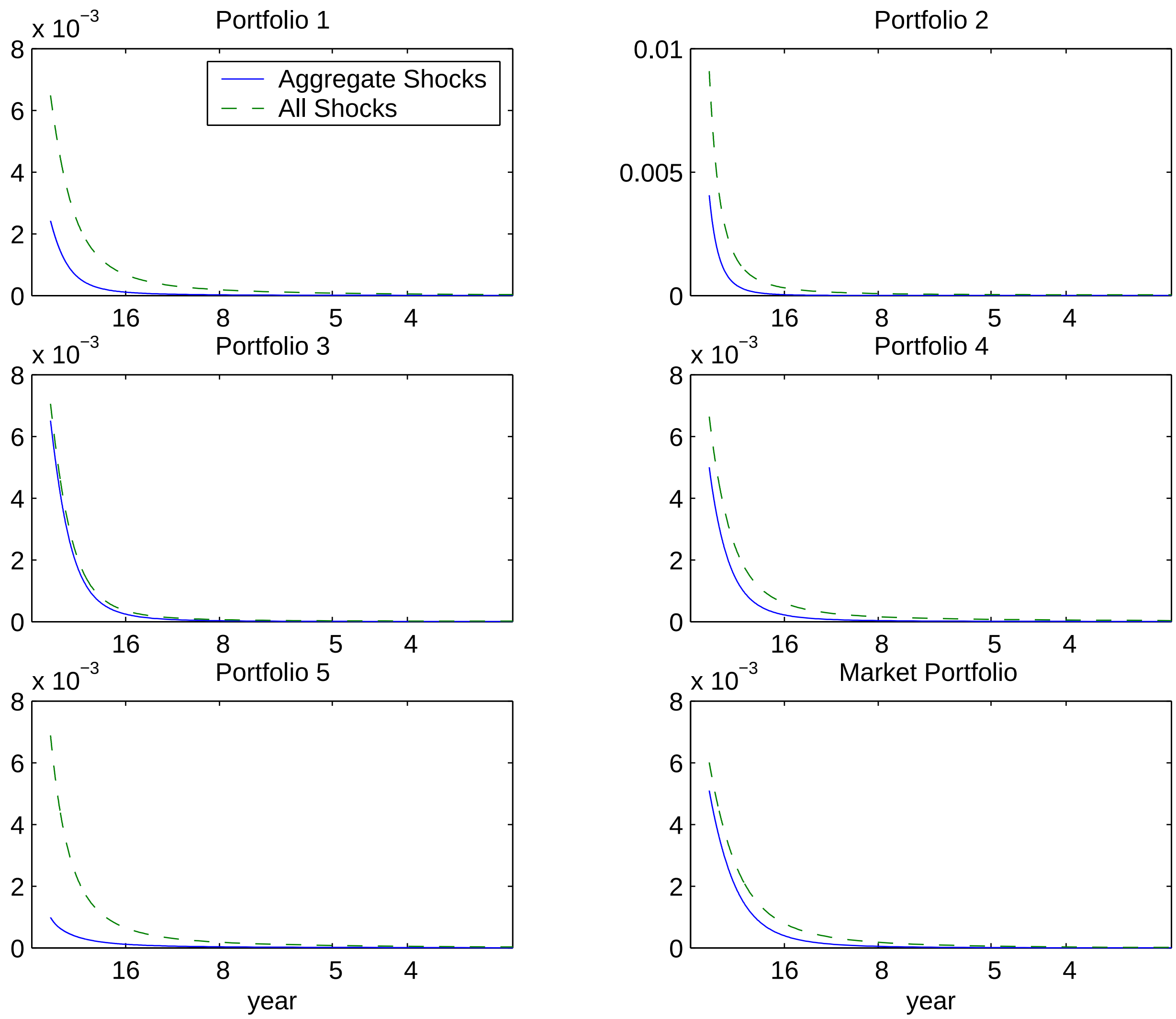

Figure 13: Implied spectral density for dividend growth from VAR's with consumption and aggregate earnings and no time trend in dividends. Densities are normalized by the standard deviation of dividend growth implied by the model. The solid lines give the spectral densities implied by the aggregate shocks to consumption and corporate earnings alone. The dashed lines give the implied spectral densities from the complete model. The x-axes give the number of years that correspond to each frequency.

Table 2: Confidence intervals for λ and implied time trends

	Portfolio					
	1	2	3	4	5	Market
			No Time Trend, AIC			
λ	(-4.2, 0.6)	(-10^*, 10^*)	(0.6, 1.1)	(0.6, 1.0)	(0.6, 4.1)	(0.3, 0.5)
			No Time Trend, BIC			
λ	(-10^*, 10^*)	(-10^*, 10^*)	(0.5, 1.3)	(0.5, 1.4)	(-10^*, 10^*)	(0.1, 0.6)
			Time Trend, AIC			
λ	(-5.6, -1.9)	(-10^*, -2.0)	(-1.3, 3.0)	(1.3, 2.9)	(-1.3, 10)	(-0.9, 1.2)
ζ	(.021, .052)	(.022, .093)	(-.019, .019)	(-.019, -.005)	(-.073, .024)	(-.007, .011)
			Time Trend, BIC			
λ	(-10^*, -0.9)	(-10^*, -0.6)	(-6.1, 3.9)	(0.4, 3.7)	(-10^*, 10^*)	(-2.8, 1.9)
ζ	(.011, .088)	(.009, .093)	(-.027, .062)	(-.026, .003)	(-.073, .093)	(-.013, .027)

Confidence intervals for λ constructed from implied deterioration in the concentrated likelihood as a function of λ. For the intervals labelled "AIC" the maximal deterioration in the twice the log-likelihood ratio was 2 yielding a confidence level of 84%. For the intervals labelled "BIC" the maximal deteriorate was 5.4 yielding a confidence level of 98%. Variation in λ was restricted to the interval $(-10, 10)$. Confidence interval endpoints reaching these limits are denoted by *. The implied estimates of ζ from the concentrated likelihood are decreasing in λ.

This seems to be particularly true for the models we fit to portfolios 1 and 2. In figure 14 we report the time series trajectory implied by the initial conditions alone with time trends included. These trajectories are depicted by the dot-dashed lines and are generated using the conditional maximum likelihood estimates. The initial conditions appear to be far from the trend lines for portfolios 1 and 2, and as a consequence have a nontrivial trajectory. By conditioning, the maximum likelihood estimates allow for this feature of the model fit. Do we really believe that investors have confidence at the beginning of the sample in such a trajectory? We suspect not.

How do we fix this estimation problem? Sims (1996) suggests that the implied distributions for the initial contributions be included in the likelihood function construction. Including this initial term will penalize large deviations of the initial conditions from the implied stationary distribution. Since the cointegration model has a unit root by construction, this correction must necessarily focus on the linear combinations of initial y's that are implied to be stationary. We suspect that modifying the likelihood function in this manner will have an important impact on its overall shape and its implication for the cointegration between consumption and dividends. The implied role for consumption might well be diminished. Conditional likelihood estimation has the potential to mis-measure long-run risk, especially when time trends are introduced into the estimation.

6 Conclusion

Long run or growth rate variation in consumption or cash flows can have important consequences in asset valuation. Some recent time series evidence supports so called *consumption-based* models by appealing to long-run consumption risk.

Using statistical methods to measure directly long-run cash flow variation is a challenging endeavor, however. Statistical methods typically rely on extrapolating the time series model to infer how cash flows respond in the long-run to shocks. This extrapolation depends on details of the growth configuration of the model, and in many cases these details are defended primarily on statistical grounds. Moreover, the simple linear models we consider are likely to be misspecified. There is pervasive statistical evidence for growth rate changes or breaks in trend lines, but this statistical evidence is difficult to use directly in models of decision-making under uncertainty without some rather specific ancillary assumptions.

There are two complementary responses to this conundrum. One is to resort to the use of highly structured, but easily interpretable, models of long-run growth variation. The other is to exploit the fact that asset values encode information about long-run growth. To break this code requires a reliable economic model of the long-run risk-return relation. We suggest model-based methods for economic characterizations of this relation. These methods give us clues as to what types of models feature or amplify the role of long-run risk. Unfortunately, as yet there is not an empirically well grounded, and economically relevant model of asset pricing to use in deducing investors beliefs about the long-run from values of long-lived assets. Much progress has been made in our understanding of models, but less in understanding the precise nature of long-run growth rate risk in the underlying economy.

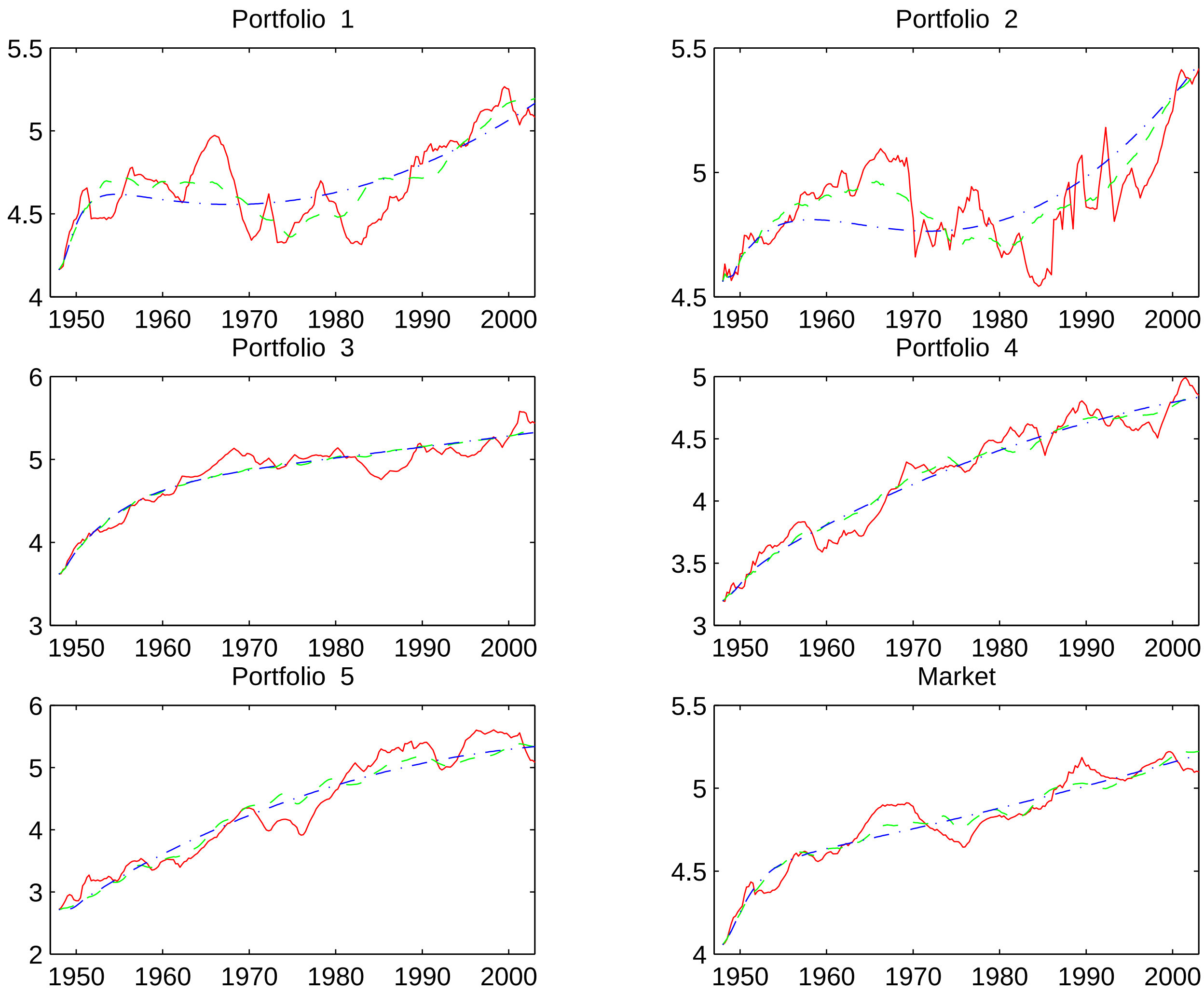

Figure 14: *Portfolio Dividends and Fitted Values.* Solid Lines —— display the data. Dashed lines − − − are the fitted values based on consumption shocks alone. Dot-dashed lines −·−· are fitted values with all shocks set to zero .

A Expansion

We compute the first-order expansion:

$$v_t \approx v_t^1 + (\rho - 1) D v_t^1$$

where v_t^1 is the continuation value for the case in which $\rho = 1$. We base this calculation on the approximate recursion:

$$v_t \approx \beta \left[\mathcal{Q}_t(v_{t+1} + c_{t+1} - c_t) + (1-\rho) \frac{\mathcal{Q}_t(v_{t+1} + c_{t+1} - c_t)^2}{2} \right].$$

Then

$$v_t^1 = \beta \mathcal{Q}_t(v_{t+1}^1 + c_{t+1} - c_t),$$

which is the $\rho = 1$ exact recursion and

$$\begin{aligned} D v_t^1 &= -\beta \frac{\mathcal{Q}_t(v_{t+1}^1 + c_{t+1} - c_t)^2}{2} + \beta \tilde{E}(D v_{t+1}^1 | \mathcal{F}_t) \\ &= -\frac{(v_t^1)^2}{2\beta} + \beta \tilde{E}(D v_{t+1}^1 | \mathcal{F}_t) \end{aligned} \tag{14}$$

where $\tilde{E}$ is the distorted expectation operator associated with the density

$$\frac{\left(V_{t+1}^1\right)^{1-\theta}}{E\left[\left(V_{t+1}^1\right)^{1-\theta} | \mathcal{F}_t\right]}.$$

Consider example 2.1. Then

$$(v_t^1)^2 = (x_t)' U_v (U_v)' x_t + 2 \mu_v U_v \cdot x_t + (\mu_v)^2.$$

Write:

$$D v_t^1 = (x_t)' \Upsilon_d x_t + U_d \cdot x_t + \mu_d.$$

From (14),

$$\begin{aligned} \Upsilon_d &= -\frac{1}{2\beta} U_v (U_v)' + \beta A' \Upsilon_d A \\ U_d &= -\frac{1}{\beta} \mu_v U_v + \beta(1-\theta) A' \Upsilon_d B \gamma(\beta) + \beta A' U_d \\ \mu_d &= -\frac{1}{2\beta} (\mu_v)^2 + \beta(1-\theta)^2 \gamma(\beta)' B B' \gamma(\beta) + \beta(1-\theta) U_d \cdot [B\gamma(\beta)] + \beta \mathrm{trace}(B' \Upsilon_d B) + \beta \mu_d \end{aligned} \tag{15}$$

The first equation in (15) is a Sylvester equation and is easily solved. Given Υ_d, the solution for U_d is:

$$U_d = (I - \beta A')^{-1} \left[\mu_v U_v + \beta(1-\theta) A' \Upsilon_d B \gamma(\beta) \right],$$

and given Υ_d and U_d the solution for μ_d is:

$$\mu_d = \frac{(\mu_v)^2 + \beta(1-\theta)^2\gamma(\beta)'BB'\gamma(\beta) + \beta(1-\theta)U_d \cdot [B\gamma(\beta)] + \beta \text{trace}(B'\Upsilon_d B)}{1-\beta}.$$

Finally, consider the first-order expansion of the logarithm of the stochastic discount factor:

$$s_{t+1,t} \approx s^1_{t+1,t} + (\rho - 1)Ds^1_{t+1,t}.$$

Recall that the log discount factor is given by:

$$\begin{aligned} s_{t+1,t} &= -\beta - \rho\,(c_{t+1} - c_t) + (\rho - \theta)\,[v_{t+1} + c_{t+1} - \mathcal{Q}_t(v_{t+1} + c_{t+1})] \\ &= -\beta - \rho\,(c_{t+1} - c_t) + (\rho - \theta)\,[v_{t+1} + c_{t+1} - c_t - \mathcal{Q}_t(v_{t+1} + c_{t+1} - c_t)] \end{aligned}$$

Differentiating with respect to ρ gives:

$$\begin{aligned} Ds^1_{t+1,t} &= -\,(c_{t+1} - c_t) + \left[v^1_{t+1} + c_{t+1} - c_t - \mathcal{Q}_t(v^1_{t+1} + c_{t+1} - c_t)\right] \\ &\qquad +(1-\theta)\left[Dv^1_{t+1} - \tilde{E}\left(Dv^1_{t+1}|\mathcal{F}_t\right)\right] \\ &= v^1_{t+1} - \tfrac{1}{\beta}v^1_t + (1-\theta)\left[Dv^1_{t+1} - \tilde{E}\left(Dv^1_{t+1}|\mathcal{F}_t\right)\right]. \end{aligned}$$

Note that

$$v^1_{t+1} - \mathcal{Q}_t\left(v^1_{t+1} + c_{t+1} - c_t\right) = U_v \cdot x_{t+1} - \frac{1}{\beta}U_v \cdot x_t + \left(1 - \frac{1}{\beta}\right)\mu_v$$

and

$$\begin{aligned} Dv^1_{t+1} - \tilde{E}\left(Dv^1_{t+1}|\mathcal{F}_t\right) &= \gamma(\beta)'B'\Upsilon_d B\gamma(\beta) + 2(Bw_{t+1})'\Upsilon_d\,[Ax_t + B\gamma(\beta)] \\ &+ U_d \cdot [B\gamma(\beta) + Bw_{t+1}]\,. \end{aligned}$$

References

Alvarez, F. and U. Jermann (2001). The size of the permanent component of asset pricing kernels. University of Chicago.

Anderson, E. W., L. P. Hansen, and T. J. Sargent (2003). A quartet of semi-groups for model specification, detection, robustness and the price of risk. *Journal of European Economic Association 1*, 66–123.

Bansal, R., R. Dittmar, and C. Lundblad (2002). Consumption, dividends, and the cross-section of equity returns. Duke University.

Bansal, R. and B. Lehmann (1997). Growth optimal portfolio restrictions on asset pricing models. *Macroeconomic Dynamics 1*, 1135–1171.

Bansal, R. and A. Yaron (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance forthcoming*.

Campbell, J. (1996). Understanding risk and return. *Journal of Political Economy 104*, 298–345.

Campbell, J. and J. Cochrane (1999). By force of habit. *Journal of Political Economy 107*, 205–251.

Campbell, J., C. Polk, and T. Vuolteenaho (2003). Growth or glamour? Harvard University.

Campbell, J. and R. Shiller (1988). The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies 1*, 195–227.

Campbell, J. and T. Vuolteenaho (2003). Bad beta, good beta. Harvard University.

Constantinides, G. (1990). Habit formation: A resolution of the equity premium puzzle. *Journal of Political Economy 104*, 519–543.

Epstein, L. and S. Zin (1989a). Substitution, risk aversion and the temporal behavior of stock returns: An empirical investigation. *Journal of Political Economy 99*, 263–286.

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

Fama, E. F. and K. R. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics 33*, 3–56.

Gabaix, X. and D. Laibson (2002). The 6d bias and the equity premium puzzle. In B. Bernanke and K. Rogoff (Eds.), *NBER Macroeconomics Annual*. MIT Press.

Gibbons, M. and W. Ferson (1985). Testing asset pricing models with changing expectations and an unobservable market portfolio. *Journal of Financial Economics 14*, 217–236.

Hansen, L. P., J. Heaton, and N. Li (2004). Intangible risk. University of Chicago.

Hansen, L. P. and S. J. Richard (1987). The role of conditioning information in deducing testable implications of asset pricing models. *Econometrica 55*(3), 587–613.

Hansen, L. P. and T. Sargent (1995, May). Discounted linear exponential quadratic gaussian control. *IEEE Transactions on Automatic Control 40*(5), 968–971.

Hansen, L. P. and J. Scheinkman (2003). Semigroup pricing. Princeton University.

Hansen, L. P. and K. Singleton (1983). Stochastic consumption, risk aversion, and the temporal behavior of asset returns. *Journal of Political Economy 91*, 249–265.

Heaton, J. (1995). An empirical investigation of asset pricing with temporally dependent preference specifications. *Econometrica 63*, 681–717.

Lettau, M. and S. Ludvigson (2001a). Consumption, aggregate wealth and expected stock returns. *Journal of Finance 56*, 815–849.

Lettau, M. and S. Ludvigson (2001b). Resurrecting the (c)capm: A cross-sectional test when risk premia are time-varying. *Journal of Political Economy 109*, 1238–1287.

Lucas, R. E. (1978). Asset prices in an exchange economy. *Econometrica*.

Lynch, A. (1996). Decision frequency and syncrhonization across agents: Implications for aggregate consumption and equity returns. *Journal of Finance 51*, 1479–1497.

Menzly, L., T. Santos, and P. Veronesi (2004). Understanding predictability. *Journal of Political Economy 112*(1), 1–47.

Parker, J. A. and C. Julliard (2003). Consumption risk and cross-sectional returns. Princeton University.

Santos, T. and P. Veronesi (2001). Labor income and predictable stock returns. Graduate School of Busines, University of Chicago.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics 6*, 461–464.

Sims, C. A. (1991). *Journal of Applied Econometrics 6*, 423–434.

Sims, C. A. (1996). Inference for multivariate time series models with trend. Originally presented at the August 1992 ASSA Meetings.

Sundaresan, S. (1989). Intertemporally dependent preferences and volatility of consumption and wealth. *Review of Financial Studies 3*, 73–89.

Tallarini, T. (1998). Risk sensitive business cyles. *Journal of Monetary Economics 45*, 507–532.

Weil, P. (1990). Nonexpected utility in macroeconomics. *Quarterly Journal of Economics 105*, 29–42.