# Consumption Strikes Back?: Measuring Long-Run Risk *

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

First draft: May 2004
Current draft: June 2005

**Abstract**

We characterize and measure a long-run risk return tradeoff for the valuation of financial cash flows that are exposed to fluctuations in macroeconomic growth. This tradeoff features cash flow components that are realized far into the future but are still reflected in current asset values. We use the recursive utility model with empirical inputs from vector autoregressions to quantify this tradeoff; and we study the long-run risk differences in aggregate securities and in portfolios constructed based on the ratio of book equity to market equity. We isolate features of the economic model needed for the long run valuation differences among these portfolios to be sizable. Finally, we show how the resulting measurements vary when we consider alternative statistical specifications of cash flow and consumption growth.

*We thank Fernando Alvarez, Ravi Bansal, John Cochrane, Ken Judd, Jonathan Parker, Tano Santos, Tom Sargent, Chris Sims, Pietro Veronesi and Kenji Wada for valuable comments. We also thank workshop participants at Chicago, Columbia, Duke, Emory, INSEAD, LBS, MIT, Montreal, NBER, NYU, and Northwestern. 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

§School of Business, National University of Singapore

# 1 Introduction

Applied time series analysts have studied extensively how macroeconomic aggregates respond in the long run to underlying economic shocks. For instance, Cochrane (1988) used time series methods to measure the importance of permanent shocks to output and discussed the associated measurement challenges. Blanchard and Quah (1989) advocated using restrictions on long run responses to identify economic shocks and measure their importance. Many researchers have subsequently applied the Blanchard-Quah approach to the study of macroeconomic time series. This paper develops and applies methods for integrating asset valuation of cash flows with stochastic growth into this analysis. The unit root contributions measured by macroeconomists are a source of long-run risk that should be reflected in the valuation of cash flows. To quantify the role of long-run risk, we use tools that are complementary to the methods developed by Campbell and Shiller (1988). Our analysis is motivated in part by other recent research seeking to construct *cash flow betas*, (*e.g*, see Bansal, Dittmar, and Lundblad (2005)), but our interpretation and justification for such objects is novel.

Valuation of cash flows reflect expected growth, discounting and riskiness. In Markov economies with stationary growth, the value of cash flows in the distant future declines as the horizon increases at a rate that is approximately constant. When this decay rate is small, future cash flows have a durable contribution to current values.

As is known from the Gordon growth model, a small decay rate in the contribution to value reflects in part cash flow growth. When cash flows grow relatively quickly, their contribution to value is more persistent. Since dividend growth rates projected far into the future are approximately constant, there is a well defined adjustment for cash flow growth. By adding the dividend growth rate to the value decay rate, we extract a risk-adjusted discount rate. This rate is applicable to the components of the cash flow growth process that are realized in the distant future. The risk adjustment comes from two sources. One is the direct random fluctuation in the growth rates of the cash flow, and the other is the riskiness that is imputed by the valuation of this cash flow. Our paper focuses on the characterization and measurement of this long-run risk relation.

We characterize long-run cash flow risk by exploiting a mathematical formulation of asset valuation developed in Hansen and Scheinkman (2005). This method computes a long-run dominant pricing component for cash flows that grow stochastically. A family of valuation operators indexed by the payoff horizon share common factors called eigenfunctions. One of these eigenfunctions is dominant in the long run in a well defined sense. It isolates value movements due to cash-flow riskiness far into the future. By applying this apparatus, we quantify what model ingredients have important influences on the valuation of growth components.

Value decompositions of the type just described require a specific economic model and empirical inputs to characterize the growth and riskiness of cash flows. The calculations in this paper are based on a well specified, albeit highly stylized, model. Following Epstein and Zin (1989b), Weil (1990), Tallarini (1998), Bansal and Yaron (2004) and many others, we use a recursive utility framework of Kreps and Porteus (1978). For these preferences, the intertemporal composition of risk matters to the decision maker. Risk cannot simply be reduced or averaged out. Instead, the timing of when information is revealed about intertemporal consumption lotteries matters in the implied preference ordering. As emphasized by Epstein and Zin (1989b), these preferences also offer a convenient and appealing way to break the preference link between risk aversion and intertemporal substitution. Bansal and Yaron (2004) showed that predictable components in consumption growth can amplify the risk premia in security market prices. To study how long-run risk depends on intertemporal substitution and on risk aversion and predictable components to consumption growth, we extend an analytical approach suggested by Kogan and Uppal (2001). Formally, we expand the equilibrium pricing relation around an economy in which investor preferences have a unitary elasticity of intertemporal substitution (EIS).

While we focus on a recursive utility specification, the intertemporal timing of risk matters in other models as well, including models that feature habit persistence (*e.g.* Constantinides (1990), Heaton (1995), and Sundaresan (1989)) and models of staggered decision-making (*e.g.* see Lynch (1996) and Gabaix and Laibson (2002).) The approach we adopt in this paper can be extended to apply to these models as well.

In addition to an economic model, our value decompositions require statistical inputs that quantify long-run stochastic growth in macroeconomic variables, particularly in consumption. The decompositions also require knowledge of the long-run link between stochastic cash flows and the macroeconomic risk variables. These components of financial risk cannot be fully diversified and hence require nontrivial risk adjustments. The long-run nature of these risks adds to the statistical challenges just as it does in the related macroeconomic literature.

In this paper like many others, we study the intertemporal composition of risk using (log) linear vector autoregressive (VAR) models of consumption and cash flows. These models are designed to accommodate transient dynamics in a flexible way. They are convenient time series models that allow us to explore the statistical accuracy of the risk measurements along with the sensitivity of these measurements to changes in the model specification. Our focus on long-run risk deliberately stretches the VAR methods beyond their ability to capture transient dynamics. As a consequence, this paper explores 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? In addition to providing a long-run valuation counterpart to the familiar risk-return tradeoff, this paper examines the sensitivity of the measurements to estimation and model uncertainty.

In section 2 we use a finite state Markov chain to illustrate our methods. In section 3 we use the recursive utility model to show why the intertemporal composition of risk might matter to an investor. We also develop a general approximation to the model's solution. In section 4 we identify important aggregate shocks that affect long-run consumption. Section 5 develops a notion of risk based on the low frequency properties of cash flows and consumption, and section 6 constructs the implied measures of the risk-return relation. Section 7 explores the valuation sensitivity of alternative specifications of the long-run statistical relationship between consumption and portfolio cash flows. Section 8 concludes.

## 2 Markov Chain Example

Prior to developing a model economy with empirical inputs, we illustrate our analytical approach when the dynamic evolution can be captured by a finite state Markov chain. This allows us to use the familiar theory of matrices to depict our results.

Suppose that the dynamics of cash flows and consumption are determined by an $N$ state Markov chain. State $n$ of this Markov chain is denoted $\mathsf{x}_n$, and the probabilities of transiting from one state to another are given by:

$$\mathsf{a}_{m,n} = \text{Prob}(x_{t+1} = \mathsf{x}_n | x_t = \mathsf{x}_m) \ .$$

We assume that the resulting probability matrix is irreducible. That is, for some integer $\tau$, the entries of $\mathcal{A}^\tau$ are strictly positive, where $\mathcal{A}$ is formed from the $\mathsf{a}_{m,n}$'s.

Each entry $\mathsf{a}_{m,n}$ is scaled by two objects. One is the stochastic discount factor between adjacent dates. Exploiting the Markov property, the discount factor for the next period state $\mathsf{x}_n$ conditioned on the current state being $\mathsf{x}_m$ is $\mathsf{s}_{m,n}$ and is assumed to be strictly positive.[1]

The second object in our scaling is a stochastic cash flow growth factor. This scaling reflects our interest in the effects of different specifications of long-run risk. This second factor can also be state dependent. Conditioned on being in state $\mathsf{x}_m$ in the current period, the next period growth factor is $\mathsf{d}_{n,m}$ and assumed to be strictly positive. This leads us to study a new matrix $\mathcal{P}$ with entries:

$$\mathsf{p}_{m,n} = \mathsf{a}_{m,n}\mathsf{s}_{m,n}\mathsf{d}_{m,n}. \tag{1}$$

For each row, the sums across columns (sums over n) are typically not unity. In other words, we cannot interpret each row of $\mathcal{P}$ as a vector of probabilities.

We are interested in the discounted cash flows:

$$\frac{P_t}{D_t} = E\left[\sum_{j=1}^{\infty}\left(\prod_{\tau=1}^{j} S_{t+\tau,t+\tau-1}\right)\frac{D_{t+j}}{D_t}|x_t\right] \tag{2}$$

where $\{D_{t+j} : j \geq 0\}$ is a stochastic cash flow process with price $P_t$ at date $t$ and $S_{t+1,t}$ is a stochastic discount factor process between date $t$ and date $t+1$. We use the matrix $\mathcal{P}$ to compute and decompose this valuation by the payoff horizon [index $j$ in formula (2). Suppose the cash flow is always positive and that its ratio $\frac{D_{t+1}}{D_t}$ depends only $x_t$ and $x_{t+1}$.

[1]This transformation of the probabilities is familiar from asset pricing where the "risk-neutral" distribution is obtained from the pricing model and the objective distribution. We do not, however, rescale the discount factors to behave as probabilities.

Given this limited dependence on the Markov state, we use this ratio to construct $\mathsf{d}_{m,n}$. As a consequence, term $j$ in the infinite sum on the right-hand side of (2) can be expressed as:

$$E\left[\left(\prod_{\tau=1}^{j} S_{t+\tau,t+\tau-1}\right)\frac{D_{t+j}}{D_t}|x_t = \mathsf{x}_m\right] = e_m(\mathcal{P})^j\mathbf{1}_N$$

where $e_m$ is a row vector of zeros with a one in the $m^{th}$ column and $\mathbf{1}_N$ is an $N$-dimensional column vector of ones.

The long-run characteristics of cash flows are encoded in the matrix $\mathcal{P}$, and the contributions of these characteristics to value are determined by the properties of this matrix raised to the power $j$. There are temporary components to cash flows as well that, for valuation purposes, may be dominated by the long-run characteristics of cash flows. These latter characteristics are determined by the behavior of $\mathcal{P}^j$ as $j$ gets large and hence can be evaluated by examining the eigenvalues of $\mathcal{P}$.

Raising a matrix to a power preserves the eigenvectors. Eigenvalues are altered but in a straightforward way. The original eigenvalues are raised to the same power as the matrix. Since the entries of the matrix $\mathcal{P}$ raised to some power are strictly positive, there is *principal* eigenvalue that is positive and a corresponding eigenvector with positive entries.[2] The principal eigenvalue has the largest magnitude among all eigenvalues of $\mathcal{P}$, and as a consequence it dominates in the long run.

The principal eigenvector, $f^*$, solves:

$$-\nu = \max_{f\geq 0}\min_m \log(e_m\mathcal{P}f) - \log(e_m f)$$

where $f$ is an $N$ dimensional column vector. The objective is the logarithm of the ratio of payoff value to payoff where $f$ is the vector of payoffs in each of the $N$ Markov states. The objective is the one-period counterpart to the logarithm of the price-dividend ratio. The objective is to guarantee that this ratio is large in all states. It is a standard result (and straightforward to show) that the solution to this problem equalizes the objective:

$$\log(u_m\mathcal{P}f) - \log(u_m f)$$

across the alternative $j$ states. As a consequence, this solution $f^*$ is also an eigenvector with positive entries associated with a positive eigenvalue:

$$Pf^* = \exp(-\nu)f^*.$$

where $-\nu$ is the optimized objective.

The positive eigenvalue $\exp(-\nu)$ strictly dominates all others. In particular, let $g^*$ be the row eigenvector associated with the eigenvalue $\exp(-\nu)$:

$$g^*\mathcal{P} = \exp(-\nu)g^*.$$

[2] This is known from the Perron-Frobenius theory of matrices.

For any positive $f$

$$\lim_{\tau \to \infty} \exp(\nu \tau) (\mathcal{P})^{\tau} f = \frac{(g^* f)}{(g^* f^*)} f^*.$$

Thus as the valuation horizon gets large, the vector of values are approximately proportional to $f^*$, provided of course that $(g^* f)$ is not zero. Moreover, when $f$ has strictly positive entries,

$$\lim_{\tau \to \infty} -\frac{1}{\tau} [\log (\mathcal{P})^{\tau} f] = \nu \mathbf{1}_N.$$

Thus $\nu$ is the asymptotic decay rate in valuation, and the eigenvector $f^*$ gives the limiting distribution of values across states.

From the Gordon growth model, we know that the decay rate $\nu$ is influenced by two factors, asymptotic (risk adjusted) discount rates and asymptotic growth rates in cash flows. When cash flows grow faster values decay slower. Thus to produce a risk adjusted discount rate, we need to adjust $\nu$ for dividend growth. To measure this, we form a matrix $\mathcal{G}$ with entries $\mathsf{a}_{mn}\mathsf{d}_{mn}$. The asymptotic cash flow growth rate is the logarithm $\epsilon$ of the dominant eigenvalue of this matrix, and the implied discount rate is $\epsilon + \nu$. This discount rate includes of an adjustment for long-run risk.

To compute asymptotic discount rates, as we saw in formula (1) we must specify three objects: a) the transition probabilities $\mathsf{a}_{mn}$, b) the stochastic discount factors $\mathsf{s}_{mn}$ implied by an economic model, and c) the growth factors $\mathsf{d}_{mn}$ for the cash flow growth. In the next section, section 3, we describe a parameterized class of economic models for the stochastic discount factor.

To characterize the implications for long-run risk, we are interested in how value decay rates and discount rates change as we alter the long-run risk exposure. As we change the specification $\mathsf{d}_{mn}$, we alter the implied discount rate giving rise to a long-run risk return relation. For instance, consider a specification:

$$\mathsf{d}_{mn} = \begin{cases} \tilde{\mathsf{d}} & \text{if } \ \mathrm{n} = \ell \\ 1 & \text{otherwise} \end{cases}.$$

This corresponds to a stochastic growth specification that features Markov state $\ell$. Growth at a rate $\log \tilde{\mathsf{d}}$ only occurs when state $\ell$ is realized, otherwise there is no growth. By changing $\ell$ and possibly $\tilde{\mathsf{d}}$, we may characterize a long-run risk return relation. For instance, this allows us to quantify the discount rate differences between low growth and high growth states of the Markov chain.

In section 6 we use familiar securities from financial economics as cash flows in our characterization long-run risk. The pertinent components of cash flows that determine the long-run returns are martingale components to these cash flows. Valuation is dictated by the importance of the macroeconomic shocks with long-run consequences to the martingale component of cash flows.

In the remainder of the paper we will use linear Markov processes instead of Markov chains. We do this so that we can explore temporal dependence in a more flexible manner. To support this application, we extend the approach just described by replacing matrices with operators that integrate over continuous states. This extension is given in section 5.

# 3 Asset Pricing

Models of asset pricing link investor preferences and opportunities to deduce equilibrium relations for returns and prices. These models explain return heterogeneity by the existence of risk premia. Investors require larger expected returns as compensation for holding riskier portfolios. Alternative asset pricing models imply alternative risk-return tradeoffs. Equivalently [*e.g*, see Hansen and Richard (1987)] they imply an explicit model of stochastic discount factors, the market determined 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.

## 3.1 Preferences

We follow Epstein and Zin (1989b) and Weil (1990) 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 it provides a convenient separation between risk aversion and the elasticity of intertemporal substitution [see Epstein and Zin (1989b)]. This separation allows us to examine the effects of changing risk exposure with modest consequences for the risk-free rate. Many of the measurement challenges that emerge in this economic model carry over to others as well, including any model that features the intertemporal composition of risk, including models in which investor preferences display intertemporal complementarity or "habit persistence."

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

$$V_t = \left[(1-\beta)\left(C_t\right)^{1-\rho} + \beta \mathcal{R}_t(V_{t+1})^{1-\rho}\right]^{\frac{1}{1-\rho}}. \tag{3}$$

The random variable $V_{t+1}$ is the continuation value of a consumption plan from time $t+1$ forward. The recursion incorporates the current period consumption $C_t$ and makes a risk adjustment $\mathcal{R}_t(V_{t+1})$ to the date $t+1$ continuation value. We use a CES specification for this risk adjustment as well:

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

where $\mathcal{F}_t$ is the current period information set. The outcome of the recursion is to assign a continuation value $V_t$ at date $t$.

When there is perfect certainty, the value of $1/\rho$ determines the *elasticity of intertemporal substitution* (EIS). A measure of risk aversion depends on the details of the gamble being considered. As featured by Kreps and Porteus (1978), preferences like these relax the restriction that intertemporal compound lotteries can be *reduced* by simply integrating out the uncertainty conditioned on current information. Instead the intertemporal composition of risk matters. As we will see, this will be reflected explicitly in the equilibrium asset prices that we characterize. On the other hand, the aversion to simple wealth gambles is given by $\theta$.

Under a Cobb-Douglas specification ($\rho = 1$), recursion (3) becomes:

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

In what follows, the case of $\rho = 1$ will receive special attention because of its analytical tractability.

To include stochastic growth in consumption we study an alternative recursion that scales continuation values by consumption:

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

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 (3) as

$$v_t = \frac{1}{1-\rho} \log\left((1-\beta) + \beta \exp\left[(1-\rho)\, \mathcal{Q}_t(v_{t+1} + c_{t+1} - c_t)\right]\right), \tag{4}$$

where $\mathcal{Q}_t$ is the so-called *risk-sensitive recursion*:[3]

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

The risk sensitive recursion is convenient for our subsequent characterizations.

## 3.2 Shadow Valuation

Consider the shadow valuation of a given consumption process. The utility recursion gives rise to a corresponding valuation recursion and implies stochastic discount factors used to represent this valuation. In light of the intertemporal budget constraint, the valuation of consumption in equilibrium coincides with wealth.

The first utility recursion (3) is homogeneous of degree one in consumption and the future continuation utility. Use Euler's Theorem to write:

$$V_t = (MC_t)C_t + E\left[(MV_{t+1})V_{t+1} | \mathcal{F}_t\right] \tag{5}$$

[3]See Hansen and Sargent (1995) and Tallarini (1998) and the relations they show to the risk sensitive control literature.

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 (5) 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 (5) 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. A claim to next period's consumption is valued using

$$S_{t+1,t} = \frac{MV_{t+1} MC_{t+1}}{MC_t} = \beta \left(\frac{C_{t+1}}{C_t}\right)^{-\rho} \left(\frac{V_{t+1}}{\mathcal{R}_t(V_{t+1})}\right)^{\rho-\theta} \tag{6}$$

as a stochastic discount factor. There are two (typically highly correlated) contributions to the stochastic discount factor in formula (6). One is the direct consumption growth contribution familiar from the Lucas (1978) and Breeden (1979) 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 $\rho$ and $\theta$ differ.

Given the homogeneity in the recursion used to depict preferences, equilibrium wealth is given by $W_t = \frac{V_t}{MC_t}$. Substituting for the marginal utility of consumption, the wealth-consumption ratio is:

$$\frac{W_t}{C_t} = \frac{1}{1-\beta} \left(\frac{V_t}{C_t}\right)^{1-\rho}.$$

Taking logarithms, we find that

$$\log W_t - \log C_t = -\log(1-\beta) + (1-\rho) v_t \tag{7}$$

When $\rho = 1$ we obtain the well known result that the wealth consumption ratio is constant.

A challenge in using this model empirically is to measure the continuation value, $V_{t+1}$, which is linked to future consumption via the recursion (3). 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 wealth. 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; or other components can be included such as human capital with assigned market or shadow values (see Campbell (1994)). In addition to requiring the use of a market measure of wealth, this approach precludes the special case in which $\rho = 1$. Since the consumption wealth ratio is constant when $\rho = 1$, we cannot infer the continuation value from wealth and consumption. Moreover, when $\rho$ is close to one any volatility in the stochastic discount factor attributed to wealth should also be reflected in consumption volatility. This implication is typically ignored even when consumption and wealth are used simultaneously.

In this investigation, like that of Restoy and Weil (1998), we maintain the direct link between the continuation value and the stochastic process governing future consumption. In the case of logarithmic intertemporal preferences ($\rho = 1$), the link between future consumption and the continuation value easily can be calculated as we demonstrate in the next section. It is well understood that $\rho = 1$ leads to substantial simplification in the equilibrium prices and returns (*e.g.* see Schroder and Skiadas (1999).)

Approximate characterization of equilibrium pricing for recursive utility have been produced by Campbell (1994) and Restoy and Weil (1998). In what follows we use a distinct but related approach. While Campbell (1994) and Restoy and Weil (1998) use log-linear approximation of budget constraints, we follow Kogan and Uppal (2001) by approximating around an explicit equilibrium computed when $\rho = 1$. Our approximation is in the parameter $\rho$.[4] Campbell and Viceira (2002) (chapter 5) show the close connection between approximation around the utility parameter $\rho = 1$ and approximation around a constant consumption-wealth ratio for portfolio problems.

Our application in what follows is to the study of a simple model of equilibrium price determination. We find some useful and intriguing contrasts between approximation methods. There are interesting conceptual differences in the implied one-period risk prices and the implied consumption-wealth ratios. Since our aim to study the implications for $\rho$ and $\theta$ for tail returns, we will compute $\rho$ derivatives of long-run risk for alternative values of $\theta$. Before turning to a discussion of the $\rho$ expansion, we consider the special case in which $\rho = 1$.

### 3.3 The special case in which $\rho = 1$

As in many papers in asset pricing, we use a $\rho = 1$ specification as a convenient benchmark. Campbell (1996) argues for less intertemporal substitution and Bansal and Yaron (2004) assume more. We will explore such deviations subsequently. The $\rho = 1$ is convenient for our purposes because when consumption has a log linear time series evolution, we can solve the for the continuation value. This feature gives us the flexibility to include important low frequency time series components in the model solution.

The $\rho = 1$ limit in recursion (4) for continuation values:

$$v_t = \beta \mathcal{Q}_t(v_{t+1} + c_{t+1} - c_t). \tag{8}$$

The stochastic discount factor in this special case is:

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

[4] Strictly speaking, $\rho = 1$ is ruled out in the parameterization considered by Restoy and Weil (1998) including the return-based Euler equation exploited in their calculations. The economy we study is different from that Kogan and Uppal (2001), but they suggest that extensions in the directions that interest us would be fruitful.

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. Recursion (8) was used by Tallarini (1998) in his study of *risk sensitive* business cycles.

To make our formula for the marginal rate of substitution operational, we need to compute $V_{t+1}$ 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 $\gamma_j$ 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 to be correct.

Guess a solution:

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

Rewrite recursion (8) 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 $\upsilon$ must solve:

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

which in particular implies that

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

Solving for $\upsilon$ and $\mu_v$:

$$\upsilon(z) \quad = \quad \beta\frac{\gamma(z) - \gamma(\beta)}{z - \beta}$$

$$\mu_v \;=\; \frac{\beta}{1-\beta}[\mu_c + \frac{(1-\theta)}{2}\gamma(\beta)\cdot\gamma(\beta)].$$

The formula $v(z)$ is familiar from Hansen and Sargent (1980) and is the solution to the forecasting problem:

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

pervasive in the literature on linear rational expectations. The risk parameter $\theta$ enters only the constant term of continuation value process.

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 $\beta$ 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. For instance, the price of payoff $\phi(w_{t+1})$ is given by:

$$E\left[\exp(s_{t+1})\phi(w_{t+1})|\mathcal{F}_t\right] \;=\; E\left[\exp(s_{t+1})|\mathcal{F}_t\right]\frac{E\left[\exp(s_{t+1})\phi(w_{t+1})|\mathcal{F}_t\right]}{E\left[\exp(s_{t+1})|\mathcal{F}_t\right]}$$

The first term is a pure discount term and the second the is the expectation of $\phi(w_{t+1})$ under the so-called risk neutral probability distribution. The logarithm of the first term is:

$$\log E\left[\exp(s_{t+1})|\mathcal{F}_t\right] = -\delta - \sum_{j=0}^{\infty}\gamma_{j+1}w_{t-j} - (1-\theta)\gamma(\beta)\cdot\gamma_0 + \frac{\gamma_0\cdot\gamma_0}{2},$$

which is minus the yield on a discount bond. The $w_{t+1}$ coefficient on the innovation to the logarithm $s_{t+1,t}$ of the stochastic discount factor is

$$-\gamma_0 + (1-\theta)\gamma(\beta).$$

This vector is also the mean of the normally distributed shock $w_{t+1}$ under the risk-neutral distribution. The adjustment $-\gamma_0$ is familiar from Hansen and Singleton (1983) and the term

$(1-\theta)\gamma(\beta)$ is the adjustment for the intertemporal composition of consumption risk implied by the Kreps and Porteus (1978) specification of recursive utility. Large values of the risk parameter $\theta$ enhance the importance of this component. This latter effect is featured in the analysis of Bansal and Yaron (2004).[5]

The following Markov example will be used in our calculations.

**Example 3.1.** *Suppose that consumption evolves according to:*

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

*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*

$$v_t = U_v x_t + \mu_v$$

*where*

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

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

*and*

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

*The logarithm of the stochastic discount factor is:*

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

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). This complexity will challenge further the measurements we describe.

[5] Anderson, Hansen, and Sargent (2003) suggest a different interpretation for the parameter $\theta$. 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.

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 (2005) 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 $\gamma$ is now measured using a market return.[6] In this paper we instead follow Hansen and Singleton (1983), Restoy and Weil (1998), Bansal and Yaron (2004) and others by focus on consumption dynamics. Continuation values are thus a central ingredient in our analysis.

### 3.4 Intertemporal substitution ($\rho \neq 1$)

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 $Dv^1_t$ to use in a first-order approximation:

$$v_t \approx v^1_t + (\rho - 1)Dv^1_t$$

where $V^1_t$ is the continuation value for an economy in which $\rho = 1$ and $v^1_t = logV^1_t - c_t$. In appendix A, we derive the following recursion for the derivative:

$$Dv^1_t = -\frac{(1-\beta)(v^1_t)^2}{2\beta} + \beta\tilde{E}(Dv^1_{t+1}|\mathcal{F}_t)$$

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

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

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

[6] Campbell and Vuolteenaho (2003) refer to this second term as the *bad* $\beta$ term.

zero, it has a risk adjusted mean of $(1-\theta)\gamma(\beta)$. The derivative $Dv_t^1$ is negative because it is the (distorted) expectation of the sum of negative random variables.

When $\rho$ is different from one, the wealth-consumption ratio is not constant. A first-order expansion of the continuation value implies a second-order expansion of the consumption-wealth ratio. This can be seen directly from (7):

$$\begin{aligned}\log W_t - \log C_t &= -\log(1-\beta) + (1-\rho)\left[v_t^1 + (\rho-1)Dv_t^1\right] \\ &= -\log(1-\beta) - (\rho-1)v_t^1 - (\rho-1)^2 Dv_t^1.\end{aligned}$$

The term $v_t^1$ is very similar (but not identical to) the term typically used when taking log-linear approximations.[7] Recall that this term is the expected discounted value of consumption growth with an additive term constant term that adjusts for variability. In the first-order approximation of the wealth-consumption ratio, $v_t^1$ shows how the wealth-consumption ratio is altered with the intertemporal substitution parameter $\rho$.

The first-order term captures the well known property that when consumption growth rates are predictable, this predictability should be reflected in the consumption wealth ratio. Forecasts that a geometric average of future consumption will be higher than current consumption imply a higher wealth-consumption ratio when $\rho$ exceeds one and a lower one $\rho$ is less than one. This is evident because the immediate response of $v_t^1$ to a shock $w_{t+1}$ is given by $[\gamma(\beta)-\gamma_0]w_{t+1}$, which is the difference between the discounted response and the instantaneous response. In contrast, the risk parameter $\theta$ alters the constant term in $v_t^1$. This implication of intertemporal substitution is familiar from previous literature (*e.g.* see Campbell (1996) and Restoy and Weil (1998)). By construction, the second-order term adjusts the wealth consumption ratio in a manner that is symmetric about $\rho = 1$. When $\rho$ deviates from one, this second-order correction is positive.

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

$$s_{t+1,t} \approx s_{t+1,t}^1 + (\rho-1)Ds_{t+1,t}^1$$

where

$$Ds_{t+1,t}^1 = v_{t+1}^1 - \frac{1}{\beta}v_t^1 + (1-\theta)\left[Dv_{t+1}^1 - \tilde{E}\left(Dv_{t+1}^1|\mathcal{F}_t\right)\right].$$

Recall that in Example 3.1, $c_{t+1}-c_t$ has conditional mean: $\mu_c + U_c x_t$ and a shock contribution: $\gamma_0 w_{t+1}$. Using the parameterization, the logarithm of the continuation value/consumption ratio is:

$$\begin{aligned}v_{t+1}^1 &= U_v x_{t+1} + \mu_v \\ &= U_v H w_{t+1} + U_v G x_t + \mu_v.\end{aligned}$$

In appendix A, we show that

$$Dv_{t+1}^1 = -\frac{1}{2}x_{t+1}{}'\Upsilon_{dv}x_{t+1} + U_{dv}x_{t+1} + \mu_{dv}$$

[7] In log-linear approximation the discount rate in this approximation is linked to the mean of the wealth consumption ratio. In the $\rho$ expansion, the subjective rate of discount is used instead.

where formulas for $\Upsilon_{dv}$, $U_{dv}$ and $\mu_{dv}$ are given in appendix A.

We could use this expansion to produce approximations to equilibrium prices, in particular the implied risk neutral prices. In the example economy the first-order approximation of the stochastic discount factor implies that the risk neutral distribution for $w_{t+1}$ remains normal but with an enhanced covariance matrix and an alternative mean. In a continuous-time approximation, only the mean adjustment is present. The first-order expansion of the altered mean can be expressed as:

$$-\rho\gamma(0) + (\rho - \theta)\gamma(\beta) + (\rho - 1)(1 - \theta)(U_{dv} - H'\Upsilon_{dv} G x_t)].$$

The term $-\rho\gamma(0)$ is familiar from the work of Hansen and Singleton (1983) and the term $(\rho - \theta)\gamma(\beta)$ is the approximate adjustment for recursive utility. The third term $(\rho - 1)(1 - \theta)(U_{dv} - H'\Upsilon_{dv} G x_t)$ is new relative to the more typical log-linear approximation. It is time varying when $\rho$ and $\theta$ are distinct from unity and consumption growth rates are predictable. The variation is present even though the consumption process in the example economy is homoskedastic (in logarithms). For the estimated laws of motion we consider, this source of variation is small.

Our interest in sensitivity to preference parameters extends beyond one-period risk adjustments. We will use this first-order expansion in the stochastic discount factor to compute derivatives of the logarithms of security prices with positive payoffs at different horizons including ones far in to the future.

# 4 Shocks and Vector Autoregressions

As in much of the empirical literature in macroeconomics, we use vector autoregressive (VAR) models to identify interesting aggregate shocks and estimate $\gamma(z)$. We consider a specification of the VAR that is rich enough to allow experimentation with different long-run assumptions and different variables that may be important in identifying the long-run consequences of macroeconomic shocks.

In our initial model we let consumption be the first element of $y_t$ and corporate earnings be the second element:

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

This vector is presumed evolve as a VAR of order $\ell$. In the results reported subsequently, $\ell = 5$. The least restrictive specification we consider is:

$$A_0 y_t + A_1 y_{t-1} + A_2 y_{t-2} + ... + A_\ell y_{t-\ell} + B_0 = w_t\,, \tag{9}$$

The vector $B_0$ two-dimensional, and similarly the square matrices $A_j, j = 1, 2, ..., \ell$ are two by two. 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:

$$\gamma(\beta) \;=\; (1-\beta)u_c A(\beta)^{-1}$$

where $u_c$ is a column vector with a one in the first position and a zero in the second entry.

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. Our inclusion of corporate earnings in the VAR is motivated by the work of Lettau and Ludvigson (2001) and Santos and Veronesi (2001). The second time series is mean to capture aggregate exposure to stock market cash flows. It is used as a predictor of consumption and as an additional source of maroeconomic risk. We measure corporate earnings from NIPA and convert this series to real terms using the implicit price deflator for nondurables and services.

Following Hansen, Heaton, and Li (2005), we consider two specifications of the evolution of $y_t$. In one case the model is estimated without additional restrictions, 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 $\alpha$ 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. In this system both series respond in the same way to shocks in the long run. Specifically, the limiting response of consumption and earnings to a shock at date 0 is the same. 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.

Our use of a second time series is to identify additional sources of long-run risk beyond just a single consumption innovation. Whereas Bansal and Yaron (2004) consider multivariate specifications of consumption risk, they seek to infer this risk from a single aggregate time series on consumption or aggregate dividends. With flexible dynamics, such a model is not well identified from time series evidence. On the other hand, while our shock identification allows for flexible dynamics, it requires that we specify *a priori* the important sources of macroeconomic risk.

In our analysis, we will not be concerned with the usual shock identification familiar from the literature on structural VAR's. This literature assigns structural labels to the underlying shocks and imposes *a priori* restrictions to make this assignment. While we have restricted $C$ to be lower triangular, this is just a normalization. This restriction leads to the identification of two shocks, but other shock configurations with an identity as a covariance matrix can be constructed by taking linear combinations of the initial two shocks we identify. Sometimes we will construct a temporary shock, a linear combination of shocks that has no long run impact on consumption and corporate earnings, and a permanent shock, a shock that has the same long-run impact on consumption and earnings. The permanent shock is uncorrelated with the temporary shock. This construction is much in the same spirit as Blanchard and Quah (1989). What primarily interests us, however, is the intertemporal composition of consumption risk and not the precise labels attached to individual shocks.

We report impulse responses for estimates of the VAR with and without the cointegration restriction in figure 1. When cointegration is imposed, corporate earnings relative to consumption identifies an important long-run response to both shocks. The long-run impact of the first *consumption shock* is twice that of the impulse on impact. While the second *earnings shock* is normalized to have no immediate impact on consumption, its long-run impact is sizeable. We demonstrated in the recursive utility model, that the geometrically weighted average response of consumption to the underlying shocks is a key ingredient in the stochastic discount factor. As the subjective discount rate converges to zero, this average coincides with the limiting consumption response.

Notice from the impulse responses in figure 1, that when the cointegration restriction is not imposed, the estimated long-run consumption responses are substantially smaller. The imposition of the cointegration restriction is critical to locating an important low frequency component in consumption. Moreover, in the absence of this restriction, the overall feedback from earnings shocks to consumption is substantially weakened. The earnings shocks have little impact on consumption for the *no cointegration* specification.

Using the cointegration specification, we explore the statistical accuracy of the estimated responses. Following suggestions of Sims and Zha (1999) and Zha (1999), we impose Box-Tiao priors on the coefficients of each equation and simulate histograms for the parameter estimates. This provides approximation for Bayesian posteriors with a relatively diffuse (and improper) prior distribution. These “priors” are chosen for convenience, but they give us a simple way for us to depict the sampling uncertainty associated with the estimates.

In the model of Hansen and Singleton (1983), it is the immediate innovation in consumption in consumption that matters for pricing one-period securities. Figure 2 gives a histogram for the standard deviation of this estimate. In other words it gives the histogram for the estimate of the $(1,1)$ entry of $A_0$.

For comparison we also report the histogram for a long-run response using the permanent-transitory decomposition just described. Figure 2 also gives a histogram for the long-run consumption response to a long-run shock. The permanent shock is normalized to have unit standard deviation, so that we can compare magnitudes across the long-run and short run responses.

As might be expected, the short run response estimate is much more accurate than the long-run response. Notice that the horizontal scales of histogram differ by a factor of ten. In particular, while the long-run response is centered at a higher value and it also has a substantial right tail. Consistent with the estimated impulse response functions, the median long-run response is about double that of the short-term response. In addition nontrivial probabilities are given to substantially larger responses.[8] Thus from the standpoint

[8]The accuracy comparison could be anticipated in part from the literature on estimating linear time series

## Impulse Response of Consumption and Earnings

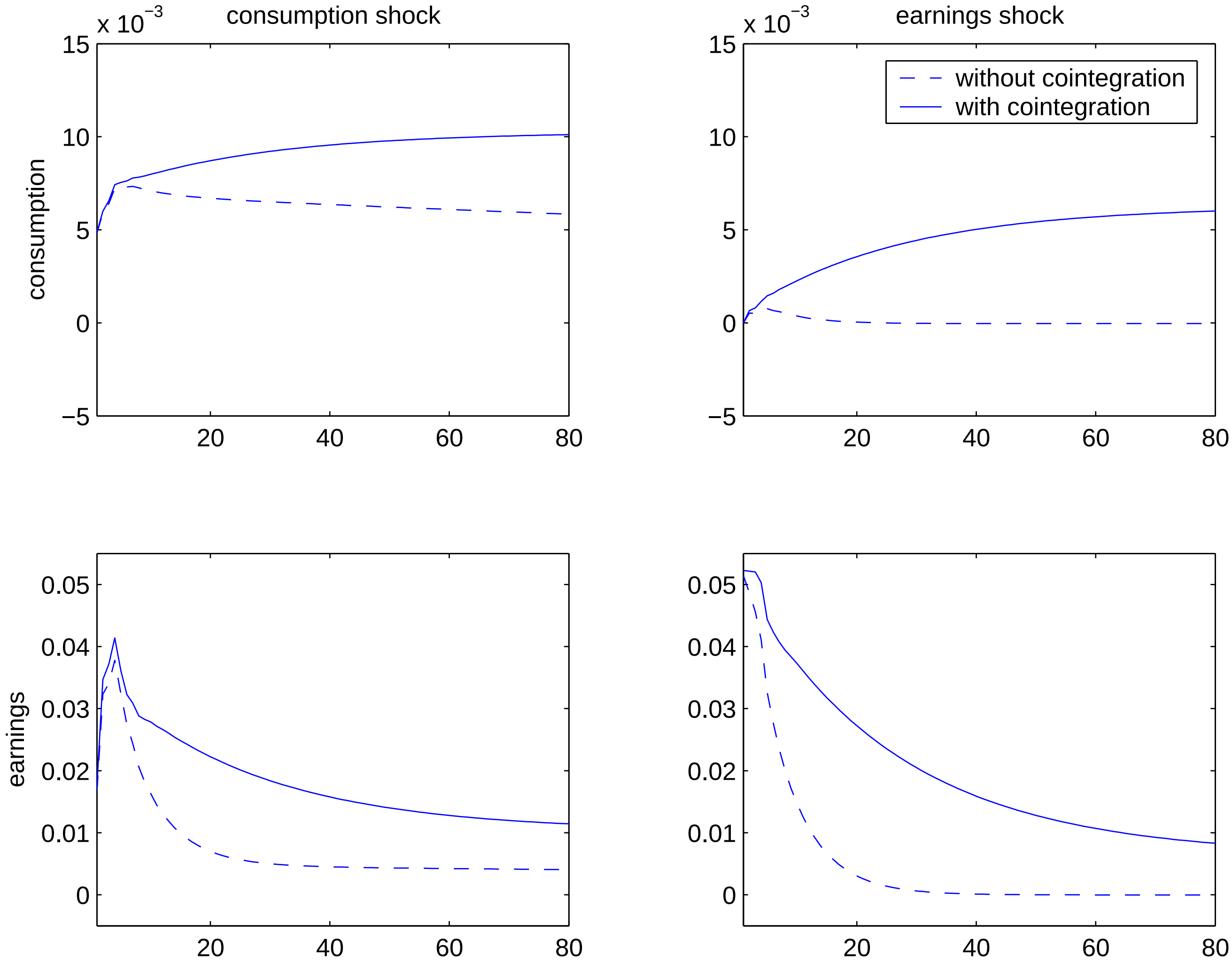



Figure 1: The impulse responses without imposing cointegration were constructed from a bivariate VAR with entries $c_t$, $e_t$. These responses are given by the dashed lines $---$. Solid lines —— are used to depict the impulse responses estimated from a cointegrated system. The impulse response functions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

of sampling accuracy, the long-run response could be even more than double that of the immediate consumption response. 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 the evolution of consumption aids our understanding of the equity premium.

The cointegrated specification with a known cointegrating coefficient imposes a restriction on the VAR. To explore the statistical plausibility of this restriction, we free up the cointegration relation by allowing consumption and earnings to have different long-run responses. To assess statistical accuracy we simulate the posterior distribution for the cointegrating coefficient imposing a Box-Tiao prior for each VAR conditioned on the cointegrating coefficient. The resulting histogram is depicted in figure 3. For sake of computation, we used a uniform prior over the interval $[-2, 2]$ for the cointegrating coefficient. This figures suggests that the *balanced growth* coefficient of unity is plausible.[9]

# 5 Long-Run Cash Flow Risk

We have seen evidence for an important long-run component in consumption when combined with the preference specification of section 3.1. We now investigate how growth rate risk in cash flows is encoded in asset prices. Specifically, we consider when riskiness about long-run cash flow growth can have an important contribution to current value.

To explore this issue first consider a stationary Markov specification for $\{x_t\}$, a process used to depict the underlying valuation. The logarithm of consumption evolves according to:

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

This model nests the specifications we have considered so far as special cases.

In what follows we consider cash flows that may not grow proportionately with consumption as in Campbell and Cochrane (1999), Bansal, Dittmar, and Lundblad (2005), Lettau, Ludvigson, and Wachter (2004), and others. For example, the sorting method we use in constructing some of our portfolios can induce permanent differences in dividend growth. For this reason we allow cash flows or dividends to risky securities to be *levered* claims on consumption in the long run. In our first specification of cash flows we therefore consider,

$$d_{t+1} - d_t = \mu_d(x_t) + \sigma_d(x_t) w_{t+1}.$$

where $d_t$ is the logarithm of the cash flow.

---

models using a finite autoregressive approximation to an infinite order model (see Berk (1974)). The *on impact* response is estimated at the parametric rate, but the *long-run* response is estimated at a considerably slower rate that depends on how the approximating lag length increases with sample size. Our histograms do not confront the specification uncertainty associated with approximating an infinite order autoregressions, however.

[9]The model with cointegration imposes two restrictions on the matrix $A(1)$. Twice the likelihood ratio for the two models is 5.9. The Bayesian information or Schwarz criterion selects the restricted model.

**Approximate Posterior Distributions for Responses**

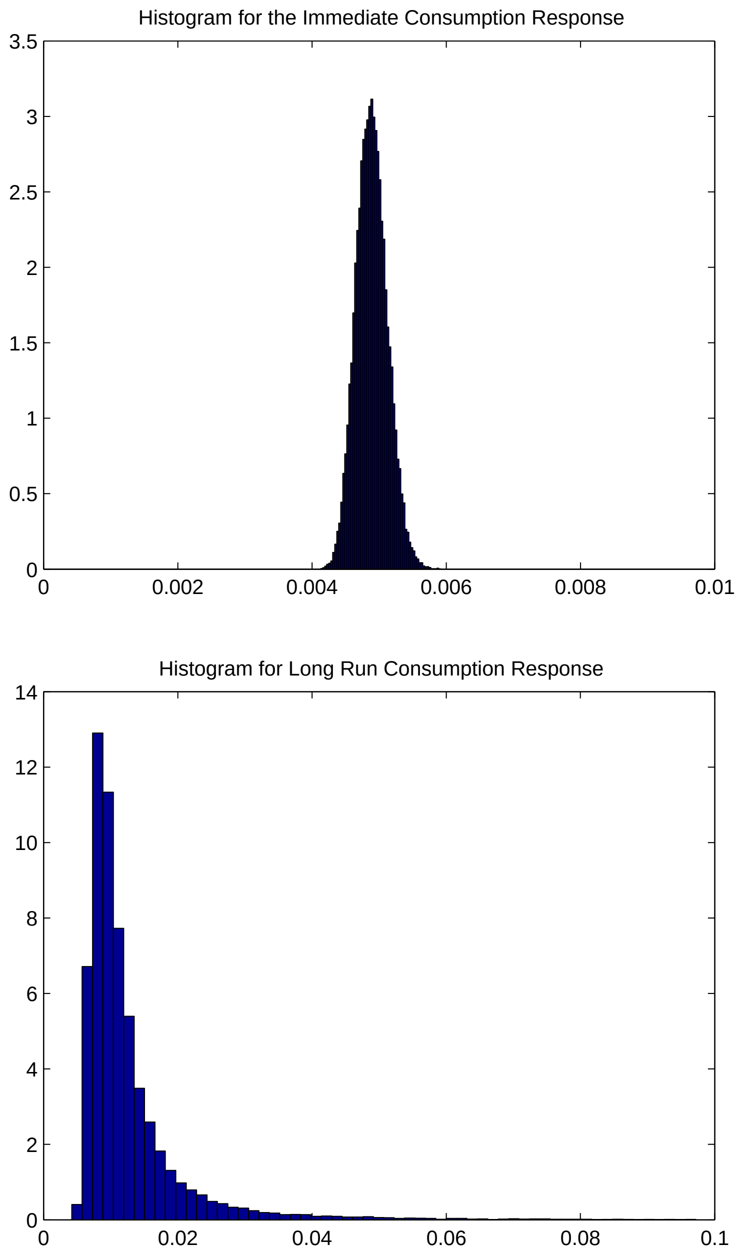



Figure 2: The top panel gives the approximate posterior for the immediate response to consumption and the bottom panel the approximate posterior for the long-run response of consumption to the permanent shock. The histograms have sixty bins with an average bin height of unity. They were constructed using using Box-Tiao priors for each equation. Vertical axes are constructed so that on average the histogram height is unity.

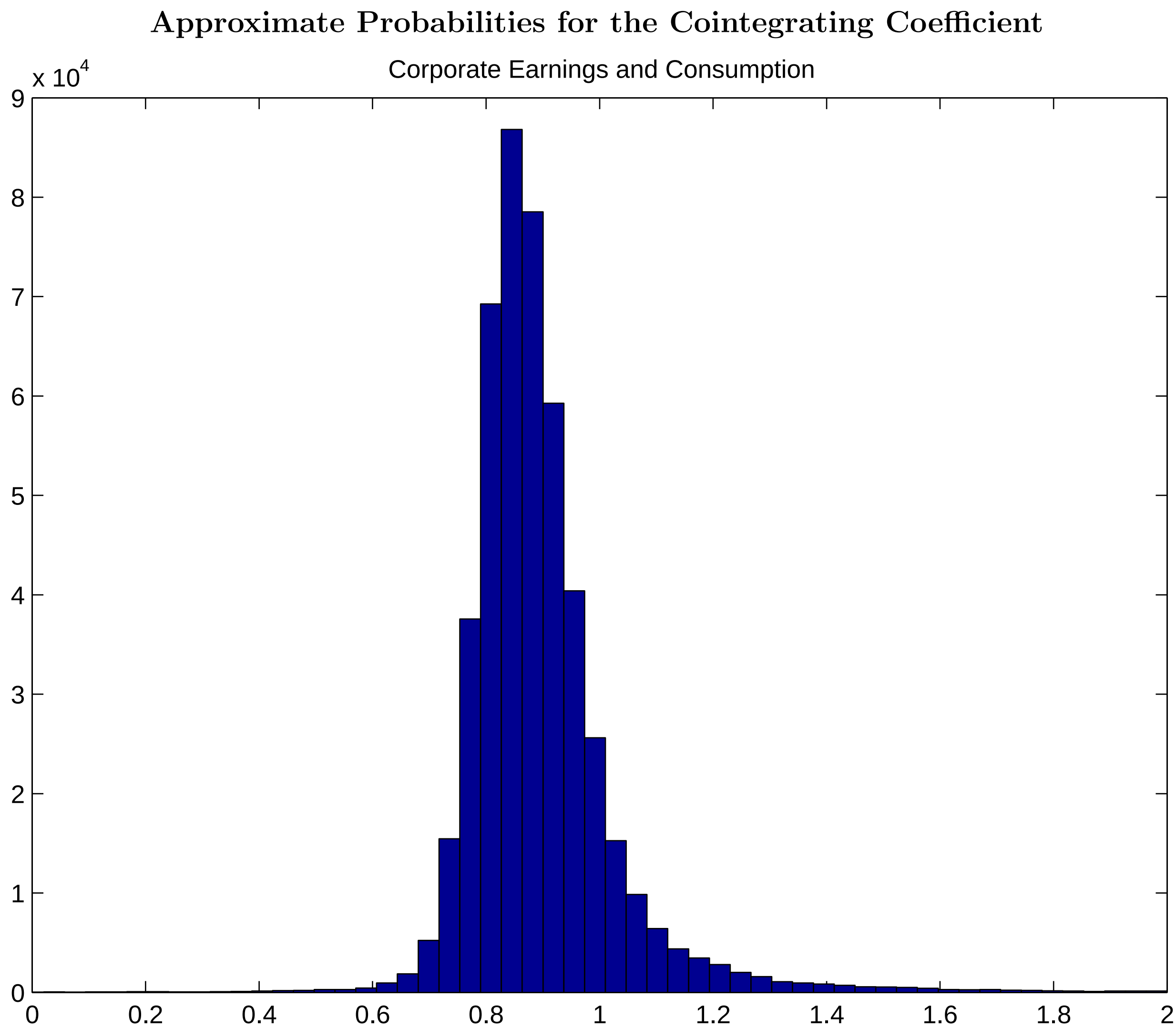



Figure 3: Box-Tiao priors are imposed on the regression coefficients and innovation variances conditioned on the cointegrating coefficient. Posterior probabilities are computed by simulating from a Markov chain constructed from the conditional likelihood function.

In what follows we will focus exclusively on log linear specifications, but the apparatus we describe allows for nonlinearities as well.

## 5.1 A useful time series decomposition

Prior to characterizing the value contribution of long-run cash flow risk, we employ a useful martingale decomposition of processes that display stochastic growth. Following example 3.1, let the Markov process follows a stationary first-order vector autoregression with mean zero. The joint dynamics of consumption and dividends are given by:

$$\begin{aligned} c_{t+1} - c_t &= \mu_c + U_c x_t + \gamma_0 w_{t+1} \\ d_{t+1} - d_t &= \mu_d + U_d x_t + \iota_0 w_{t+1}. \end{aligned}$$

Alternatively, we may specify the process in moving-average form:

$$\begin{aligned} c_{t+1} - c_t &= \mu_c + \gamma(L) w_{t+1} \\ d_{t+1} - d_t &= \mu_d + \iota(L) w_{t+1}. \end{aligned}$$

where

$$\iota(z) = \sum_{j=0}^{\infty} \iota_j z^j$$

Martingale approximation of the stationary increment processes for log consumption and log dividends formally depicts the target time series as the sum of a martingale and the first difference of a stationary process. It is commonly used in establishing central limit approximations (*e.g.* see Hall and Heyde (1980)), and it is not limited to linear processes (*e.g.* see Hansen and Scheinkman (1995) for a nonlinear Markov version.) For scalar linear time series, it coincides with the decomposition of Beveridge and Nelson, but it is applicable much more generally. In the case of consumption, this decomposition is given by:

$$c_{t+1} - c_t = \mu_c + \gamma(1) w_{t+1} + U_c^* x_{t+1} - U_c^* x_t$$

where

$$\gamma(1) = [U_c(I - G)^{-1} H + \gamma_0] = \sum_{j=0}^{\infty} \gamma_j$$

and

$$U_c^* \doteq -U_c(I - G)^{-1}.$$

Then $c_t$ has a growth component $\mu_c t$ and a martingale component with increment: $\gamma(1) w_t$. The vector $\gamma(1)$ is sum of the impulse response vectors of consumption growth rates to the shocks. Equivalently, it is the long-run response of consumption to these shocks. The remaining component of $c_t - c_{t-1}$ is the first difference of stationary process, or equivalently it is a stationary component of log consumption.

Similarly, we write

$$d_{t+1} - d_t = \mu_d + \iota(1) w_{t+1} + U_d^* x_{t+1} - U_d^* x_t \tag{10}$$

Thus $d_t$ has a growth rate $\mu_d$ and a martingale component with increment: $\iota(1) w_t$. The vector $\iota(1)$ is the long run response vector for log dividends. As an example, consumption and dividends are cointegrated when

$$\iota(1) = \lambda\gamma(1), \quad \mu_d = \lambda\mu_c$$

for some $\lambda$. As we will see, it is $\mu_d t$ and the martingale increments that are the cash flow inputs into the long run contribution to value. The long-run contribution does not depend on the stationary component and hence is invariant to $U_d^*$.

So far, we have considered a cash flow process with log-linear dynamics. Suppose instead we consider a *share model* as in Santos and Veronesi (2001). The discrete-time version of such a model can be depicted as:

$$D_t = C_t \Psi(x_t)$$

where $\Psi(x_t)$ is restricted to be between zero and one and gives the dividend share of aggregate consumption. Thus

$$d_t - d_{t-1} = c_t - c_{t-1} + \log \Psi(x_t) - \log \Psi(x_{t-1})$$

By construction this share model assumes that log consumption and log dividends share the same stochastic growth, so that the long-run dividend risk is the same as that of consumption. The counterpart to $\iota(1)$ is the long-run consumption response $\gamma(1)$. While physical claims to resources may satisfy balanced growth restrictions, financial claims of the type we investigate need not. Share models are not attractive models of the cash flows we consider in the next section unless the share process is allowed to have a very pronounced low frequency component.

## 5.2 Operator Valuation

We study the long-run effect of cash flow risk on pricing using valuation operators. These operators are the counterparts to the matrices constructed in section 2. We are required to use an operator formulation to exploit the mathematical convenience of the continuous state Markov process implied by the VAR statistical model of section 4. There we found that important elements of the state variable are current and lagged values of consumption and corporate earnings and that these state variables identified a low-frequency component of consumption. To capture this component of consumption would require a very large number of states in a discrete approximation. The VAR specification is also convenient as it allows us to consider alternative statistical assumptions in a simple way. To accommodate an environment in which the states are continuous, the operators are constructed instead using integrals. We study the behavior of these valuation operators using the counterpart to the eigenvalue methods described in section 2.

The counterpart to the matrix $\mathcal{P}$ used in section 2 is the one-period valuation operator given by:

$$\mathcal{P}\psi(x) = E\left(\exp\left[s_{t+1,t} + \zeta + \pi w_{t+1}\right]\psi(x_{t+1})|x_t = x\right).$$

As before we call this a *valuation operator* because it depends on the value of $\pi$ and $\zeta$ which account for the long-run components of cash flows. Formally, we view this operator as mapping functions of the Markov state into functions of the Markov state. In particular, it is well defined for functions that are bounded functions of the Markov state, but it is well defined for other functions as well. This operator takes a payoff at date $t+1$ of the form:

$$D_{t+1} = \exp\left((t+1)\zeta + \pi \sum_{\tau=1}^{t+1} w_\tau\right)\psi(x_{t+1})D_0, \tag{11}$$

and maps it into a price today scaled by $D_t$. Typically $D_0$ is initialized to be positive. Since payoffs and prices are scaled by a growth factor, the valuation operator depends on the choice of $\pi$ and $\zeta$ used in the scaling. The log-linear specification of dividend growth given in (10) is a special case of this model in which $\pi = \iota(1)$, $\zeta = \mu_d$, and $\psi(x_t) = \exp(U_d^* x_t)$.[10]

Multi-period prices can be inferred from this one-period pricing operator through iteration. The value of a date $t+j$ cash flow (11) is given by:

$$D_t\left[\mathcal{P}^j\psi(x_t)\right].$$

The notation $\mathcal{P}^j$ denotes the application of the one-period valuation operator $j$ times, and it is the counterpart to raising a matrix to the $j$ power.

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

$$\frac{P_t}{D_t} = \frac{\sum_{j=1}^{\infty}\mathcal{P}^j\psi(x_t)}{\psi(x_t)} \tag{12}$$

provided that $\psi(x_t)$ is strictly positive. The term

$$\frac{\mathcal{P}^j\psi(x_t)}{\psi(x_t)}$$

is the contribution of the date $t+j$ cash flow 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 define operators that we use to measure these rates and the limiting growth behavior. Let

$$\mathcal{G}\psi(x) = E\left[\exp\left(\zeta + \pi w_{t+1}\right)\psi(x_{t+1})|x_t = x\right].$$

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) is:

$$D_t\left[\mathcal{G}^j\psi(x_t)\right].$$

[10]The operator formulation allows $\psi$ to be negative with positive probability and thus allows for the study of cash flows that are sometimes negative.

### 5.3 Limiting Behavior

As in section 2, we study the limiting behavior by constructing positive eigenfunctions. Consider first the solution, $\phi$, to:

$$-\nu = \sup_{\psi>0} \inf_{x} \left(\log \mathcal{P}\psi - \log \psi\right).$$

The solution equalizes the objective across states at $-\nu$. The objective is the logarithm of the value payoff ratio, and problem is to guarantee a high value of this objective across all states. Thus $\phi$ satisfies the equation:

$$\mathcal{P}\phi = \exp(-\nu)\phi,$$

and is only well defined up to scale.

Recall that the left eigenvector of a matrix is the right eigenvector of its transpose. Here $\varphi$ is the eigenfunction of the adjoint of the operator $\mathcal{P}$ where the adjoint is the operator equivalent of a transpose. As shown by Hansen and Scheinkman (2005), whenever $E\left(\psi\varphi\right)$ and $E\left(\varphi\phi\right)$ are well defined and finite:

$$\lim_{j\to\infty} \exp(\nu j)\mathcal{P}^j\psi(x) = \frac{E(\varphi\psi)}{E(\varphi\phi)}\phi(x). \tag{13}$$

Thus when $E(\varphi\psi) > 0$,

$$\lim_{j\to\infty} \frac{\log\left[\mathcal{P}^j\psi(x)\right]}{j} = -\nu.$$

This calculation gives us an asymptotic decay rate that depends on both cash flow growth through the specification of $\pi$ and $\zeta$, on the economic value associated with that growth, but not the particular function $\psi$ that dictates the transient contribution to cash flows. The eigenfunction $\phi$ is dominant as it gives the limiting state dependence of the values as reflected in formula (13). Thus the pair $(\nu, \phi)$ measures how long-run prospects about dividends contribute to value. The $\psi$ contribution is transient and does not alter the asymptotic decay rate or the appropriately scaled limiting value.

The asymptotic cash-flow growth is characterized by an analogous eigenfunction-eigenvalue pair. A straightforward calculation shows that the dominant eigenfunction of $\mathcal{G}$ is one and that

$$\mathcal{G}\psi = \exp\left(\zeta + \frac{|\pi|^2}{2}\right)\psi$$

for $\psi = 1$.[11] Thus

$$\epsilon = \zeta + \frac{|\pi|^2}{2}.$$

is the implied asymptotic rate of growth for the cash flow.

In what follows we will motivate the study of $\epsilon + \nu$. This sum depends on $\pi$ but not $\zeta$. It also depends on the preference parameters $\delta$, $\theta$ and $\rho$ through the stochastic discount factor $s_{t+1}$.

[11] When the martingale approximation for the cash flow has heteroskedastic increments, this calculation ceases to have a trivial solution.

### 5.4 Securities constructed from dominant eigenfunctions

We aim to approximate dividend price ratios and returns to a security with dividends payments only far off into the future.

First we use the dominant eigenfunction to construct a valuation process and the corresponding return. A valuation process $\{J_t : t = 1, 2, ...\}$ is one for which the date $t$ price of the security with liquidation value $J_{t+1}$ is $J_t$. We can construct such securities by supposing that dividends are continually reinvested. Form:

$$J_{t+1} = \exp\left[(\nu + \zeta)(t+1) + \pi \sum_{\tau=1}^{t+1} w_j\right] \phi(x_{t+1}).$$

Since $\phi$ is a positive eigenfunction, the date $t$ value of the payoff $J_{t+1}$ is indeed $J_t$, verifying that $J_{t+1}$ is indeed a valuation process. Notice that the riskiness of the one period return depends on $\pi$ and the response of $\log \phi$ to the underlying shocks. In our calculations, $\pi = \iota(1)$ which is extracted as the permanent component to cash flows. The implied return includes an additional value contribution captured by the logarithm of the dominant eigenfunction, $\log \varphi$.

The $k$ period return is:

$$R^k_{t+k} = \frac{J_{t+k}}{J_t} = \exp\left[(\zeta + \nu)k + \pi \sum_{\tau=1}^{k} w_{\tau+t}\right] \frac{\phi(x_{t+k})}{\phi(x_t)}.$$

Take expectations and logarithms:

$$\lim_{k\to\infty} \frac{1}{k} \log E\left(R^k_{t+k} | \mathcal{F}_t\right) = \nu + \lim_{k\to\infty} \frac{1}{k} \log \mathcal{G}^k \phi(x_t) = \nu + \epsilon \tag{14}$$

provided that $E\phi\kappa$ is finite where $\kappa$ is the eigenfunction of the adjoint of the operator $\mathcal{G}$.

Consider next a security with a dividend process of the from (11) using the eigenfunction $\phi$ in place of $\psi$. This security has a constant price/dividend ratio. Using the eigenvalue property and formula (12), the price-dividend ratio is

$$\frac{\exp(-\nu)}{1 - \exp(-\nu)}, \tag{15}$$

which does not vary across states. As in the Gordon growth model, the factor $\exp(-\nu)$ includes both a pure discount factor (adjusted for risk) and a dividend growth factor. The implied discount rate is $\nu + \epsilon$ since the asymptotic dividend growth factor for dividends with long-run risk is: $\exp(\epsilon)$. The one-period return on this constructed equity is the same as that of the valuation process described previously.

These constructed securities are related to the returns to holding equity with the initial dividend processes stripped out. Instead we consider the return on equity with the original cash flow (11), but with the initial payoff date far into the future. While this cash flow depends on the choice of $\varphi$, the initial payoff date is well into the future and the dependence on $\varphi$ is small. Formally, from formula (13), the date $t$ valuation of payoff of the original cash flow (11) $j$-periods into the future is approximately

$$\exp(-\nu j) \exp\left[\zeta(t) + \pi \sum_{\tau=1}^{t} w_\tau\right] \frac{E[\varphi(x_t)\psi(x_t)]}{E[\varphi(x_t)\phi(x_t)]}\phi(x_t).$$

for large $j$. Adding over horizons $j \geq k$ for some large $k$ gives

$$\hat{P}_t^k = \frac{\exp(-\nu k)}{1 - \exp(-\nu)} \frac{E[\varphi(x_t)\psi(x_t)]}{E[\varphi(x_t)\phi(x_t)]} \exp\left[\zeta(t) + \pi \sum_{\tau=1}^{t} w_\tau\right] \phi(x_t)$$

The approximate one period return on this security is:

$$\frac{\hat{P}_{t+1}^{k-1}}{\hat{P}_t^k} = \exp(\nu) \exp\left(\zeta + \pi w_{t+1}\right) \frac{\phi(x_{t+1})}{\phi(x_t)} = \frac{J_{t+1}}{J_t}$$

since $k$ exceeds one. This is the approximate return to holding an equity for which all but the components far into the future are excluded. Not surprisingly, it is the long-run martingale component of the cash flow that dictates what component of the cash is valued when the payoff is far into the future. The dominant eigenfunction adds an addition source of return risk necessary to understand how cash flow risk is transmitted into return risk far into the future. In what follows we will refer to the constructed return $\frac{J_{t+1}}{J_t}$ as a valuation return associated with a cash flow with risk vector $\pi$.

Characterizing the dependence of $\nu + \epsilon$ on $\pi$ gives a long-run risk return relation. The vector $\pi$ gives the cash flow weights on the underlying shocks and $\nu + \epsilon$ gives the implied expected rate of return. By setting $\pi = 0$, we obtain a benchmark return that this the long-run counterpart to the riskfree return. The resulting return is the maximal growth return of Bansal and Lehmann (1997). Alvarez and Jermann (2001) study of the holding period returns to long-horizon discount bonds. The return to a long horizon bond is:

$$\frac{\mathcal{P}^{k-1}\psi(x_{t+1})}{\mathcal{P}^k\psi(x_t)}$$

for $\psi = 1$ and $\mathcal{P}$ constructed using $\pi = 0$ and $\zeta = 0$. Thus the approximate one period return is:

$$\exp(\nu)\frac{\phi(x_{t+1})}{\phi(x_t)}$$

constructed using the $\pi = 0$ and $\zeta = 0$ for the associated dominant eigenvalue and function. In this case, $\epsilon$ is zero by construction.

### 5.5 Long-Run Risk Return Tradeoff

When the intertemporal substitution parameter is unity, we can provide a complete characterization of the long-run risk-return tradeoff and how this tradeoff depends on parameter values.

**Theorem 5.1.** *Suppose that $\rho = 1$ and consumption follows the first-order dynamics given in example (3.1).*

1. *The dominant eigenfunction $\phi$ is a scale multiple of* $\exp(-\omega x)$ *where*

$$\omega \doteq U_c(I - G)^{-1}.$$

2. *The dominant eigenvalue is* $\exp(-\nu)$ *where*

$$\nu = \mu_c - \zeta + \delta - \frac{|\pi - \gamma(1)|^2}{2} + (\theta - 1)\gamma(\beta) \cdot [\pi - \gamma(1)].$$

3. *The dominant eigenfunction, $\varphi$, of the adjoint of $\mathcal{P}$ is a scale multiple of* $\exp(-\omega^* x)$ *where the formula for $\omega^*$ is given in appendix B.*[12]

4. *The expected valuation rate of return is:*

$$\epsilon + \nu = \varsigma^* + \pi^* \cdot \pi$$

*where*

$$\begin{aligned} \pi^* &\doteq (\theta - 1)\gamma(\beta) + \gamma(1) \\ \varsigma^* &\doteq \mu_c + \delta + (1 - \theta)\gamma(1) \cdot \gamma(\beta) - \frac{|\gamma(1)|^2}{2}. \end{aligned}$$

*Proof.* See appendix B. □

Recall that there are two sources of risk in the implied valuation returns: a long-run cash flow risk $\pi w_{t+1}$ and an additional valuation risk $-\omega H w_{t+1}$ contributed by the dominant eigenfunction. For the $\rho = 1$ economy, the dominant eigenfunction does not depend on the risk aversion parameter $\theta$ or on the vector $\pi$ that weights the shocks; but this result is not true in general.

Since $\omega$ does not depend on $\pi$ in this case, we may view $\pi^*$ as the vector of valuation risk prices. As in the pricing of one-period securities, there is a recursive utility contribution constructed from the discounted consumption response vector $\gamma(\beta)$. There is a second contribution that replaces the single period consumption risk vector $\gamma_0$ with its long run counterpart $\gamma(1)$. As the subjective discount factor $\beta$ tends to unity, the two components become proportional.

[12] Analogously the eigenfunction, $\kappa$, of the adjoint of $\mathcal{G}$ is also log-linear in $x$. This guarantees that the limit in (14) exists.

When $\rho \neq 1$, we expand the dominant eignenvalue of the valuation operator around $\rho = 1$. This allows us to characterize the long-run risk return tradeoff for different values of $\rho$. The expansion we use is given by:

$$\nu = \nu^1 + (\rho - 1)\frac{d\nu}{d\rho}|_{\rho=1}$$

where $\nu^1$ is the dominant eigenvalue when $\rho = 1$. In appendix B we show how to evaluate the derivative of $\nu$ with respect to $\rho$.

### 5.6 Long Horizon Returns

An alternative approach to long-run risk is based entirely on returns, and not cash flows. Consider a one-period return process $\{R_t\}$ and construct the reference growth process by compounding returns:

$$D_t^* = \prod_{j=1}^{t} R_j \tag{16}$$

Thus $D_t^*$ is the time $t$ payment when dividends are continually reinvested in the security and the security is sold at time $t$. The logarithm, $\epsilon$, of the dominant eigenvalue for the corresponding growth operator $\mathcal{G}$ gives the long-run expected rate of return. While we are primarily interested in cash flow valuation, we will also use this return construction for comparison.[13]

## 6 Long-Run Valuation Risk

In this section we report estimates of long-run dividend growth and the risk associated with that growth.

### 6.1 Dividend Dynamics

We identify dividend dynamics and, in particular, the martingale component $\iota(1)$ using VAR methods. Consider a VAR with three variables: consumption, corporate earnings and dividends (all in logarithms). Consumption and corporate earnings are modelled as before in a cointegrated system. We use the cointegrated system because it identifies a long-run consumption risk component that is distinct from the one-step-ahead forecast error of consumption. In addition to consumption and earnings, we include in sequence the dividend series from each of the five book-to-market portfolios and from the market. Thus the same two shocks as were identified previously remain shocks in this system because consumption

[13] If we construct the valuation operator $\mathcal{P}$ using a gross return in place of the dividends, then $\mathcal{P}$ should be a distorted conditional expectation operator. Its dominant eigenfunction is unity as is the corresponding eigenvalue.

and corporate earnings remain an autonomous system. An additional shock is required to account for the remaining variation in dividends beyond what is explained by consumption and corporate earnings.

Formally, we append a dividend equation

$$A_0^* y_t^* + A_1^* y_{t-1} + A_2^* y_{t-2} + ... + A_\ell^* y_{t-\ell} + B_0^* = w_t^* , \tag{17}$$

to equation system (9). In this equation the vector of inputs is

$$y_t^* \doteq \begin{bmatrix} y_t \\ d_t \end{bmatrix} = \begin{bmatrix} c_t \\ e_t \\ d_t \end{bmatrix}$$

and the shock $w_t^*$ is scalar with mean zero and unit variance. This shock is uncorrelated with the shock $w_t$ that enters (9). The third entry of $A_0^*$ is normalized to be positive. We refer to (17) as the dividend equation, and the shock $w_t^*$ as the dividend shock. As in our previous estimation, we set $\ell = 5$. W presume that this additional shock has a permanent impact on dividends by imposing the linear restriction:

$$A^*(1) = \begin{bmatrix} \alpha^* & -\alpha^* & 0 \end{bmatrix} .$$

In the next section we will explore sensitivity to alternative specifications of long-run stochastic growth in the cash flows.

A stationary counterpart to this log level specification can be written in terms of the $c_t - c_{t-1}, e_t - c_t, d_t - d_{t-1}$. We estimated the VAR using these transformed variables with four lags of the growth rate variables and five lags of the logarithmic differences between consumption and earnings.

## 6.2 Long-run Risk in Aggregate Securities

Next we use the VAR estimates to measure long-run risk components of cash flows. The implied sample estimates of $\pi$ and $\zeta$ for each of the cash flows are inputs into these calculations.

In table 1 we report long-run expected rates of return using the dividends from the CRSP value weighted equity. We explore sensitivity as we alter $\theta$, and display derivatives with respect to the intertemporal substitution parameter $\rho$. We compare expected rates of return to those of implied by consumption and those implied by a long-run riskless rate of return. This latter return is used as the reference point for computing expected excess returns and it is the long-run riskless return considered by Alvarez and Jermann (2001).

As is evident from this table, the implied differences in expected returns across securities are small even when $\theta$ is as large twenty. The derivatives of the returns with respect to $\rho$ are large while the derivatives of the excess returns are small. According to the derivatives, increasing $\rho$ by $\epsilon$ adds over three times $\epsilon$ percentage points to the expected rates of returns. While larger values of $\rho$ increase long-run riskless return rate, this increase can be offset

**Valuation Returns for Aggregates**

| Portfolio | Return | Excess Return | Return Derivative | Excess Return Derivative |
|---|---|---|---|---|
| | | | $\theta = 1$ | |
| market | 6.60 | .06 | 3.52 | .00 |
| consumption | 6.58 | .05 | 3.51 | .00 |
| long bond | 6.54 | 0 | 3.51 | 0 |
| | | | $\theta = 5$ | |
| market | 6.69 | .30 | 3.43 | -.01 |
| consumption | 6.58 | .19 | 3.43 | -.01 |
| long bond | 6.39 | 0 | 3.44 | 0 |
| | | | $\theta = 20$ | |
| market | 7.00 | 1.19 | 3.09 | -.07 |
| consumption | 6.58 | .77 | 3.13 | -.04 |
| long bond | 5.81 | 0 | 3.17 | 0 |

Table 1: The excess returns are measured relative to the return on the long horizon discount bond. The derivative entries in columns four and five are computed with respect to $\rho$ and evaluated at $\rho = 1$.

by simultaneously reducing $\delta$.[14] The expected excess returns to valuation are essentially proportional to $\theta$. Quadrupling $\theta$ ($\theta = 5$ to $\theta = 20$), approximately quadruples the numbers in the "Excess Return" column. This approximation is to be expected. The proportionality would be exact if $\gamma(\beta) = \gamma(1)$, and we have chosen our discount factor to be close to unity. Overall, the long-run rate of return heterogeneity is small, even when risk aversion parameter is set to a large number.

## 6.3 Book to Market Portfolios

Next we use five portfolios constructed based on a measure of book equity to market equity, and characterize the time series properties of the dividend series as it covaries with consumption and earnings. We follow Fama and French (1993) and construct portfolios of returns by sorting stocks according to their book-to-market values. We use a coarser sort

[14] Of course there is a limit to this reduction, when $\delta$ is restricted to be positive.

**Properties of Portfolios Sorted by Book-to-Market**

| | Portfolio | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | Market |
| Avg. Return (%) | 7.91 | 8.32 | 9.86 | 10.61 | 12.69 | 8.42 |
| Long-Run Return (%) | 8.16 | 7.97 | 9.96 | 10.27 | 12.15 | 8.40 |
| Avg. B/M | 0.32 | 0.62 | 0.84 | 1.12 | 2.00 | 0.79 |
| Avg. P/D | 49.8 | 33.3 | 27.4 | 24.3 | 25.5 | 33.6 |

Table 2: 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. "Avg. B/M" for each portfolio is the average portfolio book-to-market over the period computed from COMPUSTAT. "Avg. P/D" gives the average price-dividend for each portfolio where dividends are in annual units.

into 5 portfolios to make our analysis tractable.

Summary statistics for these portfolios are reported in table 2. The row labeled "Avg. Return" gives the logarithm of the expected quarterly gross return to holding each security where the expectation is multiplied by four to produce results in annual units. The expected returns are constructed by adding a third equation for the logarithm of gross returns to the VAR for consumption and earnings of section 4. We impose the restriction that consumption and earnings are not Granger Caused by the returns and we estimated a separate VAR for each portfolio. One period expected gross returns are calculated conditional on being at the mean of the state variable implied by the VAR. In the row labeled "Long-Run Return," we also report the logarithm of the dominant eigenvalue of the operator $\mathcal{G}$ implied by the VAR and the compound return process (16). These results are also reported in annual units.

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. Both one-period and long-run average returns generally follow this sort. For example, portfolio 1 has much lower average returns than portfolio 5. It is well known that the differences in these average returns are not well explained by exposure to contemporaneous covariance with consumption.

In this section we are particularly interested in the behavior of dividends from the constructed portfolios. The constructed dividend processes accommodate changes in the classification of the primitive assets and depend on the relative prices of the new and old asset in the book-to-market portfolios. Monthly dividend growth for each portfolio are constructed from the gross returns to holding each portfolio with and without dividends. Using the initial price-dividend ratio for the series, these growth rates are used to construct monthly dividend levels. Dividends on a quarterly basis are constructed as an accumulation of the monthly dividends during the quarter. Our measure of quarterly dividends in quarter $t$ is then constructed by taking an average of the logarithm of dividends in quarter $t$ and over the previous three quarters $t-3$, $t-2$ and $t-1$. This last procedure removes the pronounced seasonal in dividend payments. Details of this construction are given in Hansen, Heaton, and Li (2005), which follows the work of Bansal, Dittmar, and Lundblad (2005), and Menzly, Santos, and Veronesi (2004).

We estimate $\iota(1)$ from the dividend regression, and use this as a measure of $\pi$. We then explore the limiting valuation and rates of returns using the eigenvector methods described previously. Table 3 gives long-run average rates of return for the five book-to-market portfolios. Again we explore formally sensitivity to the risk aversion parameter $\theta$ and report derivatives with respect to the intertemporal elasticity parameter $\rho$.

Complementary to many other asset pricing studies, differences in the average rates of return on long-run valuation securities are small except for large values of the risk aversion parameter $\theta$, say $\theta = 20$. In contrast to aggregate securities, the implied heterogeneity in the valuation returns are now substantial, for large values of $\theta$. Again changing $\theta$, alters the expected excess returns proportionately.

When $\theta = 20$, differences between the valuation returns of portfolios 1 and 5 are similar to the observed differences reported in table 2. The level of the returns in table 3 are lower than those in the data, however. These levels could easily be reconciled by a different choice of the discount factor $\beta$.

As with the aggregate returns, derivatives with respect to $\rho$ are similar across securities so that modest movements in $\rho$ have little impact on the excess long-run returns.

It is of also of interest to study the implied logarithm of the price/dividend ratio decomposed and scaled by horizon. These are reported in figure 4. The lower panel of this figure depicts the dividend growth rate by horizon. The figures are computed assuming that the Markov state is set to its unconditional mean. The limiting values are inputs into the return calculations. When $\theta = 1$, the risk adjustments are very small and the value decomposition is a direct reflection of the dividend growth. Moreover, the values for portfolio one are dominated by those for portfolio five across all horizons.

It is only with high values of the risk aversion parameter $\theta$ that the value decomposition for the low book to market portfolios eventually exceed those of the high book to market portfolios as required by the data (see table 2). To see this, recall that the limiting values of the lines plotted in the upper panels of figure 4 give values $-\nu$ corresponding to each portfolio and $\theta$ combination. Relation (15) maps these values of $-\nu$ into a corresponding long run or tail notion of a price-dividend ratio. When $\theta = 1$ the tail price-dividend ratios (in annual units) for portfolios 1 and 5 are 22.5 and 102.2 respectively. When $\theta = 20$, however, these values are 45.1 and 21.2, respectively, which matches the data more closely.

Finally, in our calculations it can take many time periods for the valuations to approximate their limiting values. Even fifteen years (sixty quarters) is not be long enough to approximate well the limit in some cases. Thus the expected rates of return to valuation do

**Valuation Returns for Portfolios**

| Portfolio | Return | Excess Return | Return Derivative | Excess Return Derivative |
|---|---|---|---|---|
| | | | $\theta = 1$ | |
| 1 | 6.48 | -.06 | 3.51 | .00 |
| 2 | 6.55 | .01 | 3.51 | .00 |
| 3 | 6.67 | .13 | 3.52 | .01 |
| 4 | 6.71 | .17 | 3.52 | .01 |
| 5 | 6.78 | .22 | 3.53 | .02 |
| | | | $\theta = 5$ | |
| 1 | 6.11 | -.27 | 3.45 | .01 |
| 2 | 6.42 | .03 | 3.43 | .00 |
| 3 | 7.01 | .61 | 3.40 | -.04 |
| 4 | 7.18 | .79 | 3.41 | -.03 |
| 5 | 7.43 | 1.03 | 3.40 | -.04 |
| | | | $\theta = 20$ | |
| 1 | 4.74 | -1.07 | 3.23 | .07 |
| 2 | 5.94 | 0.13 | 3.16 | .00 |
| 3 | 8.25 | 2.44 | 2.97 | -.20 |
| 4 | 8.95 | 3.14 | 2.98 | -.19 |
| 5 | 9.90 | 4.09 | 2.92 | -.24 |

.

Table 3: The excess returns are measured relative to the return on a long horizon discount bond. The derivative entries in columns four and five are computed with respect to $\rho$ and evaluated at $\rho = 1$.

indeed extrapolate the implied consumption/dividend dynamics very far into the future.

The price-dividend decomposition include expected growth and expected return contributions. We form the expected excess rate of return by horizon by computing:

$$\frac{400}{\tau}[\log(\mathcal{P}^{\tau}\mathbf{1}) - \log(\mathcal{P}^{\tau}\psi) + \log(\mathcal{G}^{\tau}\psi)].$$

As $\tau$ gets arbitrarily large, the limits converge to the corresponding expected excess return limits given in tables 1 and 3. Figure 5 shows how these expected excess rates of returns change with $\theta$ and with modest movements in $\rho$ for different values of $\tau$. Consistent with our characterization of the limit points, small changes in $\rho$ have little impact on this decomposition. We only consider values of $\rho$ close to unity because the approximation we employ is local to $\rho = 1$. While expected rates of return for portfolio five and the market increase with horizon, those of portfolio one eventually decrease. The portfolio excess rates of return are more responsive to changes in $\theta$ than the market return, consistent with the limiting calculations in tables 1 and 3.

## 6.4 Statistical Accuracy

We consider sampling uncertainty in some of inputs used for long run risk. Recall that these inputs are based in part extrapolation of VAR systems fit to match transition dynamics. As in the related macroeconomics literature, we expect a substantial degree of sampling uncertainty. We now quantify how substantial this is for our application.

When $\rho = 1$, the expected excess returns are approximately equal to:

$$\theta\gamma(1) \cdot \pi.$$

We now investigate the statistical accuracy of $\gamma(1) \cdot \pi$ for the five portfolio, and for the difference between portfolios one and five. The vector $\pi$ is measured using $\iota(1)$. In table 4 we report the approximate posterior distribution for $\gamma(1) \cdot \pi$ computed using an approach advocated by Sims and Zha (1999) and Zha (1999) based on Box-Tiao priors. While there is a considerable amount of statistical uncertainty in these risk measures, there are important differences the expected excess value returns between portfolios one and five.

**Value and Growth Decompositions for Two Portfolios**

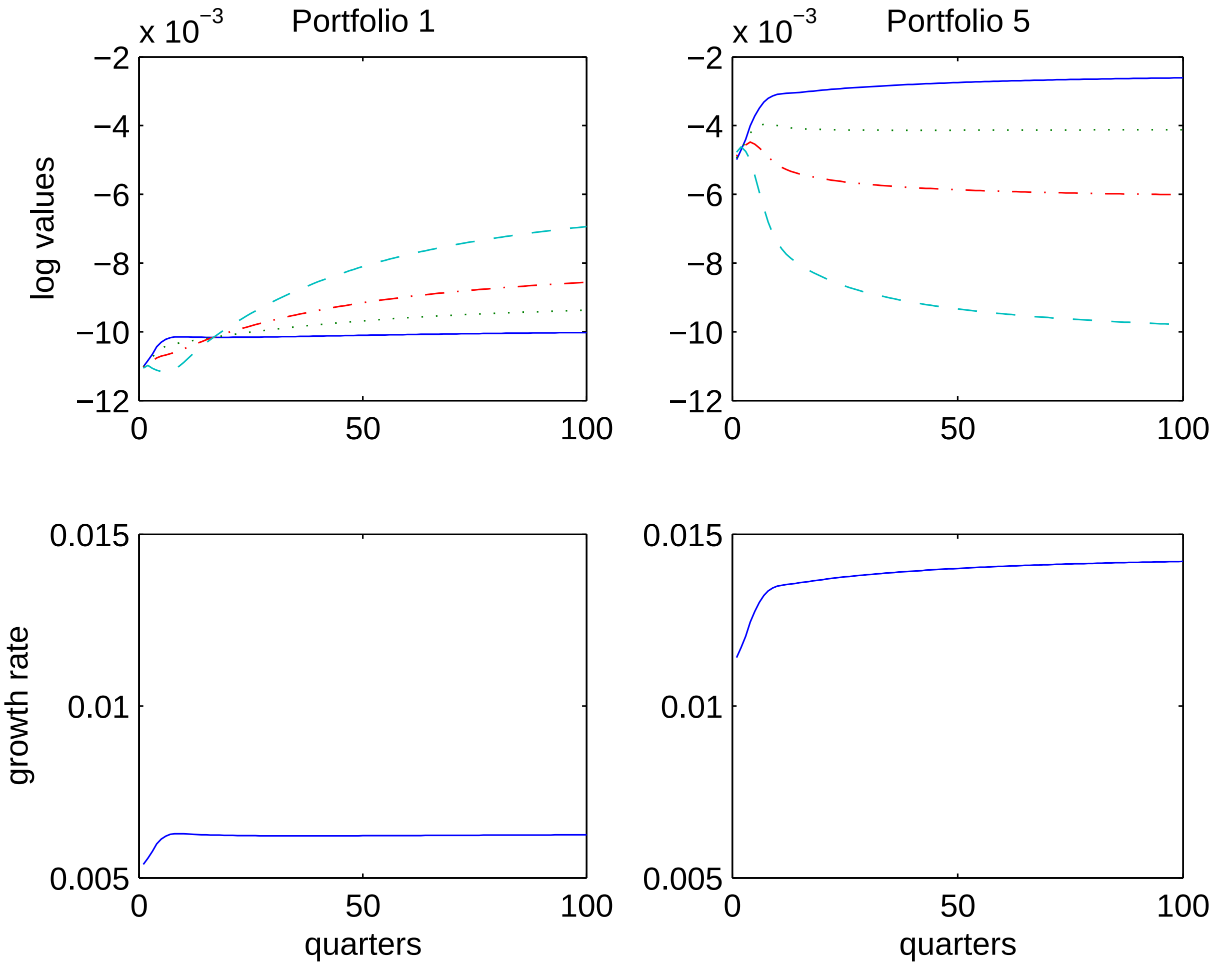



Figure 4: In the top two panels, the __ curve is computed using $\theta = 1$, the $\cdot\cdot$ curve assumes $\theta = 5$, the $-.$ curve assumes $\theta = 10$ and the $--$ curve assumes that $\theta = 20$.

**Excess Return Decompositions by Horizon**

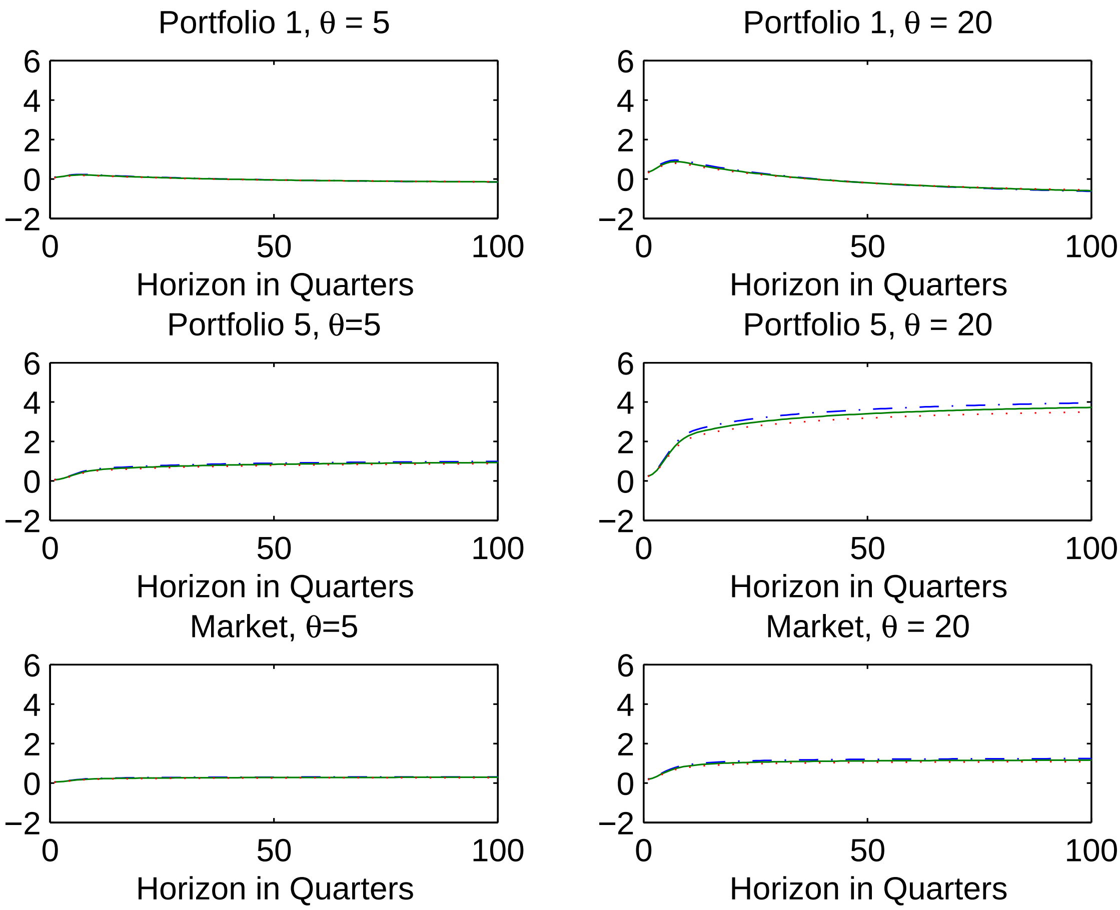



Figure 5: In the top two panels, the __ curves impose $\rho = 1$, the ·· curves impose $\rho = .5$ and the −. curves impose $\rho = 1.5$. The curves for $\rho \neq 1$ were computed using linear approximation around the point $\rho = 1$.

**Accuracy of Risk Measures**

| Portfolio | Quantile | | | | |
|---|---|---|---|---|---|
| | .05 | .25 | .5 | .75 | .95 |
| 1 | -.63 | -.13 | -.05 | -.01 | .10 |
| 2 | -.19 | -.03 | .01 | .04 | .22 |
| 3 | .01 | .06 | .12 | .28 | 1.42 |
| 4 | .04 | .10 | .17 | .32 | 1.46 |
| 5 | .04 | .12 | .21 | .42 | 1.88 |
| market | -.01 | .03 | .06 | .12 | .58 |
| 5-1 | .05 | .15 | .27 | .55 | 2.41 |

Table 4: Quantiles were computed by simulating 100,000 times using Box-Tiao priors. The quantiles were computed using only simulation draws for which the absolute values of the eigenvalues were all less than .999. The fraction of accepted draws ranged from .986 to .987. The quantiles were computed using VAR's that included consumption, corporate earnings and a single dividend series with one exception. To compute quantiles for the $5-1$ row, dividends for both portfolios were included in the VAR.

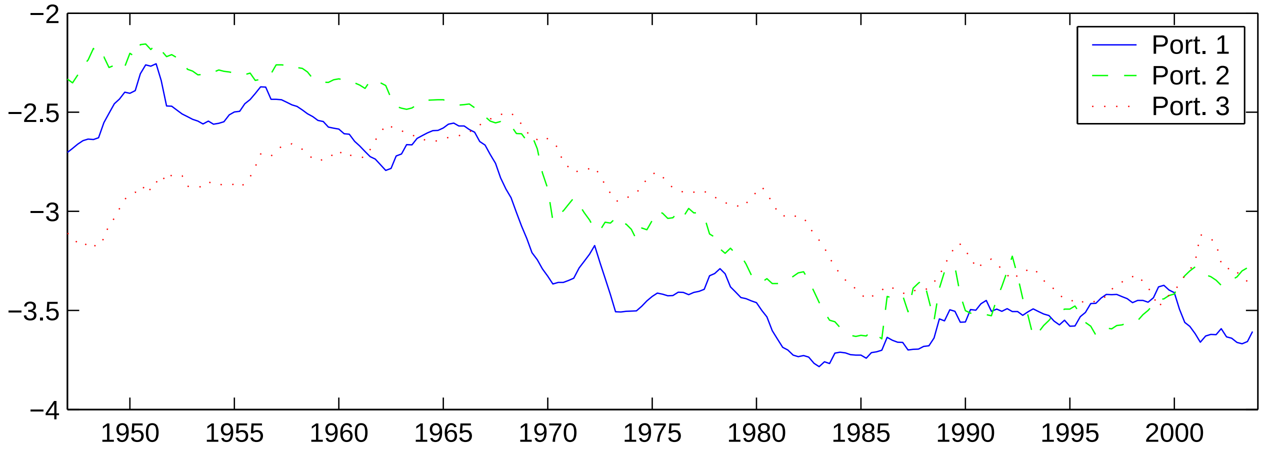
# 7 Alternative Models of Cash Flow Growth

Our calculations so far have been based on one model of cash flow growth. We now explore some alternative specifications used in other research and check for sensitivity. These specifications continue to capture the fact that 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.[15] Notice that the first three portfolios appear to grow slower than consumption, and even market dividends display this same pattern. Portfolios four and five show more pronounced growth than consumption.

## 7.1 Dividend Dynamics

In the previous section, we identified dividend dynamics and, in particular, the martingale component $\iota(1)$ using VAR methods. We used a VAR with three variables: consumption, corporate earnings and dividends (all in logarithms). Consumption and earnings were restricted to have the same long-run response to permanent shocks. We now consider two

[15] 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.



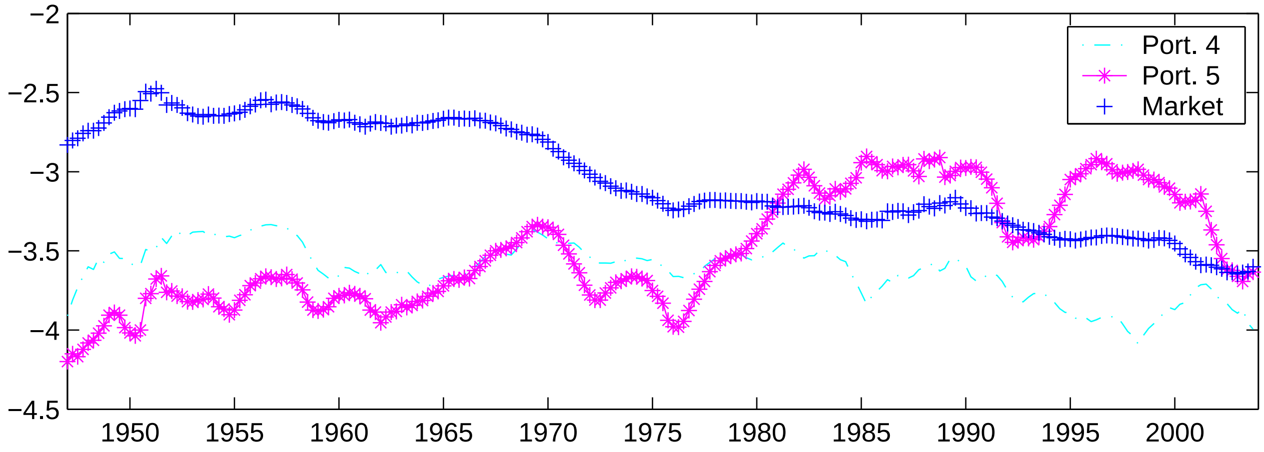




Figure 6: *Log of Ratios of Portfolio Dividends to Consumption*

alternative specifications of dividend growth to assess sensitivity to model specification. Both are restrictions on the equation:

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

where the shock $w_t^*$ is scalar with mean zero and unit variance and uncorrelated with the shock vector $w_t$ that enters (9). The third entry of $A_0^*$ is normalized to be positive. As in our previous estimation, we set $\ell = 5$.

## 7.2 Cointegration

The first specification restricts that the trend coefficient $B_1^*$ equal zero, and is the model used by Hansen, Heaton, and Li (2005). Given our interest in measuring long-run risk, we measure the permanent response of dividends to the permanent shock. While both consumption and corporate earnings are restricted to respond to permanent shocks in the same manner, the dividend response is left unconstrained. We let $\lambda^*$ denote the ratio of the long-run dividend response to the long-run consumption response. We measure this for each of the five portfolios. In this case we allow the matrix:

$$\begin{bmatrix} A(1) & 0 \\ \multicolumn{2}{c}{A^*(1)} \end{bmatrix}$$

to have rank two where

$$A^*(z) \doteq \sum_{j=0}^{\ell} A_j^* z^j .$$

The cointegrating vector $(1, 1, \lambda^*)$ is in the null space of this rank two matrix. For this model, the vector $\pi$ is

$$\pi = \iota(1) = \lambda^* \gamma(1) \tag{18}$$

and $\zeta = \mu_d = \lambda^* \mu_c$.

The second specification includes a time trend by freely estimating $B_1^*$. A model like this, but without corporate earnings, was used by Bansal, Dittmar, and Lundblad (2005). We refer to this as the time trend specification. In this model the time trend introduces a second source of dividend growth. While $\pi$ is constructed as in model (18), $\mu_d = \zeta$ is now left unrestricted.

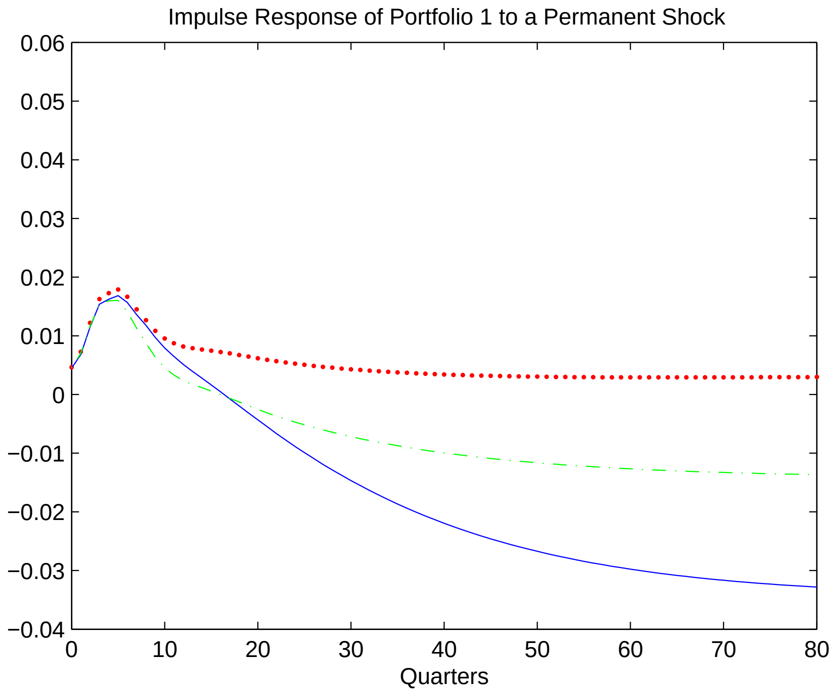
The role of specification uncertainty is illustrated in the impulse response figure 7. This figure features the responses of portfolio one and five to a permanent shock. For each portfolio, the measured responses obtained for each of the three growth configurations are quite close up to about three to four years and then they diverge. Both portfolios initially respond positively to this shock with peak responses occurring in about seven time periods. The response of portfolio one is much larger in this initial phase. The limiting responses differ substantially depending on the growth configuration that is imposed in estimation. The estimated response of portfolio one is eventually negative when time trends are included or an additional stochastic growth factor is included. The time trend model leads to lower limits for both portfolios. It is interesting, however, that the long-run differential responses between portfolio one and five are approximately the same for the time trend model and the dividend growth model.[16]

To better understand the importance of alternative growth configurations, figure 8 plots both the level of dividends for portfolios one and five and the fitted values implied by the "aggregate" innovations to consumption and corporate earnings alone. Results are reported for all three growth configurations. The presence of a deterministic trend in a log levels specification allows the VAR model to fit the low frequency movements of dividends for portfolio 1 much better than either of the other two models.[17] In contrast the fitted values are quite similar across growth configurations for portfolio 5.

[16]Bansal, Dittmar, and Lundblad (2005) use their estimates with a time trend model as inputs into a cross sectional return regression. While estimation accuracy and specification sensitivity may challenge these regressions, the consistency of the ranking across methods is arguably good news, as emphasized to us by Ravi Bansal. As is clear from our previous analysis, we are using the economic model in a more formal way and in way that departs in a substantial way from running cross-sectional regressions.

[17]Results for portfolio 2 are very similar to those for portfolio 1.

**Impulse Response Functions for Two Portfolios**



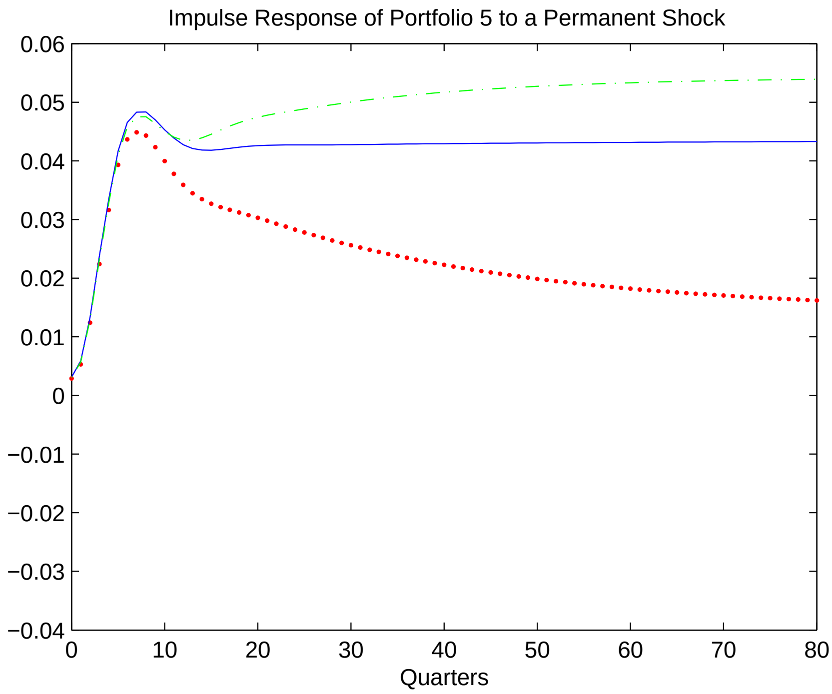




Figure 7: The · · · curve is generated from the level specification for dividends; the — is generated from the level specification with time trends included; and the -·- curve is generated from the first difference specification.

**Portfolio Dividends and Fitted Values**

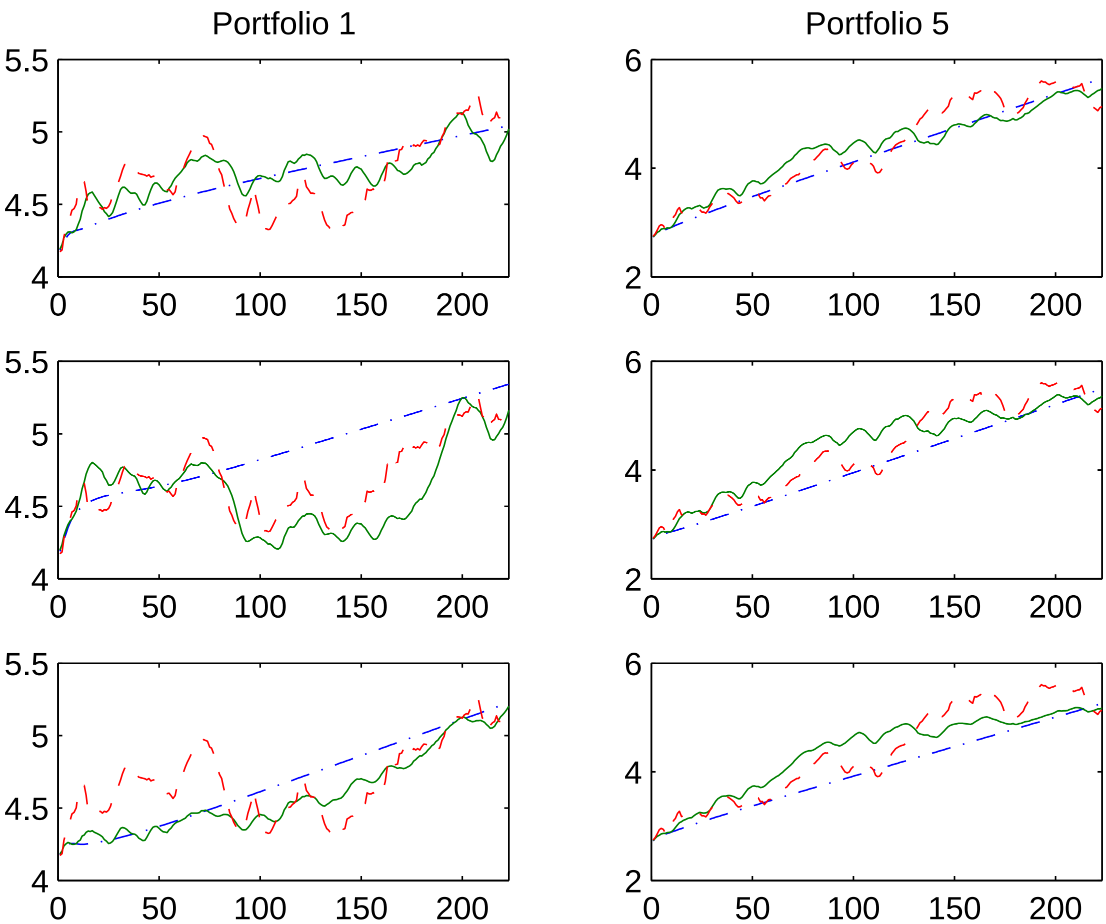



Figure 8: Dashed lines −− display the data. Solid lines — are the fitted values based on consumption shocks alone. Dot-dashed lines −· are fitted values with all shocks set to zero. Row one gives results for the cointegrated model without time trends, row two for the cointegrated model with time trends, and row three for the model in which an additional unit root is imposed on the dividend evolution.

Up until now, we have taken the linear cointegration model with time trends literally. 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 $\lambda^*$ 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 cointegrating relation is hard to defend.

There is a potential pitfall in estimation methods that 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 over fit the the initial time series, ascribing it to a transient component far from the trend line. As Sims argues,

> ... that the estimated model implies that future deviations as great as the initial deviation will be extremely rare.

This impact is evident for portfolio 1 as seen in figure 8. This figure includes trajectories simulated from the initial conditions alone. When the time trend is included, the *deterministic* simulation tracks well the actual dividend data for the first few years. There is sharp upward movement in the initial phase of this deterministic simulation when a time trend is included in the dividend evolution. The increase is much more muted when time trends are excluded.[18] Did investors have confidence at the beginning of the sample in such a trajectory? We suspect not. In contrast, this phenomenon is not present in deterministic simulation for portfolio 5. Instead the deterministic trajectory is very similar across the three time series models.

In summary, while there is intriguing heterogeneity in the long run cash flow responses and implied returns, the implied risk measures are sensitive to the growth specification as is the case in the related macroeconomics literature. Given the observed cash flow growth, it is important to allow for low frequency departures from a balanced growth restriction. The simple cointegration model introduces only one free growth parameter for each portfolio, but results in a modest amount of return heterogeneity. The time trend growth models impose additional sources of growth. The added flexibility of the time trend specification may presume too much investor confidence in a deterministic growth component, however. The dividend growth specification that we used in our previous calculations, while *ad hoc*, presumes this additional growth component is stochastic and is a more appealing specification to us.

[18] Again portfolio 2 behaves similarly to portfolio 1.

## 7.3 Adding Price Information

In the specifications we have considered so far, we have ignored any information for forecasting future consumption that might be contained in asset prices. Our model of asset pricing implies a strict relationship between cash flow dynamics and prices so that price information should be redundant. Prices, however, may reveal additional components to the information set of investor and hence a long-run consumption risk that cannot be identified from cash flows. For these reasons we consider an alternative specification of the VAR where we include consumption, corporate earnings, dividends as well as prices.

Parker and Julliard (2004) argue that it is the differential ability of the returns to growth and value portfolios in forecasting future consumption that is an important feature in the data. We therefore include dividends and prices for portfolios one and five simultaneously in this analysis. We continue to impose a unit root in consumption and the restriction that consumption and corporate earnings are cointegrated. We allow each dividend series to have its own stochastic growth path, but the prices of each portfolio are assumed to be cointegrated with their corresponded dividends. Finally, to assess the ability of portfolio prices to forecast future consumption we relax the assumption that consumption and corporate earnings are not Granger caused by portfolio cash flows or prices.

Figure 9 reports results for excess returns by horizon as in figure 4. The general character of the results are not changed. For large values of $\theta$ the model predicts substantial differences between portfolio excess returns at long-horizon. The exact patterns are different when prices are included, however. For example the excess returns to portfolio one, when $\theta = 20$, are larger at long horizons. Further there is more sensitivity to the parameter $\rho$ when $\theta = 20$.

**Excess Return Decompositions by Horizon**

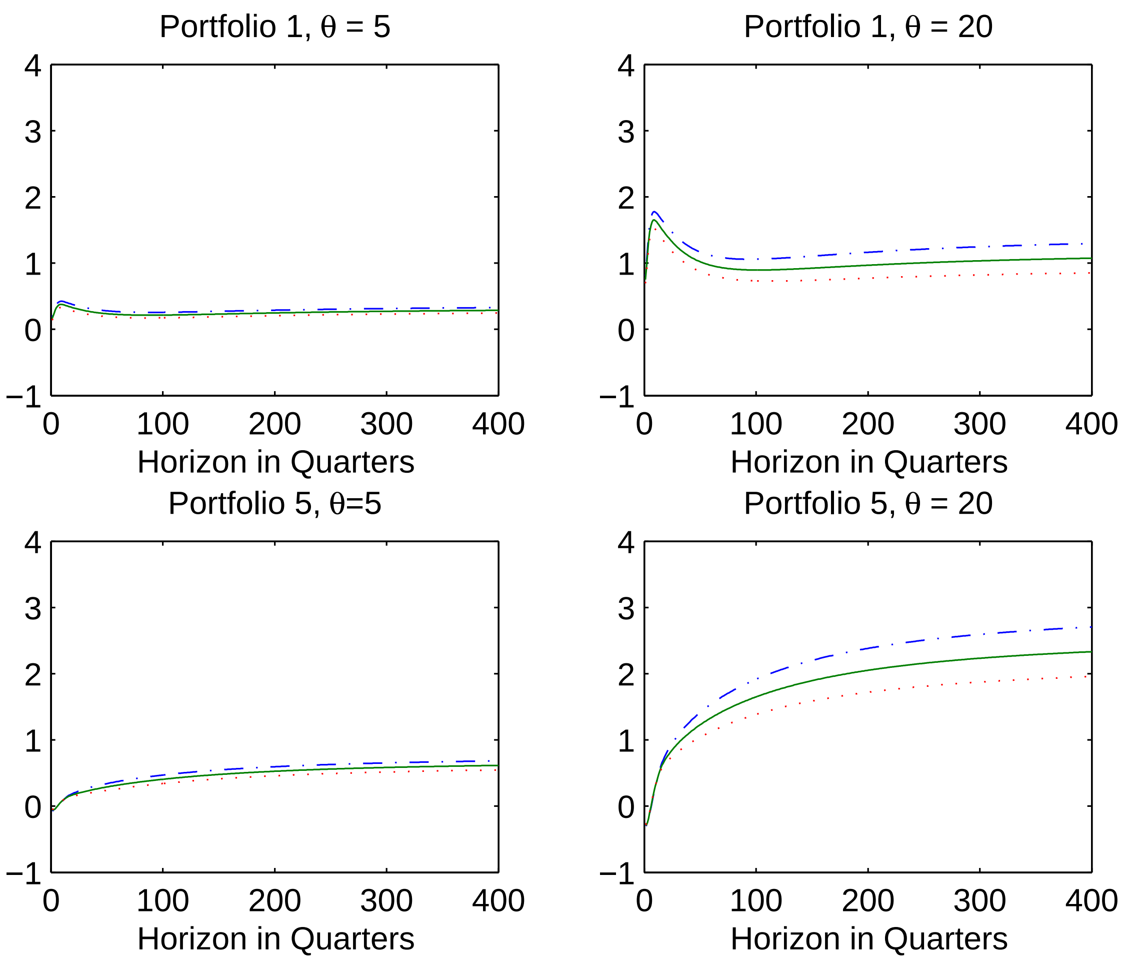



Figure 9: In the top two panels, the __ curves impose $\rho = 1$, the ·· curves impose $\rho = .5$ and the −. curves impose $\rho = 1.5$. The curves for $\rho \neq 1$ were computed using linear approximation around the point $\rho = 1$.

# 8 Conclusion

Growth rate variation in consumption or cash flows have important consequences for asset valuation. In this paper we analyzed formally the valuation implications through the guises of a commonly used consumption-based benchmark model. We made operational a notion of long-run valuation risk, and we studied measurement accuracy of the inputs needed to characterize the implied risk-return tradeoff. The methods on display in this paper produced a useful formalization of the long-run contribution to value of the stochastic components of discount factors and cash flows. We used them to isolated features of the economic environment that have important consequences for long run valuation and heterogeneity across cash flows.

Important inputs into our calculations are the long-run riskiness of cash flows and consumption. While these are crucial inputs, they are hard to measure in practice. Using standard methods from linear time series analysis to measure directly long-run variation in growth rates is a challenging endeavor. 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 hard to defend on purely statistical grounds.

The linear models we investigate are likely to be misspecified. For simplicity we closed down one potentially important channel for long-run risk by abstracting from volatility changes. These changes can induce an additional source of risk, but also pose additional statistical challenges of how to model and measure this volatility in a flexible way. While the direct evidence from consumption data for time varying volatility is modest, the implied evidence from asset pricing is intriguing.[19] Moreover, 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 about investor beliefs. Many of the statistical challenges that plague econometricians presumably also plague market participants. Naive application of rational expectations equilibrium concepts may endow investors with too much knowledge about future growth prospects.

There are two complementary responses to the measurement and modelling 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. While we explored one model-based method for extracting economic characterizations of this relation, we resorted in part to high risk aversion to produce heterogeneity in the dominant valuation components to portfolio cash flows. 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. While the methods we have proposed aid in our understanding of asset-pricing models, they

[19]For example, see Lettau and Wachter (2005) for a long-run risk characterization that features the consequences of heteroskedasticity using an *ad hoc* stochastic discount factor model.

also expose the important measurement challenges in quantifying the long-run risk-return tradeoff. Much progress has been made in our understanding of models, but there remain important challenges in understanding the precise nature of long-run growth rate movements in the underlying economy.

# 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 $\rho = 1$. To construct an appropriate recursion for $Dv_t^1$ we construct an approximate recursion by expanding the logarithm and exponential functions in (4) and including up to second-order terms in $\mathcal{Q}_t$. The approximate recursion is:

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

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

$$\begin{aligned} Dv_t^1 &= -\beta(1-\beta) \frac{\mathcal{Q}_t(v_{t+1}^1 + c_{t+1} - c_t)^2}{2} + \beta \tilde{E}(Dv_{t+1}^1 | \mathcal{F}_t) \\ &= -\frac{(1-\beta)(v_t^1)^2}{2\beta} + \beta \tilde{E}(Dv_{t+1}^1 | \mathcal{F}_t) \end{aligned} \tag{19}$$

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 3.1. Then

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

Write:

$$Dv_t^1 = -\frac{1}{2} {x_t}' \Upsilon_{dv} x_t + U_{dv} x_t + \mu_{dv}.$$

From (19),

$$\begin{aligned} \Upsilon_{dv} &= \frac{(1-\beta)}{\beta} U_v' U_v + \beta G' \Upsilon_{dv} G \\ U_{dv} &= -\frac{(1-\beta)}{\beta} \mu_v U_v - \beta(1-\theta)\gamma(\beta) H' \Upsilon_{dv} G + \beta U_{dv} G \\ \mu_{dv} &= -\frac{(1-\beta)}{2\beta}(\mu_v)^2 - \frac{\beta(1-\theta)^2}{2} \gamma(\beta) H' \Upsilon_{dv} H \gamma(\beta)' \\ &\quad + \beta(1-\theta) U_{dv} H \gamma(\beta)' - \frac{\beta}{2} \text{trace}(H' \Upsilon_{dv} H) + \beta \mu_{dv} \end{aligned} \tag{20}$$

The first equation in (20) is a Sylvester equation and is easily solved. Given $\Upsilon_{dv}$, the solution for $U_{dv}$ is:

$$U_{dv} = -(I - \beta G')^{-1} \left[ \frac{1-\beta}{\beta} \mu_v U_v + \frac{\beta(1-\theta)}{2} G' \Upsilon_{dv} H \gamma(\beta)' \right],$$

and given $\Upsilon_{dv}$ and $U_{dv}$ the solution for $\mu_{dv}$ is:

$$\mu_{dv} = \frac{-\frac{(1-\beta)}{2\beta}(\mu_v)^2 - \frac{\beta(1-\theta)^2}{2}\gamma(\beta) H' \Upsilon_{dv} H \gamma(\beta)' + \beta(1-\theta) U_{dv} H \gamma(\beta)' - \frac{\beta}{2}\text{trace}(H' \Upsilon_{dv} H)}{1-\beta}$$

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

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

Recall that the log discount factor is given by:

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

Differentiating with respect to $\rho$ gives:

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

Note that

$$\begin{aligned} v^1_{t+1} - \frac{1}{\beta} v^1_t &= U_v x_{t+1} - \frac{1}{\beta} U_v x_t + \left(1 - \frac{1}{\beta}\right)\mu_v \\ &= U_v\left(G - \frac{1}{\beta} I\right) x_t + \left(1 - \frac{1}{\beta}\right)\mu_v + U_v H w_{t+1}. \end{aligned}$$

and

$$\begin{aligned} Dv^1_{t+1} - \tilde{E}\left(Dv^1_{t+1} | \mathcal{F}_t\right) &= -\frac{1}{2}(H w_{t+1})' \Upsilon_{dv} H w_{t+1} - (H w_{t+1})'[\Upsilon_{dv} G x_t - U_{dv}] \\ &\quad + \frac{1}{2}(1-\theta)^2 \gamma(\beta) H' \Upsilon_{dv} H \gamma(\beta)' + (1-\theta)\gamma(\beta) H' [\Upsilon_{dv} G x_t - U_{dv}] \\ &\quad + \frac{1}{2}\text{trace}(H' \Upsilon_{dv} H) \end{aligned}$$

Combining these expressions we obtain:

$$Ds^1_{t+1,t} = \frac{1}{2} {w_{t+1}}' \Theta_0 w_{t+1} + {w_{t+1}}' \Theta_1 x_t + \vartheta_0 + \vartheta_1 x_t + \vartheta_2 w_{t+1}$$

where

$$\Theta_0 = (\theta - 1) H' \Upsilon_{dv} H$$

$$
\begin{aligned}
\Theta_1 &= (\theta-1)H'\Upsilon_{dv}G \\
\vartheta_0 &= \left(1-\frac{1}{\beta}\right)\mu_v + \frac{1}{2}(1-\theta)^3\gamma(\beta)H'\Upsilon_{dv}H\gamma(\beta)' - (1-\theta)^2\gamma(\beta)H'U_{dv}' + \frac{1-\theta}{2}\text{trace}(H'\Upsilon_{dv}H) \\
\vartheta_1 &= U_v\left(G-\frac{1}{\beta}I\right) + (\theta-1)^2\gamma(\beta)H'\Upsilon_{dv}G \\
\vartheta_2 &= (1-\theta)U_{dv}H + U_vH
\end{aligned}
$$

The mean under the risk neutral measure for $w_{t+1}$ is

$$
\begin{aligned}
[I \quad + \quad & (\rho-1)(\theta-1)H'\Upsilon_{dv}H]^{-1} \\
& [-\rho\gamma(0) + (\rho-\theta)\gamma(\beta) + (\rho-1)(\theta-1)\left(U_{dv} - H'\Upsilon_{dv}Gx_t\right)].
\end{aligned}
$$

This mean can be interpreted as the negative of a risk premia. A component of this mean is the undiscounted (by the risk free rate) price an investor is willing to pay for contingent claim to the corresponding component of the shock $w_{t+1}$. In a continuous time approximatoin, this formula simplifies to the one reported in the paper.

# B Calculating Eigenvalues

## B.1 Proof of Theorem 5.1

Consider the first-order autoregressive specification in example 3.1:

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

where $G$ has eigenvalues with absolute values that are strictly less than one. The consumption dynamics are:

$$c_{t+1} - c_t = \mu_c + U_cx_t + \gamma_0 w_{t+1}$$

First we study the $\rho = 1$ benchmark. For simplicity, we set $\zeta = 0$, since a nonzero $\zeta$ has a known impact on the eigenvalue and no impact on the eigenfunction. Write:

$$s^1_{t+1,t} + \pi w_{t+1} = \xi_0 + \xi_1 x_t + \xi_2 w_{t+1}$$

where

$$
\begin{aligned}
\xi_0 &\doteq -\delta - \mu_c - \frac{(1-\theta)^2|\gamma(\beta)|^2}{2} \\
\xi_1 &\doteq -U_c \\
\xi_2 &\doteq (1-\theta)\gamma(\beta) - \gamma_0 + \pi \\
\gamma(\beta) &= \gamma_0 + \beta U_c(I-\beta G)^{-1}H.
\end{aligned}
$$

### B.1.1 Dominant Eigenfunction and Eigenvalue

We seek an eigenfunction that is log-linear:

$$\log \phi(x) = -\omega x.$$

This eigenfunction should satisfy the equation:

$$E\left[\exp[s^1_{t+1,t} + \pi w_{t+1}]\phi(x_{t+1})|x_t\right] = \exp(-\nu)\phi(x_t).$$

where $\exp(-\nu)$ is the eigenvalue associated with eigenfunction $\psi(x_t)$. For a log linear eigenfunction to exist, the vector $\omega$ necessarily satisfies:

$$\xi_1 - \omega G = -\omega,$$

which implies that

$$\omega = \xi_1(G - I)^{-1} = U_c(I - G)^{-1}.$$

The negative logarithm of the eigenvalue is

$$\nu = -\xi_0 - \frac{|\xi_2 - \omega H|^2}{2}.$$

Plug in the formulas for $\xi_0$, $\xi_1$ and $\xi_2$, and add back $\zeta$, the we have

$$\nu = \delta + \mu_c - \zeta - \frac{|\pi - \gamma(\beta)|^2}{2} + (\theta - 1)\gamma(\beta) \cdot [\pi - \gamma(1)]\,.$$

### B.1.2 Eigenfunction for adjoint of the pricing operator

Next we need to compute an eigenfunction for the adjoint of the pricing operator. Guess an eigenfunction of the form:

$$\log \varphi(x) = -\omega^* x$$

then this eigenfunction should satisfy the equation:

$$E\left[\exp[s^1_{t+1,t} + \pi w_{t+1}]\varphi(x_t)|x_{t+1}\right] = \exp(-\nu^*)\varphi(x_t).$$

where $\exp(-\nu^*)$ is the eigenvalue associated with eigenfunction $\psi^*(x_t)$, and we will show later $\nu$ and $\nu^*$ are the same.

First compute the reverse time evolution of $x_t$,

$$x_t = G^* x_{t+1} + H^* w^*_t.$$

where $w^*_t$ is a multivariate standard normal, independent of $x_{t+1}$.

The matrix $G^*$ can be inferred by standard least squares formulas:

$$\Sigma G' \Sigma^{-1} = G^*.$$

and the matrix $H^*$ can be inferred by factoring:

$$\Sigma - G^* \Sigma G^{*\prime}.$$

where $\Sigma$ is the unconditional variance-covariance matrix of $x_t$.

Write:

$$w_{t+1} = (H'H)^{-1} H'(x_{t+1} - Gx_t) = (H'H)^{-1} H' \left[ (I - GG^*)x_{t+1} - GH^* w_t^* \right].$$

Thus:

$$s^1_{t+1,t} + \pi w_{t+1} = \xi_0^* + \xi_1^* x_{t+1} + \xi_2^* w_t^*$$

for

$$\begin{aligned}
\xi_0^* &= \xi_0 \\
\xi_1^* &= \xi_1 G^* + \xi_2 (H'H)^{-1} H' (I - GG^*) \\
\xi_2^* &= -\xi_2 (H'H)^{-1} H' G H^* + \xi_1 H^*.
\end{aligned}$$

Then the adjoint problem solves:

$$E \left( \exp \left[ \xi_0^* + \xi_1^* x_{t+1} + \xi_2^* w_t^* - \omega^* (G^* x_{t+1} + H^* w_t^*) \right] | x_{t+1} \right) = \exp(-\nu^*) \exp(-\omega^* x_{t+1})$$

This problem has the same formal structure as the initial eigenvector problem. The solution is

$$\omega^* = \xi_1^* (G^* - I)^{-1} = U_c (I - G^*)^{-1}.$$

The negative logarithm of the eigenvalue is

$$\nu^* = -\xi_0^* - \frac{|\xi_2^* - \omega^* H^*|^2}{2}.$$

and it can be easily shown that $\nu$ and $\nu^*$ are the same.

### B.1.3 Expected valuation rate of return

As we have shown in section 5.4, the expected valuation rate of return is the sum of the decay rate of dividend growth and that of the pricing operator, $\epsilon + \nu$, where

$$\epsilon = \zeta + \frac{|\pi|^2}{2}$$

hence

$$\epsilon + \nu = \zeta + \frac{|\pi|^2}{2} - \xi_0 - \frac{|\xi_2 - \omega H|^2}{2}$$

plug in the formulas $\xi_0$, $\xi_2$ and $\omega$, and rearrage terms we have

$$\epsilon + \nu = \varsigma^* + \pi^* \cdot \pi$$

where

$$\begin{aligned}
\varsigma^* &= \gamma(1) + (\theta - 1)\gamma(\beta) \\
\pi^* &= \mu_c + \delta + (1 - \theta)\gamma(1) \cdot \gamma(\beta) - \frac{|\gamma(1)|^2}{2}
\end{aligned}$$

## B.2 Eigenvalue derivative

Consider next a derivative of the dominant eigenvalue with respect to $\rho$. Let $q$ denote the stationary density for $x_t$. This vector is normally distributed with mean zero and covariance matrix:

$$\Sigma = \sum_{j=0}^{\infty} (G^j) H H' (G^j)'.$$

We use the relation:

$$\exp(-\nu) = \frac{E\left[\exp(s_{t+1,t} + \pi w_{t+1}) \phi(x_{t+1}) \varphi(x_t)\right]}{E\left[\phi(x_t)\varphi(x_t)\right]}.$$

Write

$$Ds^1_{t+1,t} = \frac{1}{2} w_{t+1}{}' \Theta_0 w_{t+1} + w_{t+1}{}' \Theta_1 x_t + \vartheta_0 + \vartheta_1 x_t + \vartheta_2 w_{t+1}.$$

Then

$$\frac{d \exp(-\nu)}{d\rho} \mid_{\rho=1} = \frac{E\left[Ds^1_{t+1,t} \exp(s^1_{t+1,t} + \pi w_{t+1}) \phi(x_{t+1}) \varphi(x_t)\right]}{E[\phi(x_t)\varphi(x_t)]},$$

and hence

$$\frac{d\nu}{d\rho} \mid_{\rho=1} = -\frac{E\left[Ds^1_{t+1,t} \exp(s^1_{t+1,t} + \pi w_{t+1}) \phi(x_{t+1}) \varphi(x_t)\right]}{\exp(-\nu) E[\phi(x_t)\varphi(x_t)]}. \tag{21}$$

We take three steps to compute this eigenvalue derivative

### B.2.1 Step one: computing the denominator

We must compute:

$$E[\phi(x_t)\varphi(x_t)] = \int \exp[-(\omega + \omega^*) x] q(x) dx$$

From the lognormal formula, this is

$$\exp\left[\frac{(\omega + \omega^*)\Sigma(\omega + \omega^*)'}{2}\right].$$

### B.2.2 Step two: computing the numerator

We have already evaluated the denominator, but it remains to compute the numerator:

$$E\left[Ds^1_{t+1,t} \exp\left(s^1_{t+1,t} + \pi w_{t+1}\right) \phi(x_{t+1}) \varphi(x_t)\right]$$

We do so by applying the Law of Iterated Expectations, and first computing:

$$E\left[Ds^1_{t+1,t} \exp\left(s^1_{t+1,t} + \pi w_{t+1}\right) \phi(x_{t+1}) \varphi(x_t) | x_t\right].$$

Note that

$$
\begin{aligned}
& s^1_{t+1,t} + \pi w_{t+1} + \log[\phi(x_{t+1})] + \log[\varphi(x_t)] \\
= \; & \xi_0 + (\xi_1 - \omega G - \omega^*) x_t + (\xi_2 - \omega H) w_{t+1} \\
= \; & \left[ \xi_0 + \frac{|\xi_2 - \omega H|^2}{2} - (\omega + \omega^*) x_t \right] + \left[ (\xi_2 - \omega H) w_{t+1} - \frac{|\xi_2 - \omega H|^2}{2} \right].
\end{aligned}
$$

We use the second term in the square brackets to change the shock distribution. In particular, we change the mean of $w_{t+1}$ from zero to $[(\xi_2 - \omega H)]$. Thus

$$
\begin{aligned}
& E\left[ D s^1_{t+1,t} \exp\left( s^1_{t+1,t} + \pi w_{t+1} \right) \phi(x_{t+1}) \varphi(x_t) | x_t \right] \\
= \; & \exp\left[ \xi_0 + \frac{|\xi_2 - \omega H|^2}{2} - (\omega + \omega^*) x_t \right] \\
& \times \left[ \frac{1}{2} \text{trace}(\Theta_0) + (\xi_2 - \omega H)' \Theta_1 x_t + \vartheta_0 + \vartheta_1 x_t + \vartheta_2 (\xi_2 - \omega H) \right] \\
= \; & \exp(-\nu) \exp\left[ -(\omega + \omega^*) x_t \right] \\
& \times [(1/2) \text{trace}(\Theta_0) + (1/2)(\xi_2 - \omega H)' \Theta_0 (\xi_2 - \omega H) \\
& + (\xi_2 - \omega H)' \Theta_1 x_t + \vartheta_0 + \vartheta_1 x_t + \vartheta_2 (\xi_2 - \omega H)]
\end{aligned}
$$

Next we compute the unconditional expectation. Again we change probability distributions. To simply the calculation, we adopt a change in measure. We change the mean of $x_t$ from normal mean zero and covariance matrix $\Sigma$ to normal with mean

$$
\mu^*_x \doteq -\Sigma(\omega + \omega^*)'
$$

and covariance $\Sigma$. Using this transformation we find that

$$
\begin{aligned}
& E\left[ D s^1_{t+1,t} \exp\left( s^1_{t+1,t} + \pi w_{t+1} \right) \phi(x_{t+1}) \varphi(x_t) | x_t \right] \\
= \; & \exp(-\nu) \exp\left( \frac{\mu^{*\prime}_x \Sigma^{-1} \mu^*_x}{2} \right) \\
& \times [\frac{1}{2} \text{trace}(\Theta_0) + \frac{1}{2} (\xi_2 - \omega H)' \Theta_0 (\xi_2 - \omega H) \\
& + (\xi_2 - \omega H)' \Theta_1 \mu^*_x + \vartheta_0 + \vartheta_1 \mu^*_x + \vartheta_2 (\xi_2 - \omega H)].
\end{aligned}
$$

### B.2.3 Step three: combining results

We compute the right-hand side of (21) by combining numerator and denominator terms:

$$
\begin{aligned}
\frac{d\nu}{d\rho}|_{\rho=1} \; = \; & -\frac{1}{2} \text{trace}(\Theta_0) - \frac{1}{2} (\xi_2 - \omega H)' \Theta_0 (\xi_2 - \omega H) \\
& - (\xi_2 - \omega H)' \Theta_1 \mu^*_x - \vartheta_0 - \vartheta_1 \mu^*_x - \vartheta_2 (\xi_2 - \omega H).
\end{aligned}
$$

# References


Alvarez, F. and U. Jermann (2001). The size of the permanent component of asset pricing kernels. NBER working paper number 8360.

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 (2005). Consumption, dividends, and the cross-section of equity returns.

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 59*, 1481–1509.

Berk, K. N. (1974). Consistent autoregressive spectral estimates. *Annals of Statistics 2*, 289–502.

Blanchard, O. J. and D. Quah (1989). The dynamic effects of aggregate demand and supply disturbances. *American Economic Review 79*, 655–673.

Breeden, D. (1979). An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics 7*, 265–296.

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 (2005). Growth or glamour? fundamentals and systematic risk in stock returns. NBER working paper 11389.

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. NBER working paper 9509.

Campbell, J. Y. (1994). Intertemporal asset pricing without consumption data. *American Economic Review 83*, 487–512.

Campbell, J. Y. and L. M. Viceira (2002). *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*. Oxford, U. K.: Oxford University Press.

Cochrane, J. H. (1988). How big is the random walk in gnp. *Quarterly Journal of Economics 96*, 893–920.

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.

Hall, P. and C. C. Heyde (1980). *Martingale Limit Theory and its Application*. Academic Press.

Hansen, L. P., J. Heaton, and N. Li (2005). Intangible risk. In C. Corrado, J. Haltiwanger, and D. Sichel (Eds.), *Measuring Capital in the New Economy*. University of Chicago Press.

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 T. J. Sargent (1980). Formulating and estimating dynamic linear rational expectations models. *Journal of Economic Dynamics and Control 2*, 7 –46.

Hansen, L. P. and J. Scheinkman (2005). Semigroup pricing. University of Chicago.

Hansen, L. P. and J. A. Scheinkman (1995). Back to the future: Generating moment implications for continuous-time markov processes. *Econometrica 63*, 767–804.

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.

Kogan, L. and R. Uppal (2001). Risk aversion and optimal portfolio policies in partial and general equilibrium economies. Sloan School of Management and London Business School.

Kreps, D. M. and E. L. Porteus (1978). Temporal resolution of uncertainty and dynamic choice. *Econometrica 46*, 185–200.

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

Lettau, M. and J. Wachter (2005). Why is long-horizon equity less risky? a duration-based explanation of the value premium. NBER working paper number 11144.

Lettau, M. L., S. C. Ludvigson, and J. A. Wachter (2004). The declining equity premium: What role does macroeconomic policy play? New York University.

Lucas, R. E. (1978). Asset prices in an exchange economy. *Econometrica 46*, 1429–1445.

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 (2004). Consumption risk and cross-sectional returns. Princeton University.

Restoy, F. and P. Weil (1998, June). Approximate equilibrium asset prices. NBER Working Paper 6611.

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

Schroder, M. and C. Skiadas (1999). Optimal consumption and portfolio selection with stochastic differential utility. *Journal of Economic Theory 89*, 68–126.

Sims, C. A. (1991). Comment on 'to criticize the critics'. *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.

Sims, C. A. and T. Zha (1999). Error bands for impulse responses. *Econometrica 67*, 1113–1156.

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.

Zha, T. (1999). Block recursion and structural vector autoregression. *Journal of Econometrics 90*, 291–316.