NBER WORKING PAPER SERIES

THE SIZE OF THE PERMANENT COMPONENT OF ASSET PRICING KERNELS

Fernando Alvarez
Urban J. Jermann

Working Paper 8360
http://www.nber.org/papers/w8360

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2001

We thank Andy Atkeson, Lars Hansen, Pat Kehoe, Bob King, Narayana Kocherlakota, Lee Ohanian, and the participants in workshops at UCLA, the University of Chicago, the Federal Reserve Banks of Minneapolis, Chicago, and Cleveland, and Duke, Boston and Ohio State Universities for their comments and suggestions. We thank Robert Bliss for providing the data for U.S. zero-coupon bonds. Sebastian Bourgeois gave us excellent research assistance. Alvarez thanks the NSF and the Sloan Foundation for support. The latest version of this paper is available at http://www.finance.wharton.upenn.edu/~jermann/research.html. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research, the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

The Size of the Permanent Component of Asset Pricing Kernels
Fernando Alvarez and Urban J. Jermann
NBER Working Paper No. 8360
July 2001
JEL No. G12, E43, E21

ABSTRACT

We derive a lower bound for the size of the permanent component of asset pricing kernels. The bound is based on return properties of long-term zero-coupon bonds, risk-free bonds, and other risky securities. We find the permanent component of the pricing kernel to be very large; its volatility is about 100% of the volatility of the stochastic discount factor. This result implies that, if the pricing kernel is a function of consumption, innovations to consumption need to have permanent effects.

Fernando Alvarez
Department of Economics
University of Chicago
Social Science Bldg 422
1126 East 59th St.
Chicago, IL 60637
and NBER
f-alvarez1@uchicago.edu

Urban J. Jermann
Research Department
Federal Reserve Bank of Minneapolis
90 Hennepin Avenue
Minneapolis, MN 55480-0291
and Wharton School,
University of Pennsylvania
and NBER
jermann@wharton.upenn.edu

1. Introduction

Under fairly general conditions, the absence of arbitrage opportunities implies the existence of a *pricing kernel*, that is, a stochastic process that assigns values to future state-contingent payments. Knowing the properties of such processes is important for asset pricing, and it has been the focus of much recent research.[1] Given that many securities are long-lived, the low-frequency or long-term properties of pricing kernels are important determinants of their prices.

In this paper, we present and estimate a lower bound for the size of the permanent component of asset pricing kernels. The bound is based on return properties of long-term zero-coupon bonds, risk-free bonds, and other risky securities. We find the permanent component of the pricing kernel to be very large; its volatility is about 100% of the volatility of the stochastic discount factor. This result should provide guidance for the specification of asset pricing models, particularly, if the objective is to price long-lived securities.

Our results are related to the work by Hansen and Jagannathan (1991). They use no-arbitrage conditions to derive bounds on the volatility of pricing kernels as a function of observed asset prices. An important lesson from their research is that, in order to explain the equity premium, pricing kernels have to be very volatile. Our bound for the permanent component of the pricing kernel complements this approach. We find that, because term spreads for long-term bonds are so small relative to the excess returns on equity, the permanent component of the pricing kernel has to be very large.

Asset pricing models link pricing kernels to the underlying economic fundamentals. Thus, our analysis provides some insights into the long-term properties of these fundamentals and into the functions linking pricing kernels to the fundamentals. Along this dimension, we have two sets of results.

First, under some assumptions about the function of the marginal utility of wealth, we derive sufficient conditions on consumption so that the pricing kernel has no permanent innovations. We present several examples of utility functions for which the existence of an invariant distribution of consumption implies pricing kernels with no permanent innovations. Thus, these examples are inconsistent with our main findings. This result is useful for macroeconomics, because for some

[1] A few prominent examples of research in this line are Hansen and Jagannathan (1991), Snow (1991), Cochrane and Hansen (1992), Luttmer (1996), and Backus, Foresi, and Telmer (1998).

questions, the persistence properties of the processes specifying economic variables matter a great deal. Specifically, for processes with highly persistent innovations, small changes in the degree of persistence can generate large differences in the answers to quantitative questions. For instance, on the issue of the welfare costs of economic uncertainty, see Dolmas (1998) and Alvarez and Jermann (2000a); on the issue of the volatility of macroeconomic variables such as consumption, investment, and hours worked, see Hansen (1997); and on the issue of international business cycle comovements, see Baxter and Crucini (1995). On a related matter, Nelson and Plosser (1982) argue that many macroeconomic time-series are characterized by nonstationary instead of stationary processes. They interpret their findings as implying that stochastic variations due to real factors are essential in explaining macroeconomic fluctuations rather than monetary disturbances, which are assumed to have short term effects. A large body of literature has developed statistical tools to address the question of stationarity versus unit roots and to measure the size of the permanent component. The fact that most economic time-series are relatively short has been a challenge for that literature.[2] Our results complement the direct statistical analysis of macroeconomic time-series by using, among other things, the information contained in long-term bonds about how asset markets forecast long-term changes in the pricing kernel.

Second, measuring the size of the permanent component in consumption directly and comparing it to the size of the permanent component of pricing kernels provides guidance for the specification of functions of the marginal utility of wealth.[3] Specifically, we find the size of the permanent component of consumption to be lower than that of pricing kernels. This suggests the use of utility functions that magnify the permanent component.

The rest of the paper is structured as follows. Section 2 contains definitions and theoretical results. Section 3 presents empirical evidence. Section 4 concludes. Proofs are in Appendix A. Appendix B describes the data sources. Appendix C addresses a small sample bias.

---

[2] See, for instance, Hamilton (1994).

[3] See Daniel and Marshall (2001) on the related issue of how consumption and asset prices are correlated at different frequencies.

2. Definitions and Theoretical Results

Here we present our theoretical results. We start by stating some results about long-term discount bonds. Specifically, we present an inequality linking the term spread of interest rates to the excess returns on any security. This inequality holds for pricing kernels that have no permanent innovations. We then consider the case of a pricing kernel whose innovations have permanent and transitory components, and we present a lower bound for the size of the permanent component. We show how to interpret this lower bound for some classes of lognormal processes. Our second set of results extends the characterization of the stochastic process of pricing kernels to the properties of their determinants; specifically, consumption.

Let D_{t+k} be a state-contingent dividend to be paid at time $t+k$ and $V_t[D_{t+k}]$ be the current price of a claim to this dividend. Then, as can be seen, for instance, in Duffie (1996), arbitrage opportunities are ruled out in frictionless markets if and only if a strictly positive *pricing kernel* or state-price process, $\{M_t\}$, exists so that

$$V_t[D_{t+k}] = \frac{E_t[M_{t+k} \cdot D_{t+k}]}{M_t}.[4] \tag{2.1}$$

For our results, it is important to distinguish between the pricing kernel, M_{t+1}, and the *stochastic discount factor,* M_{t+1}/M_t.[5] We use R_{t+1} for the gross return on a generic portfolio held from t to $t+1$; hence,(2.1) implies that

$$1 = E_t\left[\frac{M_{t+1}}{M_t} \cdot R_{t+1}\right]. \tag{2.2}$$

We define $R_{t+1,k}$ as the gross return from holding from time t to time $t+1$ a claim to one unit of the numeraire to be delivered at time $t+k$,

$$R_{t+1,k} = \frac{V_{t+1}\left(1_{t+k}\right)}{V_t\left(1_{t+k}\right)}.$$

[4]As is well known, this result does not require complete markets, but assumes that portfolio restrictions do not bind for some agents. This last condition is sufficient, but not necessary, for the existence of a pricing kernel. For instance, in Alvarez and Jermann (2000b), portfolio restrictions bind most of the time; nevertheless, a pricing kernel exists that satisfyies (2.1).

[5]For instance, in the Lucas representative agent model, the pricing kernel M_t is given by $\beta^t U'\left(c_t\right)$, where β is the preference time discount factor and $U'\left(c_t\right)$ is the marginal utility of consumption. In this case, the stochastic discount factor, M_{t+1}/M_t, is given by $\beta U'\left(c_{t+1}\right)/U'\left(c_t\right)$.

The holding return on this discount bond is the ratio of the price at which the bond is sold, $V_{t+1}(1_{t+k})$, to the price at which it was bought, $V_t(1_{t+k})$. With this convention, $V_t(1_t) \equiv 1$. Thus, for $k \geq 2$ the return consists solely of capital gains; for $k = 1$, the return is risk free. Finally, we define the continuously compounded *term premium* for a k-period discount bond as

$$h_t(k) \equiv E_t \left\{ \log \left[\frac{R_{t+1,k}}{R_{t+1,1}} \right] \right\},$$

that is, the expected log excess return on the k-period discount bond.

Based on these definitions, simple algebra allows us to write the term premium as having two components:

$$h_t(k) = \{\log E_t M_{t+1} - E_t \log M_{t+1}\} + E_t \left\{ \log \frac{E_{t+1}[M_{t+k}]}{E_t[M_{t+k}]} \right\}.$$

The first component in braces depends only on kernels dated $t+1$, while the last term contains the dependence on $t+k$. We now define a condition for pricing kernels that turns out to be key for the properties of long-term bonds.

Definition 2.1. *We say that a pricing kernel has no permanent innovations at* t*, if*

$$\lim_{k\to\infty} E_t \left\{ \log \frac{E_{t+1}[M_{t+k}]}{E_t[M_{t+k}]} \right\} = 0. \tag{2.3}$$

Under a set of regularity conditions presented in Appendix A, this definition is equivalent to assuming that

$$\lim_{k\to\infty} \frac{E_{t+1}[M_{t+k}]}{E_t[M_{t+k}]} = 1 \tag{2.4}$$

in distribution. This can be seen intuitively by using Jensen's inequality and the law of iterative expectations:

$$\lim_{k\to\infty} E_t \left\{ \log \frac{E_{t+1}[M_{t+k}]}{E_t[M_{t+k}]} \right\} \leq \lim_{k\to\infty} \log E_t \left\{ \frac{E_{t+1}[M_{t+k}]}{E_t[M_{t+k}]} \right\} = \log(1).$$

Thus, condition (2.3) can only be satisfied if the ratio of expectations converges to its (constant) mean. We say that there are no permanent innovations because, as the forecasting horizons k become longer, information arriving at $t+1$ will not lead to revisions of the forecasts made with current period t information. Alternatively, condition (2.3) says that innovations in the forecasts of the pricing kernel have limited persistence, since their effect vanishes for large k. Formally, we will use the definition in condition (2.3) because it requires no further auxiliary assumptions; it also turns out to be easier to check in our examples.

The following proposition states an important result for zero-coupon bonds if pricing kernels have no permanent component.

Proposition 2.2. *If a pricing kernel has no permanent innovations, then*

$$h_t(\infty) \equiv \lim_{k\to\infty} h_t(k) \geq E_t\left[\log\left(\frac{R_{t+1}}{R_{t+1,1}}\right)\right], \tag{2.5}$$

where R_{t+1} is the holding return on any asset.

Proposition 2.2 states that without permanent innovations, the term spread is the highest (log) risk premium. Notice that the portfolio with the highest $E_t[\log(R_{t+1})]$ is known as the *growth optimal* portfolio.

We present here an intuitive proof of Proposition 2.2 that uses the slightly stronger notion of *no permanent innovations* than the one defined in condition (2.3). A formal proof of Proposition 2.2 is in Appendix A.

The holding return to a k-period discount bond can be written as

$$R_{t+1,k} = \frac{V_{t+1}(1_{t+k})}{V_t(1_{t+k})} = \frac{M_t}{M_{t+1}} \cdot \frac{E_{t+1}[M_{t+k}]}{E_t[M_{t+k}]}.$$

Thus, the return depends on the stochastic discount factor between time t and $t+1$ and on the innovations about the future in k periods. Using the slightly stronger version of *no permanent innovations* as defined in equation (2.4), namely,

$$\lim_{k\to\infty} \frac{E_{t+1}[M_{t+k}]}{E_t[M_{t+k}]} = 1,$$

we can write the limiting holding return as

$$R_{t+1,\infty} = \frac{M_t}{M_{t+1}}.$$

Therefore, the price of the long-term bond is risky solely because of the effect from the valuation at $t+1$. There are no innovations about the far future. Clearly, then, $R_{t+1,\infty}$ commands a positive risk premium because its return is negatively correlated with the stochastic discount factor M_{t+1}/M_t. It remains to be shown why there can be no return that commands an even higher risk premium. Specifically, the issue at hand is to find the distribution of the return that yields the highest expected log return subject to satisfying the Euler equation, that is,

$$\max_{R_{t+1}} E_t \log R_{t+1}, \text{ subject to } E_t\left(\frac{M_{t+1}}{M_t}R_{t+1}\right) = 1.$$

From the concavity of the log, we can write

$$E_t \log\left(\frac{M_{t+1}}{M_t}R_{t+1}\right) \leq \log E_t\left(\frac{M_{t+1}}{M_t}R_{t+1}\right) = \log(1) = 0. \tag{2.6}$$

Moreover, the left-hand side achieves the highest value, that is, 0, for a return for which $\frac{M_{t+1}}{M_t}R_{t+1}$ is constant. With $R_{t+1,\infty} = \frac{M_t}{M_{t+1}}$, this condition is satisfied for the asymptotic bond. Because the log is additive and $\log\frac{M_{t+1}}{M_t}$ is common to all returns, $R_{t+1,\infty}$ is the highest expected log return of all assets. A return with an even higher variance will not lead to an increase because the concavity of the log reduces the mean on the left hand-side of equation (2.6).

Proposition 2.2 essentially restates results presented in earlier studies in such a way as to allow for our subsequent extensions. Kazemi (1992) shows that in a Markov economy with a limiting stationary distribution, the return on the discount bond with the longest maturity equals the stochastic discount factor. Growth optimal returns were analyzed in Cochrane (1992) and Bansal and Lehmann (1997). Campbell, Kazemi, and Nanisetty (1999) note the relationship between the growth optimal portfolio and the return on asymptotic discount bonds.

We illustrate Proposition 2.2 for a kernel whose logarithm follows an infinite moving-average process with normal innovations. We show that if this process is covariance stationary, then condition (2.3) is satisfied, that is, there are no permanent innovations. Assume that

$$M_t = \beta(t)\exp\left(\sum_{j=0}^{\infty}\alpha_j\varepsilon_{t-j}\right),$$

with $\varepsilon_t \sim N(0,\sigma^2)$, $\alpha_0 = 1$, and $\beta(\cdot)$ a function of time. Then

$$E_t \log\frac{E_{t+1}[M_{t+k}]}{E_t[M_{t+k}]} = -\frac{1}{2}(\alpha_{k-1})^2\sigma^2.$$

If $M_t/\beta(t)$ is covariance stationary, so that the variance is finite and independent of time, we have that $\lim_{k\to\infty}(\alpha_{k-1})^2 = 0$, and the condition of *no permanent innovations* is satisfied.[6] It also follows directly that

$$E[h_t(\infty)] = \frac{\sigma^2}{2}.$$

Recall that σ is the conditional volatility of the discount factor or, equivalently, the volatility of the innovations of the pricing kernel. This last equation illustrates that if a pricing kernel has no permanent innovations, then the volatility of the innovations of the pricing kernel is tightly linked to the term premium. Hansen and Jagannathan (1991) and Cochrane and Hansen (1992) show that the conditional volatility of the discount factor is quite large, so a pricing kernel without permanent innovations will have a very large term premium.

2.1. The size of the permanent component of the pricing kernel

So far, we have focused on kernels that have innovations that are either permanent or not. We now consider the case of a kernel that has both a permanent and a transitory component. If we assume that the covariance between the two components stabilizes as maturity increases, then we can obtain a lower bound on the size of the permanent component of the pricing kernel. If, as a matter of definition, we further require that the permanent component be a martingale, then we can obtain a lower bound for the volatility of the permanent innovations relative to the overall volatility of the stochastic discount factor.

Proposition 2.3. *Assume that the kernel has a component with transitory innovations M_t^T, that is one for which (2.3) hold, and a component that has permanent innovations M_t^P, that is one for which (2.3) does not holds, so that*

$$M_t = M_t^T \, M_t^P.$$

Let $v_{t,t+k}$ be defined as

$$v_{t,t+k} \equiv \frac{cov_t\left[M_{t+k}^T,\ M_{t+k}^P\right]}{E_t\left[M_{t+k}^T\right]\ E_t\left[M_{t+k}^P\right]},$$

[6]Note that with $\log x \sim N(\mu, \sigma^2)$ $\text{var}(x) = \exp(2\mu + \sigma^2)\left[\exp(\sigma^2) - 1\right]$. Thus, if $\log x$ is stationary, x is stationary too.

and assume that

$$\lim_{k\to\infty} E_t\left[\log\frac{(1+v_{t+1,t+k})}{(1+v_{t,t+k})}\right] = 0 \text{ almost surely.} \tag{2.7}$$

Then

$$-\lim_{k\to\infty}\left\{E_t\log\frac{E_{t+1}\left[M^P_{t+k}\right]}{E_t\left[M^P_{t+k}\right]}\right\} = \{\log E_t M_{t+1} - E_t\log M_{t+1}\} - h_t(\infty) \tag{2.8}$$

$$\geq E_t\log\frac{R_{t+1}}{R_{t+1,1}} - h_t(\infty). \tag{2.9}$$

Thus, we have first an equality, (2.8), relating the size of the permanent component to a Jensen's inequality effect associated with the next period's kernel (the term in braces on the right-hand side). And second, we have an inequality, (2.9), because this Jensen effect is shown to equal the growth optimal excess return; thus, $E_t \log\frac{R_{t+1}}{R_{t+1,1}}$ for all R_{t+1} is (weakly) smaller.

Proposition 2.3 holds when the transitory and permanent components are uncorrelated, but it requires only a much weaker condition. Condition (2.7) states that the conditional covariance between the transitory and permanent components stabilizes for long forecasting horizons. To understand condition (2.7) better, we present a simple example in which it is satisfied. In the example, both components of the pricing kernel are lognormal with correlated innovations: the permanent is a random walk with drift, and the transitory is covariance stationary. This type of process has often been used in the measurement of the size of the permanent component for linear time series. See, for instance, Watson (1986) and Cochrane (1988).

Example 2.4. *Assume that*

$$\begin{aligned}\log M^P_{t+1} &= \log\beta + \log M^P_t + \varepsilon^P_{t+1},\\ \log M^T_{t+1} &= \sum_{i=0}^{\infty}\alpha_i\varepsilon^T_{t+1-i},\end{aligned}$$

where α is a square summable sequence and ε^P_t and ε^T_t are i.i.d. normal with covariance σ_{TP}. Direct computation gives

$$\log\frac{(1+v_{t+1,t+k})}{(1+v_{t,t+k})} = -\alpha_{k-1}\sigma_{TP};$$

hence, (2.3) is satisfied, since $\lim_{k\to\infty} \alpha_{k-1} = 0$ *because* α *is square summable.*[7]

The following proposition gives a clear interpretation of the expression measuring the size of the permanent component derived in Proposition 2.3.

Proposition 2.5. *Assume that* $\pi_{t+1} \equiv \log \frac{M_{t+1}^P}{M_t^P}$ *and* $\pi_{t+1} = \bar{\pi} + \sum_{i=0}^{\infty} b_i \varepsilon_{t+1-i}$*, where* b *is a square summable sequence and* ε_t *is i.i.d. normal. Assume that* $|cov_t(\pi_{t+k}, \pi_{t+k-j})| < \rho^j B$ *with* $1 < \rho < 1$ *and* $B > 0$*. Then*

$$-\lim_{k\to\infty} E_t \left\{ \log \frac{E_{t+1}\left[M_{t+k}^P\right]}{E_t\left[M_{t+k}^P\right]} \right\} = \frac{1}{2} \left\{ var(\pi_s) + 2 \sum_{j=1}^{\infty} cov(\pi_s, \pi_{s-j}) \right\}. \tag{2.10}$$

Notice that the right-hand side of equation (2.10) is the spectral density of the discount rate at frequency zero. This seems a very natural representation for the size of the permanent component.

Now we make an assumption about the process for the permanent component. In the spirit of Beveridge and Nelson (1981) and Cochrane (1988), we assume that the permanent component is a martingale. We can then bound the volatility of the permanent component of the discount factor relative to the total volatility. For this purpose, we use $J(x) \equiv \log(E[x]) - E[\log(x)]$ as a measure of volatility, defined for a positive random variable x.

Proposition 2.6. *Assume that the permanent component is a martingale, so that*

$$M_{t+t}^P = M_t^P \cdot \lambda_{t+1}^P, \text{ with } E_t \lambda_{t+1}^P = 1,$$

and under the assumptions from Proposition 2.3, then (i)

$$J_t\left(\lambda_{t+1}^P\right) \geq E_t \log \frac{R_{t+1}}{R_{t+1,1}} - h_t(\infty) \tag{2.11}$$

and (ii)

$$\frac{J\left(\lambda_{t+1}^P\right)}{J\left(\frac{M_{t+1}}{M_t}\right)} \geq \frac{E\left[\log \frac{R_{t+1}}{R_{t+1,1}}\right] - E\left[h_t(\infty)\right]}{E\left[\log \frac{R_{t+1}}{R_{t+1,1}}\right] + J\left(1/R_{t+1,1}\right)} \tag{2.12}$$

[7] Note that $(1 + v_{t,t+k}) = E_t\left(M_{t+k}^T M_{t+k}^P\right) / E_t\left(M_{t+k}^T\right) E_t\left(M_{t+k}^P\right)$. Thus, the deterministic part, represented by β, is irrelevant.

if $-E[h_t(\infty)] \leq J(1/R_{t+1,1})$*, or if* $-E[h_t(\infty)] > J(1/R_{t+1,1})$*, then* $J\left(\lambda_{t+1}^P\right) / J\left(\frac{M_{t+1}}{M_t}\right) > 1$*, where* $J_t(x_{t+1}) \equiv \log E_t x_{t+1} - E_t \log x_{t+1}$ *and* $J(x_{t+1}) \equiv \log E x_{t+1} - E \log x_{t+1}$ *are the conditional and unconditional Jensen's inequality terms associated with the log of the random variable* x*.*

We will focus on providing a lower bound for $J\left(\lambda_{t+1}^P\right) / J\left(\frac{M_{t+1}}{M_t}\right)$ as presented in equation (2.12). That is, if $-E[h_t(\infty)] \leq J(1/R_{t+1,1})$ were to hold, the size of the permanent component would be bounded below by 1; in that case, the right-hand side of equation (2.12) would be an upper bound for the size of the permanent component. Also, note that for the bound in equation (2.11), $J\left(\lambda_{t+1}^P\right) = EJ_t\left(\lambda_{t+1}^P\right)$ due to the assumption that the permanent component is a martingale. Because of this, taking unconditional expectations on both sides of (2.11), preserves the inequality.[8]

To better understand Proposition 2.6, we expand on our measure of volatility $J(x)$. Clearly, if $var(x) = 0$, then $J(x) = 0$. The reverse is not true, as higher-order moments than the variance also affect the size of this Jensen measure as further illustrated below. More specifically, the variance and $J(x)$ are related in the following way. Consider the general measure of volatility $f(Ex) - Ef(x)$, with $f(\cdot)$ a concave function. The statistic $J(x)$ is obtained by making $f(x) = \log x$, while for the variance, $f(x) = -x^2$. In a similar sense, if a random variable x_1 is more risky than x_2 in the Rothschild-Stiglitz way, then $J(x_1) \geq J(x_2)$ and, of course, $var(x_1) \geq var(x_2)$.[9]

We can also directly relate the moments of $\log x$ to $J(x)$. If $\log(x)$ has moments of all orders,

$$J(x) = \sum_{j=2}^{\infty} \kappa_j / j!,$$

where κ_j is the jth cumulant for the conditional distribution of $\log x$. Cumulants are related to the moments of the distribution of $\log x$. For instance, $\kappa_1 = \mu_1$, $\kappa_2 = \mu_2$, $\kappa_3 = \mu_3$, and $\kappa_4 = \mu_4 - 3(\mu_2)^2$, where μ_1 denotes the mean of $\log x$

[8] Note that the presence of a deterministic component in M_{t+1}^P would not affect any results; for instance, assuming that $E_t \lambda_{t+1}^P / \beta = 1$ would lead to the same equation (2.11) and equation (2.12).

[9] Recall that x_1 is more risky than x_2 in the sense of Rothschild and Stiglitz if, for $E(x_1) = E(x_2)$, $E(f(x_1)) \leq E(f(x_2))$ for *any* concave function f.

and μ_j for $j > 1$ the jth central moment.[10] Thus, the Jensen's effect summarizes information about higher moments.

To further illustrate the measure J, note that if x is lognormal, then $J(x) = 1/2\ var(\log x)$. The next two examples and the subsequent discussion illustrate the bounds derived in Proposition 2.6.

Example 2.7. *Assume that* $\log \lambda_{t+1}^P$ *is normal i.i.d. with variance* σ_P^2 *and that* $\log \frac{M_{t+1}}{M_t}$ *is normal with unconditional variance* $\sigma_{\Delta \log M}^2$*. Then*

$$J_t\left(\lambda_{t+1}^P\right) = \frac{1}{2}\sigma_P^2 \geq E_t \log \frac{R_{t+1}}{R_{t+1,1}} - h_t(\infty)$$

and

$$\frac{J\left(\lambda_{t+1}^P\right)}{J\left(\frac{M_{t+1}}{M_t}\right)} = \frac{\sigma_p^2}{\sigma_{\Delta \log M}^2} \geq \frac{E\left[\log \frac{R_{t+1}}{R_{t+1,1}}\right] - E\left[h_t(\infty)\right]}{E\left[\log \frac{R_{t+1}}{R_{t+1,1}}\right] + \frac{1}{2}\sigma_{\log R_{t+1,1}}^2}.$$

Thus, the ratio of the Jensen's effects is just the ratio of the innovation variance of the permanent component to the unconditional variance of the stochastic discount factor. On the right-hand side of the inequality, we have used the fact that, given the lognormality of the stochastic discount factor, the interest rate is lognormal itself. Beveridge and Nelson (1981) show that it is always possible to decompose a linear homoscedastic difference stationary process into a random walk component and a component that is covariance stationary. The example here falls into this category. Cochrane (1988) focuses on the ratio of the innovation variance of the random walk component to the variance of the growth rate of the time series as a measure of the permanent component in GDP.[11]

The next example covers a large class of processes typically used in finance.[12]

Example 2.8. *Assume that* λ_{t+1}^P *and* $\frac{M_{t+1}}{M_t}$ *are conditionally lognormal. Then*

$$J_t\left(\lambda_{t+1}^P\right) = \frac{1}{2}var_t\left(\log \lambda_{t+1}^P\right) \geq E_t\left[\log \frac{R_{t+1}}{R_{t+1,1}}\right] - h_t(\infty) \tag{2.13}$$

[10] Note that for the normal distribution, cumulants are zero after the first two, as can be seen in our two examples. See Billingsley (1995) for details and Backus, Foresi and Telmer (1998) for an application to the forward risk premium.

[11] See Quah (1992) about specifying the permanent component as a random walk.

[12] The affine processes used in Backus, Foresi, and Telmer (1998) fall into this category.

and

$$\frac{J\left(\lambda^P_{t+1}\right)}{E\left[J_t\left(\frac{M_{t+1}}{M_t}\right)\right]} = \frac{E\left[var_t\left(\log \lambda^P_{t+1}\right)\right]}{E\left[var_t\left(\log \frac{M_{t+1}}{M_t}\right)\right]} \geq \frac{E\left[\log \frac{R_{t+1}}{R_{t+1,1}}\right] - E\left[h_t\left(\infty\right)\right]}{E\left[\log \frac{R_{t+1}}{R_{t+1,1}}\right]}. \qquad (2.14)$$

The bound for the absolute size of the permanent component, equation (2.13), has again a straightforward interpretation as limiting the conditional variance of the permanent component. We have defined a slightly different bound for the relative size of the permanent component. Indeed, (2.14) is now for the mean of the conditional variance as opposed to the unconditional variance in equation (2.12). The right-hand side also does not require, in this case, a term related to interest rate volatility. Note that with conditional lognormality, $Evar_t\left(\log \frac{M_{t+1}}{M_t}\right) = var\left(\log \frac{M_{t+1}}{M_t}\right) - var\left(E_t \log \frac{M_{t+1}}{M_t}\right)$; for the more general case, $EJ_t\left(\frac{M_{t+1}}{M_t}\right) = J\left(\frac{M_{t+1}}{M_t}\right) - J\left(E_t \frac{M_{t+1}}{M_t}\right)$. While interest rate data allows us to estimate $J\left(E_t \frac{M_{t+1}}{M_t}\right)$ directly, and for the lognormal case, $\sigma^2_{\log R_{t+1,1}}$, with conditional lognormality, interest rates cannot pin down $var\left(E_t \log \frac{M_{t+1}}{M_t}\right)$. Thus, we have the modification in the definition of the bound.[13]

2.1.1. Yields and forward rates: Alternative measures of term spreads

For empirical implementation, we want to be able to extract as much information from long-term bond data as possible. For that reason, we show here that for asymptotic zero-coupon bonds, the unconditional expectations of the yields and the forward rates are equal to the unconditional expectations of the holding returns.

Consider forward rates. The k-period forward rate differential is defined as the rate for a one-period deposit k periods from now relative to a one-period deposit now:

$$f_t\left(k\right) \equiv -\log\left(\frac{V_{t,k+1}}{V_{t,k}}\right) - \log \frac{1}{V_{t,1}}.$$

Forward rates and expected holding returns are also closely related. They both compare prices of bonds with a one-period maturity difference, the forward rate does it for a given t, while the holding return considers two periods in a row.

[13]With conditional lognormality, $var\left(E_t \log \frac{M_{t+1}}{M_t}\right) = var\left(\log E_t \frac{M_{t+1}}{M_t} - \frac{1}{2} var_t\left(\log \frac{M_{t+1}}{M_t}\right)\right)$. Because $var_t\left(\log \frac{M_{t+1}}{M_t}\right)$ is not assumed to be constant, interest rates, $E_t \frac{M_{t+1}}{M_t}$, are not directly informative for $E_t \log \frac{M_{t+1}}{M_t}$.

Proposition 2.9. *Assume that bond prices have means that are independent of calendar time, so that $E(V_{t,k}) = E(V_{\tau,k})$ for every t and k. Then*

$$E[h_t(k)] = E[f_t(k-1)].$$

Note that there is a time shift because the holding return on a k-period bond is for the purchase of a k-period bond that becomes a $(k-1)$-period bond when it is sold. For the k-period forward, the corresponding rate compares dates between k and $k+1$ periods from now.

We define the continuously compounded *yield differential* between a k-period discount bond and a one-period risk-free bond as

$$y_t(k) \equiv \log\left(\frac{V_t[1_{t+1}]}{(V_t[1_{t+k}])^{1/k}}\right)$$

and the limiting yield differential as

$$y_t(\infty) \equiv \lim_{k\to\infty} \log\left(\frac{V_t[1_{t+1}]}{(V_t[1_{t+k}])^{1/k}}\right).$$

Proposition 2.10. *Average forward rate differentials equal yield differentials*

$$y_t(k) = \frac{1}{k}\sum_{j=0}^{k-1} f_t(j),$$

and the limiting rates are equal; that is,

$$\lim_{k\to\infty} y_t(k) = \lim_{k\to\infty} f_t(k)$$

if the two limits exist.

The proof of this proposition is trivial, because the forward rate is defined as the difference in price between a $(j+1)$- and a j-period zero-coupon, while the yield is just the per period discount of the price of the k-period bond relative to the current one-period bond.

The next proposition shows that under regularity conditions, the three measures of the term spreads are equal for the limiting zero-coupon bonds.

Proposition 2.11. *If the limits $h_t(\infty)$, $f_t(\infty)$, and $y_t(\infty)$ exist, the unconditional expectations of holding returns are independent of calendar time; that is,*

$$E[\log R_{t+1,k}] = E[\log R_{\tau+1,k}] \quad \textit{for all } t, \tau, k$$

and holding returns and yields are dominated by an integrable function,

$$E[h_t(\infty)] = E[f_t(\infty)] = E[y_t(\infty)].$$

In practice, these three measures may not be equally easy to estimate for two reasons. One is that the term premium is defined in terms of the conditional expectation of the holding returns. But this will have to be estimated from ex post realized holding returns, which are very volatile. Forward rates and yields are, according to the theory, conditional expectations of bond prices. While forward rates and yields are more serially correlated than realized holding returns, they are substantially less volatile. Overall, they should be more precisely estimated. The other reason is that, while results are derived for the limiting maturity, data is available only for finite maturities. All the previous results could have been derived for a finite k by assuming that limiting properties are reached at maturity k, except Propositions 2.10 and 2.11. In these cases, yields are equal to averages of forward rates (or holding returns), and the average only equals the last element in the limit. For this reason, yield differentials, y, might be slightly less informative for k finite than the term spreads estimated from forward rates and holding returns.

2.2. Consumption

In many models used in the literature, the pricing kernel is a function of current or lagged consumption. Thus, the stochastic process for consumption is a determinant of the process for the pricing kernel. In this section, we present sufficient conditions on consumption and the function mapping consumption into the pricing kernel so that pricing kernels have no permanent innovations. We are able to define a large class of stochastic processes for consumption that, combined with standard preference specifications, will result in counterfactual asset pricing implications. We also present two examples of utility functions in which the resulting pricing kernels have permanent innovations because of the persistence introduced through the utility functions.[14]

[14]In Section 3.3, we present evidence that the permanent components of asset pricing kernels are mainly *real*, as opposed to *nominal* (meaning driven by uncertainty in the aggregate price level). For this reason, we omit nominal risk in this section.

As a starting point, we present sufficient conditions for kernels that follow Markov processes to have no permanent innovations. We then consider consumption within this class of processes. Assume that

$$M_t = \beta(t) f(s_t),$$

where f is a positive function and that $s_t \in S$ is Markov with transition function Q:

$$\Pr(s_{t+1} \in A | s_t = s) = Q(s, A).$$

We assume that Q has an invariant distribution λ^* and that the process $\{s_t\}$ is drawn at time $t = 0$ from λ^*. In this case, s_t is strictly stationary, and the unconditional expectations are taken with respect to λ^*. We use the standard notation,

$$\left(T^k f\right)(s) \equiv \int_S f(s') Q^k(s, ds'),$$

where Q^k is the k-step ahead transition constructed from Q.

Proposition 2.12. *Assume that there is a unique invariant measure, λ^*, and that*

$$\frac{\left(T^{k-1} f\right)(s')}{\left(T^k f\right)(s)} \geq l > 0 \text{ for all } k.$$

In addition, if either (i)

$$\lim_{k\to\infty} \left(T^k f\right)(s) = \int f d\lambda^* > 0$$

or, in case $\lim_{k\to\infty} \left(T^k f\right)(s)$ is not finite, if (ii)

$$\lim_{k\to\infty} \left[\left(T^{k-1} f\right)(s') - \left(T^k f\right)(s)\right] \leq A(s)$$

for each s and s', then

$$\lim_{k\to\infty} E_t \log \frac{E_{t+1}[M_{t+k}]}{E_t[M_{t+k}]} = 0.$$

Remark 1. *For a set of conditions leading to the existence of a unique invariant measure, see Stokey and Lucas (1989, Section 11).*

Remark 2. *The uniform lower bound, l, is slightly stronger than the strict positivity implied by no-arbitrage. This bound is needed to pass the limit through the conditional expectation operator.*

As an example, consider the pricing kernel as a log-linear autoregressive process

$$\log M_{t+1} = \log \beta(t) + \rho \log M_t + \varepsilon_{t+1},$$

with $\varepsilon \sim N(0, \sigma^2)$. Simple calculations show that

$$E_t \log \frac{E_{t+1}[M_{t+k}]}{E_t[M_{t+k}]} = -\frac{1}{2}\rho^{2(k-1)}\sigma^2.$$

If $|\rho| < 1$, $\lim_{k\to\infty} E_t \log \frac{E_{t+1}[M_{t+k}]}{E_t[M_{t+k}]} = 0$, while if $\rho = 1$, $E_t \log \frac{E_{t+1}[M_{t+k}]}{E_t[M_{t+k}]} = -\frac{1}{2}\sigma^2$ for any k.[15] Thus, $\log M$ being covariance stationary with a deterministic trend is a sufficient condition for not having permanent innovations. For the nonstationary process with the unit root, innovations are permanent.[16]

We are now ready to consider consumption explicitly. Assume that

$$C_t = \tau(t)\, c_t = \tau(t)\, g(s_t),$$

where g is a positive function, $s_t \in S$ is Markov with transition function Q, and $\tau(t)$ represents a deterministic trend. We assume (a) that a unique invariant measure λ^* exists. Furthermore, using the standard notation

$$T^{*k}\lambda(A) = \int Q^k(s, A)\,\lambda(ds),$$

we assume (b) that

$$\lim_{k\to\infty} T^{*k}\lambda(A) = \lambda^*(A), \text{ for all } A.$$

Proposition 2.13. *Assume that $M_t = \beta(t) f(c_t, x_t)$, with $f(\cdot)$ positive, bounded, and continuous and that $(c_t, x_t) \equiv s_t$ satisfies properties (a) and (b) with $f(\cdot) > 0$ with positive probability. Then M_t has no permanent innovations.*

[15] Note that this process does not satisfy the lower bound used as a regularity condition in Proposition 2.12.

[16] It is straightforward to show that for a multivariate log-linear process, $\lim_{k\to\infty} E_t \log \frac{E_{t+1}[M_{t+k}]}{E_t[M_{t+k}]} = 0$, unless there is at least one unit root.

An example covered by this proposition is CRRA utility, where $f(c_t) = c_t^{-\gamma}$, with $\bar{c} \geq c_t \geq \varepsilon > 0$. Note that the bounds might not be necessary. For instance, if $\log c_t = \rho \log c_{t-1} + \varepsilon_t$, with $\varepsilon \sim N(0, \sigma^2)$ and $|\rho| < 1$, then, $\log f(c_t) = -\gamma \log c_t$, and direct calculations show that condition (2.3) defining the property of no permanent innovations is satisfied. A counter example, in which condition (2.3) is violated and that thus has a better chance of explaining term structure and return behavior would be a consumption process that does not have an invariant distribution, for instance, if $\rho = 1$.

For the CRRA case, even with consumption satisfying properties (a) and (b), condition (2.3) could fail to be satisfied because $c_t^{-\gamma}$ is unbounded if c_t gets arbitrarily close to zero with large enough probability. While this case cannot be ruled out a priori, this property would not seem to be a desirable feature, because infinitely large marginal utility is more a technical condition than a representation of consumer behavior. We are not aware of economic applications where equation (2.3) is violated while consumption is assumed to satisfy properties (a) and (b).

2.2.1. Examples with additional state variables

There are many examples in the literature in which marginal utility is a function of additional state variables; Proposition 2.13 also applies in these cases. Prominent examples are models in which the representative agent utility displays habit or durability. We present three cases below. There are also examples of preference specifications in which, even with consumption satisfying the conditions required for Proposition 2.13, the additional state variables do not have invariant distributions, and, thus, innovations to pricing kernels have permanent effects. We present two examples of this type.

For the following examples we assume that aggregate consumption satisfies the assumptions of Proposition 2.13. In addition, we assume that $c_t \in [\underline{c}, \bar{c}]$ and that the trend is geometric, $\tau(t) = \tau^t$. For the following three cases, $s_t = (c_t, x_t)$, $f(\cdot)$, and $\beta(t) = (\beta \tau^{1-\gamma})^t$ will then satisfy the assumptions of Proposition 2.13.

First, Ferson and Constantinides (1991) study a model in which the utility of the representative agent is given by

$$E_0 \left[\sum_{t=0}^{\infty} \beta^t \frac{1}{1-\gamma} \left(C_t - \sum_{j=1}^{J} \alpha_j C_{t-j} \right)^{1-\gamma} \right],$$

where $\alpha' s$ are constant coefficients. We can map the pricing kernel of this model into the Markov case as

$$\begin{aligned} f\left(c_t, x_t\right) &= \left(c_t - \sum_{j=1}^{J} \alpha_j \tau^{-j} c_{t,j}\right)^{-\gamma} \\ &\quad - \sum_{j=1}^{J} \left(\beta\tau\right)^j \alpha_j E_t \left[\left(c_{t+j} - \sum_{r=1}^{J} \alpha_r \tau^{-r} c_{t+r,j}\right)^{-\gamma}\right] \end{aligned}$$

for

$$x_t \equiv \left(c_{t-1}, c_{t-2}, ..., c_{t-J}\right).$$

Second, a related habit model is the one in which each agent compares her or his consumption with the aggregate consumption. Abel (1999) studies the case where the utility is given by

$$E_0 \left[\sum_{t=0}^{\infty} \beta^t \frac{1}{1-\gamma} \left(\frac{C_t}{\bar{C}_t^{\gamma_0} \bar{C}_{t-1}^{\gamma_1}}\right)^{1-\gamma}\right],$$

where $\bar{C}_t$ denotes average aggregate consumption and γ, γ_0, $\gamma_1 > 0$. In equilibrium, $\bar{C}_t = C_t$. We can then map the pricing kernel of this model into the Markov case as

$$f\left(c_t, x_t\right) = c_t^{-\gamma}\, c_t^{-\gamma_0(1-\gamma)} \left(\tau^{-1} c_{t-1}\right)^{-\gamma_1(1-\gamma)}$$

for $x_t \equiv c_{t-1}$.

Third, Campbell and Cochrane (1999) study a version of external habit where the utility of each identical agent is given by

$$E_0 \left[\sum_{t=0}^{\infty} \beta^t \frac{1}{1-\gamma} \left(C_t - X_t\right)^{1-\gamma}\right],$$

where X_t is taken as given by each agent and evolves as

$$\begin{aligned} X_{t+1} &= \bar{C}_{t+1} \left(1 - y_{t+1}\right) \\ \log y_{t+1} &= \left(1-\phi\right) \hat{y} + \phi \log y_t + h\left(y_t\right) \left[\log \bar{C}_{t+1} - E_t \log \bar{C}_{t+1}\right] \end{aligned}$$

for some continuous and decreasing function h, constants $0 < \phi < 1$, and $\hat{y} < 0$. Using that in equilibrium $\bar{C}_t = C_t$, we can map the pricing kernel of this model into the Markov case as

$$f\left(c_t, x_t\right) = \left(c_t \left(1 - y_t\right)\right)^{-\gamma}$$

for $x_t = y_t$.

We now present two examples of representative agent economies in which preferences are nonseparable in such a way that the pricing kernel does have permanent innovations even if detrended aggregate consumption is iid. In the first example, preferences are given by expected discounted utility displaying an extreme form of habit formation. In the second example, preferences are given by nonexpected utility.

For the first example, the representative agent's preferences are given by

$$E_0 \left[\sum_{t=0}^{\infty} \beta^t \frac{1}{1-\gamma} \left(C_t / X_t^{\phi} \right)^{1-\gamma} \right]$$

with $\gamma > 0$, and X_t is the external habit stock. Thus, the pricing kernel is

$$M_t = \beta^t C_t^{-\gamma} X_t^{\phi(\gamma-1)}.$$

We assume that detrended consumption is iid, so that $C_t = \tau^t c_t$ with c_t iid. The stock of habit evolves as

$$\log X_{t+1} = \log X_t + \log \bar{C}_t. \tag{2.15}$$

In equilibrium, $\bar{C}_t = C_t$, so that

$$M_{t+k} = \beta^k C_t^{-\gamma+\phi(\gamma-1)} \left[\Pi_{s=1}^{k} C_{t+s} \times X_t \right]^{\phi(\gamma-1)}.$$

Thus, after some algebra, we can write

$$\frac{E_{t+1} M_{t+k}}{E_t M_{t+k}} = \frac{C_{t+1}^{\phi(\gamma-1)}}{E_t \left(C_{t+1}^{\phi(\gamma-1)} \right)}.$$

Hence, if $\phi > 0$ and $\gamma \neq 1$, then the pricing kernel has permanent innovations. This result is not surprising, given the extreme amount of persistence assumed in the law of motion of the habit stock in equation (2.15).

For the second example, the representative agent has preferences represented by nonexpected utility. In particular, this class of preferences can be represented in a recursive way as

$$U_t = \phi \left(c_t, E_t U_{t+1} \right),$$

where U_t is the utility starting at time t and ϕ is an increasing concave function. For this utility function, risk aversion does not need to equal the reciprocal of the intertemporal elasticity of substitution. Epstein and Zin (1989) and Weil (1990) develop a parametric case in which the risk aversion coefficient, γ, and the reciprocal of the elasticity of intertemporal substitution, ρ, are constant. They also characterize the stochastic discount factor M_{t+1}/M_t for a representative agent economy with an arbitrary consumption process $\{C_t\}$ as

$$\frac{M_{t+1}}{M_t} = \left[\beta \left(\frac{C_{t+1}}{C_t}\right)^{-\rho}\right]^{\theta} \left[\frac{1}{R^c_{t+1}}\right]^{(1-\theta)} \tag{2.16}$$

with

$$\theta = \frac{1-\gamma}{1-\rho},$$

where β is the time discount factor and $R^c_{t+1}$the gross return on the consumption equity, that is the gross return on an asset that pays a stream of dividends equal to consumption $\{C_t\}$.

Inspection of (2.16) reveals that a pricing kernel M_{t+1} for this model is

$$M_{t+1} = \beta^{\theta(t+1)}\ Y_{t+1}^{\theta-1}\ C_{t+1}^{-\rho\theta}, \tag{2.17}$$

where

$$Y_{t+1} = R^c_{t+1} \cdot Y_t$$

and $Y_0 = 1$.

The next proposition shows that the nonseparabilities that characterize these preferences for $\theta \neq 1$ are such that, even if consumption is iid, the pricing kernel has permanent innovations. More precisely, assume that consumption satisfies

$$C_t = \tau^t c_t, \tag{2.18}$$

where $c_t \in [\underline{c}, \bar{c}]$ is iid with cdf F. Let V^c_t be the price of the consumption equity, so that

$$R^c_{t+1} = \frac{V^c_{t+1} + C_{t+1}}{V^c_t}.$$

We assume that agents discount the future enough so as to have a well-defined price-dividend ratio. Specifically, we assume that

$$\max_{c \in [\underline{c}, \bar{c}]} \beta\tau^{1-\rho} \left\{\int \left(\frac{c'}{c}\right)^{1-\gamma} dF(c')\right\}^{1/\theta} < 1. \tag{2.19}$$

Proposition 2.14. *Let the pricing kernel be given by (2.17), let the detrended consumption be iid as in (2.18), and assume that (2.19) holds. Then the price-dividend ratio for the consumption equity is given by*

$$\frac{V_t^c}{C_t} = \psi c_t^{\gamma-1}$$

for some constant $\psi > 0$; hence, V_t^c/C_t is iid. Moreover,

$$x_{t+1,k} \equiv \frac{E_{t+1}M_{t+k}}{E_t M_{t+k}} = \frac{\left(1+\frac{1}{\psi}c_{t+1}^{(1-\gamma)}\right)^{\theta-1}}{E_t\left\{\left(1+\frac{1}{\psi}c_{t+1}^{(1-\gamma)}\right)^{\theta-1}\right\}}; \tag{2.20}$$

thus the pricing kernel has permanent innovations, that is $E_t \log x_{t+1,k} < 0$, iff $\theta \neq 1$, $\gamma \neq 1$, and c_t has strictly positive variance.

Note that $\theta = 1$ corresponds to the case in which preferences are given by time separable expected discounted utility; and hence, with iid consumption, the pricing kernel has only temporary innovations. Expression (2.20) also makes clear that for values of θ close to one, the size of the permanent component is very small.

3. Empirical Evidence

In this section, we present our estimates for the size of the permanent component of pricing kernels. We use several data sets, notably U.S. zero-coupon bonds and coupon bonds, and U.K. coupon bonds. Additional results are presented. First, to illustrate our findings, we present two simple examples of processes for pricing kernels. Second, we show that the permanent component from inflation is small, suggesting that most of the permanent effects in pricing kernels are real. Third, we measure the size of the permanent component of consumption directly from consumption data.

3.1. The size of the permanent component

We estimate here the lower bound of the size of the permanent component of pricing kernels that was derived in Proposition 2.6:

$$\frac{J\left(\lambda_{t+1}^P\right)}{J\left(\frac{M_{t+1}}{M_t}\right)} \geq \frac{E\left[\log\frac{R_{t+1}}{R_{t+1,1}}\right] - E\left[h_t(\infty)\right]}{E\left[\log\frac{R_{t+1}}{R_{t+1,1}}\right] + J\left(1/R_{t+1,1}\right)}. \tag{3.1}$$

Tables 1, 2, and 3 contain the estimates of the right-hand side of (3.1) obtained by replacing each expected value with its sample analog for different data sets.

In Table 1, we present estimates using zero-coupon bonds for various maturities, k, of 25 and 29 years, and for various term spread measures. We find that the size of the permanent component is usually about 100%; none of our estimates are below 75%. For each maturity k, we present four panels, A, B, C, and D, where we use forward rates, holding returns, and yields to estimate $E[h_t(\infty)]$, since, as we have shown above, under regularity conditions, $E[f_t(k)]$ and $E[y_t(k)]$ converge to $E[h_t(\infty)]$ for large k. The data set is monthly, covering the period 1946:12 to 1999:12. In panels A, B, and C the holding period for the aggregate equity portfolio is one year, so the returns used in the estimation overlap. In panel A, forward rates are computed for a yearly period, that is, by combining the prices of zero-coupon bonds with a difference in maturity of one year. In panel B, the holding period returns on bonds are calculated using a yearly holding period. In panel D, the holding period is one month, so the returns do not overlap. Standard errors of the estimated quantities are presented in parentheses; for the size of the permanent component, we use the *delta* method. The variance-covariance of the estimates is computed by using a Newey and West (1987) window with 36 lags to account for the overlap in returns and the persistence of the different measures of the spreads.[17] When yields and forward rates are used to measure the term spreads, our estimates of the size of the permanent component are all close to 100%, with standard errors of 10% and lower. One factor that affects our estimates is the choice of the risk-free rate. When we use a holding period of one year, as in panels A, B, and C, we use an annual rate (the yield on a zero-coupon bond maturing in one year) as the risk-free return. For comparison, panel D presents results with monthly rates. Since monthly rates are about 1% below the annual rates, all excess returns increase by approximately that same amount, leading to a slight reduction in the estimate of the size of the permanent component.[18] Note

[17]For maturities longer than 15 years, we do not have a complete data set for zero-coupon bonds. In particular, long-term bonds have not been consistently issued during this period. For instance, for zero- coupon bonds maturing in 29 years, we have data for slightly more than half of the sample period, with data missing at the beginning and in the middle of our sample. The estimates of the various expected values on the right-hand side of (3.1) are based on various numbers of observations. We take this into account when computing the variance-covariance of our estimators. Our procedure gives consistent estimates as long as the periods with missing bond data are not systematically related to the magnitudes of the returns.

[18]Note that our data set does not contain the information necessary to present results for

that by estimating the right-hand side of equation (3.1) as the ratio of sample means, our estimate is consistent but biased in small samples because the denominator has nonzero variance. In Appendix C, we present estimates of this bias. They are quantitatively negligible, on the order of about 1% in absolute value terms. Finally, column 6 of Table 1 contains the asymptotic probability that the term spread is larger than the log equity premium. This would be consistent with a pricing kernel with no permanent innovations. The probability is very small, in most cases well below 1%.

In Table 2, we attempt to take into account that equation (3.1) holds with equality if R_{t+1} is the *growth optimal* return. In particular, we select portfolios to maximize $E\left[\log \frac{R_{t+1}}{R_{t+1,1}}\right]$. All the results in this table are for maturity k equal to 25 years. As a benchmark case, panel A reproduces the results of Table 1 using an aggregate equity portfolio to measure R_{t+1}. Panels B and C use different equity portfolios to measure R_{t+1}. In panel B, we present results for the return R_{t+1} on a portfolio that combines aggregate equity with the risk-free asset. Depending on the choice of the risk-free rate, $E\left[\log \frac{R_{t+1}}{R_{t+1,1}}\right]$ is up to 9% larger than the unleveraged log equity premium presented in panel A. Here, the investor is allowed to choose the amount of the aggregate market in her or his portfolio to maximize the log excess return. The investor typically chooses an equity share that is larger than 1, either 2.14 or 3.47 depending on whether the holding period is yearly of monthly. As a first-order effect, this leverage increases the mean return, but given that the log is a concave function, the ensuing increased volatility contributes negatively to the log excess return. As an illustration, assume that the return on the market, R^M, is lognormal and that the investor is allowed to choose the fraction, w, of her or his portfolio, R^P, that is allocated to the market then

$$\begin{aligned} E\left(\log\left(\frac{R^P}{R^f}\right)\right) &= E\left(\log\left(1+w\left[\frac{R^M-R^f}{R^f}\right]\right)\right) \\ &= \log E\left(1+wr\right)-\frac{1}{2}var\left(\log\left[1+wr\right]\right) \\ &\approx wE\left(r\right)-\frac{w^2}{2}var\left(r\right), \end{aligned}$$

where we have used $\log(1+x)\approx x$ and $r\equiv\frac{R^M-R^f}{R^f}$; thus, the optimal share is given by $w^*\approx E(r)/var(r)$.[19] In panel C, we let the investor choose from the

monthly holding periods for forwards rates and holding returns.

[19]Note that we do not assume lognormality for the results presented in Table 2.

menu of the 10 CRSP size decile portfolios, which leads to a log excess return of up to 22.5%.

Table 3 extends the sample period to over 100 years and adds an additional country, the United Kingdom. For the United States, given data availability, we use coupon bonds with about 20 years of maturity. For the United Kingdom, we use consols. For the United States, we estimate the size of the permanent component between 78% and 93%, depending on the time period and whether we consider the term premium or the yield differential. Estimated values for the United Kingdom are similar to the U.S. numbers.

A natural concern is whether 25- or 29-year bonds allow for good approximations of the limiting term spread, $E[h_t(\infty)]$. From Figure 1, which plots term structures for three definitions of term spreads, we take that the long end of the term structure is not increasing. This suggests, if anything, that our estimates of the size of the permanent component presented in Tables 1 and 2 are on the low side. In this figure, the standard error bands are wider for longer maturities, which is due to two effects. One is that spreads on long-term bonds are more volatile, especially for holding excess returns. The other is that for longer maturities, as discussed before, our data set is shorter.

Note that for the bound in Equation (3.1) to be well defined, specifically $J(1/R_{t+1,1})$, we have assumed that interest rates are stationary.[20] While the assumption of stationary interest rates is standard (for instance, in Ait Sahalia (1996)), many studies report the inability to reject unit roots (for instance, Hall, Anderson, and Granger (1992)). To some extent, if interest rates were nonstationary, this would seem to further support the idea that the pricing kernel itself is nonstationary.

3.2. Examples of pricing kernels

To illustrate quantitative findings, we present here two lognormal examples. The first example is a process for the pricing kernel that is roughly consistent with the overall volatility and persistence as estimated in subsection 3.1. The second example illustrates the power that bond data have to distinguish between similar levels of persistence.

For our first example, assume the following process for the pricing kernel $M =$

[20]Equation (2.11), which defines a bound for the size of the permanent component in absolute terms, does not require this assumption.

$M^T M^P$, with

$$\log M_t^P = \log M_{t-1}^P + \varepsilon_t^P, \text{ and } \log M_t^T = \rho \log M_{t-1}^T + \varepsilon_t^T,$$

where $\varepsilon_t^i \sim N(0, \sigma_i^2)$. Assume that the innovations of the permanent and transitory components are uncorrelated, that is, $\sigma_{PT}^2 = 0$. Then interest rate behavior is determined exclusively by the transitory component M_t^T. The following simple calculations show that this process can be made roughly consistent with the evidence about interest rates, growth optimal returns, and the size of the permanent component.

Simple calculations show that the volatility of the short rate and its autocorrelation are given by

$$var(\log R_{t,1}) = \frac{1-\rho}{1+\rho}\sigma_T^2, \text{ and} \tag{3.2}$$

$$corr(\log R_{t,1}, \log R_{t-1,1}) = \rho. \tag{3.3}$$

For the postwar period, the standard deviation of 1-month rates is about 2.9%, and the first-order serial correlation is about 0.968, implying that $\rho = 0.968$ and $\sigma_T^2 = 0.05$.[21] Under the stated assumptions, the variance of the log of the stochastic discount factor can be written as

$$var\left(\log \frac{M_{t+1}}{M_t}\right) = \sigma_P^2 + \frac{2}{1+\rho}\sigma_T^2. \tag{3.4}$$

As shown in Proposition 2.3

$$J\left(\frac{M_{t+1}}{M_t}\right) = E \log \frac{R_{t+1}^{GO}}{R_{t+1,1}} + J(1/R_{t+1,1}),$$

where R_{t+1}^{GO} is the growth optimal return. Thus, with lognormality,

$$var\left(\log \frac{M_{t+1}}{M_t}\right) = 2 \cdot E \log \frac{R_{t+1}^{GO}}{R_{t+1,1}} + var(\log R_{t+1,1}).$$

[21] An alternative way to pin down the volatility of the innovation of the stationary component would be to use the term premium for a long-term discount bond. The implied value of σ_T^2 for a term premium of about 1% for a 25-year bond would be even smaller for the given ρ. We discuss this case below.

Based on our estimates in Table 2, the growth optimal excess return should be at least 20%, so that $var\left(\log\frac{M_{t+1}}{M_t}\right)=0.4$; thus,

$$\sigma_P^2/var\left(\log\frac{M_{t+1}}{M_t}\right)=\frac{0.4-2/(1.968)\cdot 0.05}{0.4}=0.873.^{22}$$

Note that our measure of the size of the permanent component, as presented in this paper and restated in equation (2.12), is, in general, not constrained to be bounded by 1. This is the case, however, in the example here with uncorrelated innovations. To summarize, we have a simple example of a process that replicates the empirical magnitudes of the volatility of the stochastic discount factor and the volatility and first-order serial correlation of the short rate, as shown in equations (3.2), (3.3), and (3.4). This process also implies a permanent component of a magnitude similar to those estimated in the subsection 3.1.

The second parametric example illustrates that even for bonds with maturities between 10 and 30 years, one can obtain strong implications for the degree of persistence. Alternatively, the example shows that, in order to explain the low observed term premia for long-term bonds at finite maturities with a stationary pricing kernel, the largest root has to be extremely close to 1.

Assume that

$$\log M_{t+1}=\log\beta+\rho\log M_t+\varepsilon_{t+1}$$

with $\varepsilon_{t+1}\sim N(0,\sigma_\varepsilon^2)$. Simple algebra shows that

$$h\left(k\right)=\frac{\sigma_\varepsilon^2}{2}\left(1-\rho^{2(k-1)}\right). \tag{3.5}$$

This expression suggests that if the volatility of the innovation of the pricing kernel, σ_ε^2, is large, then values of ρ below 1 may have a significant quantitative effect on the term spread. In Table 4, we calculate the level of persistence, ρ, required to explain various levels of term spreads for discount bonds with maturities of 10, 20, and 30 years. Consistent with our estimates of a growth optimal return of at least 20%, as discussed in the previous example, we have that

$$0.4=var\left(\log\frac{M_{t+1}}{M_t}\right)=\frac{2}{1+\rho}\sigma_\varepsilon^2\simeq\sigma_\varepsilon^2,$$

[22] Note that we have omitted $var\left(\log R_{t+1,1}\right)$ because it is a mere 0.03ˆ2 = 0.0009.

where the last approximate equality holds for ρ close to 1. As is clear from Table 4, ρ has to be extremely close to 1. Otherwise, as shown in equation (3.5), ρ raised to the power of $2(k-1)$ with k equal to 10, 20, or 30 will vanish and $h(k)$ will be very large.

3.3. Nominal versus real pricing kernels

Because we have so far used bond data from nominal bonds, we have implicitly measured the size of the permanent component of nominal pricing kernels, that is, the processes that price future dollar amounts. We present now two sets of evidence showing that the permanent component is to a large extent real, so that we have a direct link between the size of the permanent component of pricing kernels and real economic fundamentals.

First, assume, for the sake of this argument, that all of the permanent movements in the (nominal) pricing kernel come from the aggregate price level. Specifically, assume that $M_t = \left(\frac{1}{P_t}\right) M_t^T$, where P_t is the aggregate price level. Because, P_t is directly observable, we can measure the size of its permanent component directly and then compare it to the estimated size of the permanent component of pricing kernels reported in Tables 1, 2, and 3. It turns out that the size of the permanent component in P_t is estimated at up to 100 times smaller than the size of the permanent component in pricing kernels. This suggests that movements in the aggregate price level have a minor importance in the permanent component of pricing kernels, and thus, permanent components in pricing kernels are primarily real.

The next proposition shows how to estimate the size of the permanent component based on the $J(.)$ measure.

Proposition 3.1. *Assume that the process X_t can be decomposed into a permanent component $X_t^P > 0$ and a transitory component $X_t^T > 0$, so that (i)*

$$X_t = X_t^P X_t^T$$

(ii) the permanent component is a martingale, that is,

$$E_t\left[X_{t+1}^P\right] = X_t^P \text{ for all } t,$$

(iii) the process X_t^T has no permanent innovations, that is,

$$\lim_{k\to\infty} E_t\left[\log \frac{E_{t+1}X_{t+k}^T}{E_t X_t^T}\right] = 0.$$

Additionally, assume the following regularity conditions: (a) the covariance between X_t^T and X_t^P stabilizes, that is,

$$\lim_{k\to\infty} E_t\left[\log\frac{(1+v_{t+1,t+k})}{(1+v_{t,t+k})}\right] = 0 \textit{ almost surely}$$

with $v_{t,t+k}$ defined as

$$v_{t,t+k} \equiv \frac{cov_t\left[M_{t+k}^T,\ M_{t+k}^P\right]}{E_t\left[M_{t+k}^T\right]\ E_t\left[M_{t+k}^P\right]},$$

(b) $\frac{X_{t+1}}{X_t}$ is strictly stationary, (c) that the following limit exists:

$$\lim_{k\to\infty} E\left[\log E_t\left[\frac{X_{t+k}}{X_t}\right] - \log E_t\left[\frac{X_{t+k-1}}{X_t}\right]\right],$$

and (d)

$$\lim_k \frac{1}{k} J\left(\frac{E_t X_{t+k}}{X_t}\right) = 0.$$

Then

$$J\left(\frac{X_{t+1}^P}{X_t^P}\right) = \lim_{k\to\infty}\frac{1}{k}J\left(\frac{X_{t+k}}{X_t}\right). \tag{3.6}$$

The usefulness of this proposition is that $J\left(X_{t+1}^P/X_t^P\right)$ is a natural measure for the size of the permanent component. However, it cannot directly be estimated if only X_t is observable, but X^P and X^T are not observable separately. Proposition 3.1 shows that $\lim_{k\to\infty}\frac{1}{k}J\left(X_{t+k}/X_t\right)$ is an equivalent measure, under some regularity conditions, and clearly, it can be estimated with knowledge of only X_t. This result is analogous to Cochrane (1988), with the difference that he uses the variance as a measure of volatility.

Cochrane (1988) proposes a simple method for correcting for small sample bias and for computing standard errors when using the variance as a measure of volatility. Thus, we will focus our presentation of the results on the variance, having established first that, without adjusting for small sample bias, the variance equals approximately one-half of the $J(.)$ estimates, which would suggest that departures from normality are second-order. Overall, we estimate the size of the permanent component of inflation to be below 0.5% based on data for 1947–99 and below 0.8% based on data for 1870–1999. This compares to the lower bound of the (absolute) size of the permanent component of the pricing kernel,

$$J\left(\frac{M_{t+1}^P}{M_t^P}\right) \geq E\left[\log\frac{R_{t+1}}{R_{t+1,1}} - h_t(\infty)\right], \tag{3.7}$$

that we have estimated to be up to about 20% as reported in column 5 in Tables 1, 2, and 3.

Table 5 contains our estimates of the permanent component of inflation. The first two rows display results based on estimating an AR1 or AR2 for inflation and then computing the size of the permanent component as one-half of the (population) spectral density at frequency zero. For the postwar sample, 1947–99, we find 0.21% and 0.15% for the AR1 and AR2, respectively. The third row presents the results using Cochrane's (1988) method that estimates $var\left(\log X_{t+1}^P/X_t^P\right)$ using $\lim_{k\to\infty}(1/k)\,var\,(\log X_{t+k}/X_t)$. For the postwar period, the size of the permanent component is 0.43% or 0.30%, depending on whether $k = 20$ or 30.[23] The table also shows that $J\left(X_{t+k}/X_t\right)/var\left(\log X_{t+k}/X_t\right)$ is approximately 0.5. Note that the roots of the process for inflation reported in Table 5 are not close to one, supporting our implicit assumption that inflation rates are stationary.

A second view about the size of the permanent component can be obtained from inflation-indexed bonds. Such bonds have been traded in the United Kingdom since 1982. Considering that an inflation-indexed bond represents a claim to a fixed number of units of goods, its price provides direct evidence about the real pricing kernel. However, because of the 8-month indexation lag for U.K. inflation-indexed bonds, it is not possible to obtain much information about the short end of the real term structure. Specifically, an inflation-indexed bond with outstanding maturity of less than eight months is effectively a nominal bond; in general, the last eight months of every payout is effectively uninsured against inflation risk. For our estimates, this implies that we will not be able to obtain direct evidence of $E\left(\log R_{t+1,1}\right)$ and $J\left(1/R_{t+1,1}\right)$ in the definition of the size of the permanent component as given in equation (2.12). Because of this, we focus on the bound for the *absolute* size of the pricing kernel as given in equation (3.7). For the nominal kernel, we use average nominal equity returns for $E\log R_{t+1}$, and for

[23] Cochrane's (1988) estimator is defined as $\widehat{\sigma}_k^2 = \frac{1}{k}\left(\frac{1}{T-k}\right)\left(\frac{T}{T-k+1}\right) \cdot \sum_{j=k}^{T}\left[x_j - x_{j-k} - \frac{k}{T}\left(x_T - x_0\right)\right]^2$, with T the sample size, $x = \log X$, and standard errors given by $\left(\frac{4}{3}\frac{k}{T}\right)^{0.5}\widehat{\sigma}_k^2$.

$E \log R_{t+1,\infty}$, we use forward rates and yields for 20 and 25 years, from the Bank of England's estimates of the zero-coupon term structures, to obtain an estimate of the right-hand side of

$$J\left(\frac{M_{t+1}^P}{M_t^P}\right) \geq E\left[\log R_{t+1} - \log R_{t+1,\infty}\right].$$

For the real kernel, we take the average nominal equity return minus the average inflation rate to get $E \log R_{t+1}$; for $E \log R_{t+1,\infty}$, we use real forwards rates and yields from a zero-coupon term structure of inflation-indexed bonds. Thus, the differences in size between nominal and real permanent components are given by the differences between, on one side, the average nominal rate, and on the other side, the average real rate plus average inflation. To the extent that there is a positive risk premium compensating investors for inflation risk in long-term nominal bonds, the size of the permanent component in real kernels will be larger than for nominal kernels.

Table 6 reports estimates for nominal and real kernels. The data are further described in Appendix B. Consistent with our finding that the size of the permanent component of inflation is very small, the differences in size of the permanent components for nominal and real kernels are very small. In fact, for three out of the four point estimates, the size of the permanent component of real kernels is larger than the estimate for the corresponding nominal kernels; for the fourth case, they are basically identical. Compare columns (3) and (6). The corresponding standard errors are always larger than the differences between the results for nominal and real kernels.

3.4. The size of the permanent component in consumption

Following our analyses in Section 2.2 of how various utility functions relate the pricing kernel to consumption, we present here estimates of the size of the permanent component of consumption, obtained directly from consumption data. We end up drawing two conclusions. One is that the size of the permanent component in consumption is about half the size of the overall volatility of the growth rate, which is lower than our estimates of the size of the permanent component of pricing kernels. This suggests that, within a representative agent asset pricing framework, preferences should be such as to magnify the size of the permanent component in consumption. The other conclusion, as noted in Cochrane (1988) for the random walk component in GDP, is that standard errors are large.

Our estimates are presented in Figures 2 and 3 for the periods 1889–1997 and 1946–97, respectively. The first panel shows $J\left(X_{t+k}/X_t\right)/var\left(\log X_{t+k}/X_t\right)$ to be close to 0.5, suggesting, as in subsection 3.3 for inflation, that we can safely use Cochrane's method based on the variance. The second panel shows the estimates of $(1/k)\,var\left(\log X_{t+k}/X_t\right)/var\left(\log X_{t+1}/X_t\right)$ with associated standard error bands. For the period 1889–1997, shown in Figure 2, the estimates seem to stabilize at about 0.5 and 0.6 for k larger than 15. For the postwar period, shown in Figure 3, standard error bands accommodate any possibly reasonable number.

4. Conclusions

In this paper, we derive and estimate a lower bound for the size of the permanent component of asset pricing kernels. The bound is based on rates of long-term bonds. These rates contain the market's forecasts for the growth rate of the marginal utility of wealth over the period corresponding to the maturity of the bond. We find that the permanent component amounts to about 100% of the total volatility of the stochastic discount factor. Standard error bands around this estimate are tight. We also relate the persistence of pricing kernels to the persistence of their determinants, notably consumption. We present sufficient conditions for consumption and preference specifications to imply a pricing kernel with no permanent innovations. We present evidence that the permanent component of pricing kernels is real to a large extent. Finally, we present some evidence that the size of the permanent component in consumption is smaller than the permanent component in pricing kernels. Within a representative agent framework, this evidence points toward utility functions that magnify the size of the permanent component.

References

Abel, Andrew B., 1999, Risk Premia and Term Premia in General Equilibrium, *Journal of Monetary Economics* 43(1): 3–33.

Ait-Sahalia, Yacine, 1996, Nonparametric Pricing of Interest Rate Derivative Securities, *Econometrica* 64(3): 527–60.

Alvarez, Fernando, and Jermann, Urban J., 2000a, Using Asset Prices to Measure the Cost of Business Cycles, available at http://finance.wharton.upenn.edu/~jermann/research.html.

Alvarez, Fernando, and Jermann, Urban J., 2000b, Efficiency, Equilibrium, and Asset Pricing with Risk of Default, *Econometrica* 68(4): 775–97.

Anderson, Nicole, and Sleath, John, 1999, New Estimates of the UK Real and Nominal Yield Curves, *Quarterly Bulletin,* Bank of England, November.

Backus, David K.; Foresi, Silverio; and Telmer, Chris, 1998, Affine Models of Currency Pricing: Accounting for the Forward Premium Anomaly, manuscript, available at http://www.stern.nyu.edu/~dbackus/.

Bansal, Ravi, and Lehmann, Bruce N., 1997, Growth-Optimal Portfolio Restrictions on Asset Pricing Models, *Macroeconomic Dynamics* 1(2): 333–54.

Baxter, Marianne, and Crucini, Mario, 1995, Business Cycles and the Asset Structure of Foreign Trade, *International Economic Review* 36(4): 821–54.

Beveridge, Stephen, and Nelson, Charles R., 1981, A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the "Business Cycle," *Journal of Monetary Economics* 7(2): 151–74.

Billingsley, Patrick, 1995, *Probability and Measure,* 3rd ed., New York: Wiley.

Bliss, Robert R., 1997, Testing Term Structure Estimation Methods, in *Advances in Futures and Options Research,* Vol. 9, Boyle, Phelim; Pennachi, George; and Ritchken, Peter, eds. Greenwich, Conn.: JAI Press, pp. 197–231.

Campbell, John Y., 1996, Understanding Risk and Return, *Journal of Political Economy* 104(2): 298–345.

Campbell, John Y., and Cochrane, John H., 1999, By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior, *Journal of Political Economy* 107(2): 205–51.

Campbell, Cynthia J.; Kazemi, Hossein B.; and Nanisetty, Prasad, 1999, Time-Varying Risk and Return in the Bond Market: A test of a New Equilibrium Pricing Model, *Review of Financial Studies* 12(3): 631–42.

Cochrane, John H., 1988, How Big Is the Random Walk in GNP? *Journal of Political Economy* 96(5): 893–920.

Cochrane, John H., 1992, Explaining the Variance of Price-Dividend Ratios, *Review of Financial Studies* 5(2): 243–80.

Cochrane, John H., and Hansen, Lars Peter, 1992, Asset Pricing Explorations for Macroeconomics, in: *NBER Macroeconomics Annual,* Blanchard Olivier-Jean, and Fischer, Stanley, eds. Cambridge and London, MIT Press, pp. 115–65.

Daniel, Kent D., and Marshall, David A., 2001, Consumption and Asset Returns at Short- and Long-Horizons, manuscript, Northwestern University, Evanston IL.

Dolmas, Jim, 1998, Risk Preferences and the Welfare Cost of Business Cycles, *Review of Economic Dynamics* 1(3): 646–76.

Duffie, Darrell, 1996, *Dynamic Asset Pricing Theory*, Princeton, N.J.: Princeton University Press.

Epstein, Larry G., and Zin, Stanley E., 1989, Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework, *Econometrica* 57(4): 937–69.

Evans, Martin D. D., 1998, Real Rates, Expected Inflation, and Inflation Risk Premia, *Journal of Finance* 53(1): 187–218.

Ferson, Wayne E., and Constantinides, George M., 1991, Habit Persistence and Durability in Aggregate Consumption: Empirical Tests, *Journal of Financial Economics* 29(2): 199–240.

Hall, Anthony D.; Anderson, Heather M.; and Granger, Clive W. J., 1992, A Cointegration Analysis of Treasury Bill Yields, *Review of Economics and Statistics* 74(1): 116–26.

Hamilton, James D., 1994, *Time Series Analysis*, Princeton, N.J.: Princeton University Press.

Hansen, Gary D., 1997, Technical Progress and Aggregate Fluctuations, *Journal of Economic Dynamics and Control* 21(6): 1005–23.

Hansen, Lars Peter, and Jagannathan, Ravi, 1991, Implications of Security Market Data for Models of Dynamic Economies, *Journal of Political Economy* 99(2): 225–62.

Ibbotson Associates, 2000, Stocks, Bonds, Bills, and Inflation—1998 Yearbook, Chicago: Ibbotson Associates, Inc.

Kazemi, Hossein B., 1992, An Intertemporal Model of Asset Prices in a Markov Economy with a Limiting Stationary Distribution, *Review of Financial Studies* 5(1): 85–104.

Luttmer, Erzo G. J., 1996, Asset Pricing in Economies with Frictions, *Econometrica* 64(6): 1439–67.

McCulloch, J. Huston, 1990, The Term Structure of Interest Rates: Appendix B: U.S. Term Structure Data, 1946–87, in: *Handbook of Monetary Economics*, Vol. 1, Handbooks in Economics, no. 8, Friedman, Benjamin M., and Hahn, Frank H., eds. Amsterdam: North-Holland, pp. 672–715.

McCulloch, J. Huston, and Kwon, Heon-Chul, 1993, U.S. Term Structure Data, 1947–1991, Working Paper 93-6, Ohio State University.

Nelson, Charles R., and Plosser, Charles I., 1982, Trend and Random Walks in Macroeconomic Time-Series: Some Evidence and Implications, *Journal of Monetary Economics*, 10(2): 139-62.

Newey, Whitney K., and West, Kenneth D., 1987, A Simple, Positive Semidefinite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55(3): 703–8.

Quah, Danny, 1992, The Relative Importance of Permanent and Transitory Components: Identification and some Theoretical Bounds, *Econometrica* 60(1): 107–18.

Shiller, Robert, 1998, *Annual Data on US Stock Market Prices, Dividends, Earnings, 1871-present with Associated Interest Rate, Price Level and Consumption Data,* http://www.econ.yale.edu/˜shiller/data/chapt26.html.

Snow, Karl N., 1991, Diagnosing Asset Pricing Models Using the Distribution of Asset Returns, *Journal of Finance* 46(3): 955–83.

Stokey, Nancy L., and Lucas, Robert E., Jr., with Edward C. Prescott, 1989, *Recursive Methods in Economic Dynamics,* Cambridge, Mass.: Harvard University Press.

Watson, Mark W., 1986, Univariate Detrending Methods with Stochastic Trends, *Journal of Monetary Economics* 18(1): 49–75.

Appendix A: Proofs

Definition 2.1. The following proposition shows that, under some regularity conditions, the notion of *no permanent innovations* can be stated in several equivalent ways.

Proposition A.1. Assume that (i) $x_k \geq 0$, $E[x_k] = 1$ for all k, (ii) $x_k \to x$ in distribution (converges weakly), and (iii) $0 < \underline{x} < x_k \leq \bar{x} < \infty$ for all k. Then the following conditions are equivalent:

a) $\lim_k E[\log x_k] = 0$,

b) $\lim_k \text{var}[\log x_k] = 0$,

c) $x_k \to 1$ in distribution (weak) or $x = 1$,

d) $\lim_k \text{var}[x_k] = 0$.

Proof. By (ii)

$$\lim_k Ef(x_k) = Ef(x)$$

for any continuous function $f : [\underline{x}, \ \bar{x}] \to R$. Hence,

$$\begin{aligned}
\lim_k E[\log x_k] &= E[\log x], \\
\lim_k \text{var}[\log x_k] &= \text{var}[\log x], \\
\lim_k \text{var}[x_k] &= \text{var}[x].
\end{aligned}$$

By (i) and the previous result, $Ex = 1$. Since log is strictly concave, by Jensen's inequality,

$$E \log x \leq \log Ex = \log Ex = \log(1) = 0$$

with strict inequality if and only if $\text{var}[x] > 0$. Since log is strictly increasing, $\text{var}[x] = 0$ if and only if $\text{var}[\log x] = 0$. Finally, $\text{var}[x] > 0$ if and only if $x \neq Ex = 1$ with positive probability.

Proposition 2.2. By definition,

$$\begin{aligned}
h_t(\infty) &= \lim_{k\to\infty} \left\{ E_t \left(\log \left[\frac{R_{t+1,k}}{R_{t+1,1}} \right] \right) \right\} = \lim_{k\to\infty} \left\{ E_t \log \left[\left(\frac{\frac{E_{t+1}[M_{t+k}]}{M_{t+1}}}{\frac{E_t[M_{t+k}]}{M_t}} \frac{E_t[M_{t+1}]}{M_t} \right) \right] \right\} \\
&= \lim_{k\to\infty} \left\{ E_t \left(\log \left[\frac{E_{t+1}[M_{t+k}]}{M_{t+1}} \frac{E_t[M_{t+1}]}{E_t[M_{t+k}]} \right] \right) \right\} \\
&= \log E_t[M_{t+1}] - E_t \log M_{t+1} + \lim_{k\to\infty} \left\{ E_t \log E_{t+1}[M_{t+k}] - \log E_t[M_{t+k}] \right\}.
\end{aligned}$$

Without permanent innovations, we have,

$$h_t(\infty) = \log E_t[M_{t+1}] - E_t \log M_{t+1}, \tag{A.1}$$
$$= \log E_t\left[\frac{M_{t+1}}{M_t}\right] - E_t \log \frac{M_{t+1}}{M_t} = -\log R_{t+1,1} - E_t \log \frac{M_{t+1}}{M_t}.$$

For any risky gross asset return R_{t+1}, we have that

$$1 = E_t\left[R_{t+1}\frac{M_{t+1}}{M_t}\right],$$

so that

$$\begin{aligned} 0 &= \log(1) = \log\left(E_t\left[R_{t+1}\frac{M_{t+1}}{M_t}\right]\right) \\ &\geq E_t\left[\log\left(R_{t+1}\frac{M_{t+1}}{M_t}\right)\right] = E_t\left[\log\left(\frac{M_{t+1}}{M_t}\right)\right] + E_t[\log(R_{t+1})], \end{aligned}$$

where the inequality follows from the concavity of log. Then

$$-E_t\left[\log\left(\frac{M_{t+1}}{M_t}\right)\right] \geq E_t[\log(R_{t+1})].$$

Combining this expression with equation (A.1), we obtain

$$h_t(\infty) \geq E_t[\log(R_{t+1})] - \log R_{t+1,1}$$

for any asset return R_{t+1}.
Proposition 2.3. First note that

$$\begin{aligned} E_{t+1}[M_{t+k}] &= E_{t+1}\left[M^T_{t+k}\right] E_{t+1}\left[M^P_{t+k}\right]\left(1 + \frac{cov_{t+1}\left[M^T_{t+k}, M^P_{t+k}\right]}{E_{t+1}\left[M^T_{t+k}\right] E_{t+1}\left[M^P_{t+k}\right]}\right) \\ &= E_{t+1}\left[M^T_{t+k}\right] E_{t+1}\left[M^P_{t+k}\right](1 + v_{t+1,t+k}) \end{aligned}$$

and likewise

$$E_t[M_{t+k}] = E_t\left[M^T_{t+k}\right] E_t\left[M^P_{t+k}\right](1 + v_{t,t+k}).$$

Hence,

$$\log \frac{E_{t+1}[M_{t+k}]}{E_t[M_{t+k}]} = \log \frac{E_{t+1}\left[M^T_{t+k}\right]}{E_t\left[M^T_{t+k}\right]} + \log \frac{E_{t+1}\left[M^P_{t+k}\right]}{E_t\left[M^P_{t+k}\right]} + \log \frac{(1 + v_{t+1,t+k})}{(1 + v_{t,t+k})}.$$

Finally, given our hypothesis about $v_{t,t+k}$ we have that

$$\begin{aligned} h_t(\infty) &= E_t[\log E_t[M_{t+1}] - E_t \log M_{t+1}] + \lim_{k\to\infty} \left\{ E_t \log \frac{E_{t+1}[M_{t+k}]}{E_t[M_{t+k}]} \right\} \\ &= E_t[\log E_t[M_{t+1}] - E_t \log M_{t+1}] + \lim_{k\to\infty} \left\{ E_t \log \frac{E_{t+1}\left[M_{t+k}^P\right]}{E_t\left[M_{t+k}^P\right]} \right\}. \end{aligned}$$

Using this equality, the proof of the proposition follows from a straightforward modification of the proof for Proposition 2.2.

Proposition 2.5. Through some algebra, we obtain

$$\begin{aligned} & E_t \left\{ \log \frac{E_{t+1}\left[M_{t+k}^P\right]}{E_t\left[M_{t+k}^P\right]} \right\} = var_{t+1}(\log P_{t+k}) - var_t(\log P_{t+k}) \\ = & -var_t(\pi_{t+k}) - 2\sum_{j=1}^{k-1} cov_t(\pi_{t+k-j}, \pi_{t+k}), \end{aligned}$$

where we have used that $var_t(\pi_{t+j}) = var_{t+1}(\pi_{t+j+1})$. Using the assumption of square summability, we have that

$$\lim_{k\to\infty} cov_t(\pi_{t+k-j}, \pi_{t+k}) = cov(\pi_{s-j}, \pi_s), \text{ and } \lim_{k\to\infty} var_t(\pi_{t+k}) = var(\pi_s).$$

Since by assumption

$$|cov_t(\pi_{t+k-j}, \pi_{t+k})| < \rho^j B,$$

then

$$\begin{aligned} & \lim_{k\to\infty} \left[var_t(\pi_{t+k}) + 2\sum_{j=1}^{k-1} cov_t(\pi_{t+k-j}, \pi_{t+k}) \right] \\ = & \; var(\pi_s) + 2\sum_{j=1}^{\infty} cov(\pi_s, \pi_{s-j}). \end{aligned}$$

Proposition 2.6. Equation (2.11) follows directly from the fact that M^P is a martingale. We then use the result that

$$J(x_{t+1}) = EJ_t(x_{t+1}) + J_t(E_t x_{t+1}),$$

which can be derived through straightforward algebra. Together with the results from Proposition 2.3, equation (2.12) follows. The direction of the inequality is obtained by differentiating with respect to the term representing the growth-optimal return.
Proposition 2.9. By definition,

$$\begin{aligned} E\left(h_t(k) - f_t(k-1)\right) &= E\left\{ \begin{array}{c} (\log V_{t+1k-1} - \log V_{t,k} + \log V_{t,1}) - \\ (\log V_{t,k-1} - \log V_{t,k} + \log V_{t,1}) \end{array} \right\} \\ &= E\left\{\log V_{t+1k-1} - \log V_{t,k-1}\right\} = 0, \end{aligned}$$

where the last line follows from the assumption of stationarity.

Proposition 2.11. By definition,

$$h_t(\infty) - y_t(\infty) = \lim_{k\to\infty} E_t \log R_{t+1,k} - \lim_{k\to\infty} (1/k) \sum_{j=1}^{k} \log R_{t+j,k-(j-1)}.$$

Taking unconditional expectations on both sides, we have that

$$E\left\{h_t(\infty) - y_t(\infty)\right\} = E \lim_{k\to\infty} E_t \log R_{t+1,k} - E \lim_{k\to\infty} (1/k) \sum_{j=1}^{k} \log R_{t+j,k-(j-1)}.$$

Since by assumption expected holding returns and yields, $E_t \log R_{t+1,k}$ and $(1/k)\sum_{j=1}^{k} \log R_{t+j,k-(j-1)}$, are dominated by an integrable random variable and the limit of the right-hand side exists, then by the Lebesgue dominated convergence theorem,

$$\begin{aligned} E \lim_{k\to\infty} E_t \log R_{t+1,k} &= \lim_{k\to\infty} E \log R_{t+1,k}, \\ E \lim_{k\to\infty} (1/k) \sum_{j=1}^{k} \log R_{t+j,k-(j-1)} &= \lim_{k\to\infty} (1/k) \sum_{j=1}^{k} E \log R_{t+j,k-(j-1)}. \end{aligned}$$

Denote the limit

$$\lim_{k\to\infty} E \log R_{t+1,k} = r, \qquad \text{(A.2)}$$

which we assume to be finite. Since, by hypothesis,

$$E \log R_{t+j,k-(j-1)} = E \log R_{t+1,k-(j-1)}$$

for all j, then

$$\lim_{k\to\infty}(1/k)\sum_{j=1}^{k}E\log R_{t+j,k-(j-1)}=\lim_{k\to\infty}(1/k)\sum_{j=1}^{k}E\log R_{t+1,k-(j-1)}=r$$

where the second inequality follows from (A.2). Thus, we have that

$$E\left\{h_t(\infty)-y_t(\infty)\right\}=\lim_{k\to\infty}E\log R_{t+1,k}-\lim_{k\to\infty}(1/k)\sum_{j=1}^{k}E\log R_{t+j,k-(j-1)}=r-r=0.$$

Proposition 2.12. Define $x_{t+1,k}=\frac{E_{t+1}M_{t+k}}{E_tM_{t+k}}$. Given that Λ_t is Markov, under the stated assumptions, we can write

$$\begin{aligned}&\lim_{k\to\infty}E_t\log x_{t+1,k}\\ =\;&\lim_{k\to\infty}\int\left[\log x_k\left(s',s\right)\right]Q\left(ds',s\right),\end{aligned}$$

where

$$x_{t+1,k}=x_k\left(s',s\right)=\frac{\left(T^{k-1}f\right)\left(s'\right)}{\int\left(T^{k-1}\right)f\left(\hat{s}\right)Q\left(d\hat{s},s\right)}.$$

By Jensen's inequality,

$$\int\left[\log x_k\left(s',s\right)\right]Q\left(ds',s\right)\leq 0$$

since

$$\int x_k\left(s',s\right)Q\left(ds',s\right)=\frac{\int\left(T^{k-1}f\right)\left(s'\right)Q\left(ds',s\right)}{\int\left(T^{k-1}\right)f\left(\hat{s}\right)Q\left(d\hat{s},s\right)}=1.$$

By our assumption, $x_k\left(s,s'\right)\geq l>0$; hence, for all k,s,s',

$$-\infty<\log l\leq\log\left(\min\left\{x_k\left(s',s\right),1+\varepsilon\right\}\right)\leq\log\left(1+\varepsilon\right)<\infty$$

for any arbitrary $\varepsilon>0$. Because $\log\left(\min\left\{x_k\left(s',s\right),1+\varepsilon\right\}\right)$ is uniformly bounded, Lebesgue dominated convergence applies. Note that we impose an artificial upper bound, $\log\left(1+\varepsilon\right)$ to get dominated convergence. With this bound, the integral can only get smaller. Thus, if we find that the integral equals zero, its upper bound, the artificial bound could not have mattered. Thus,

$$\begin{aligned} & \int \lim_{k\to\infty} \log\left(\min\left\{x_k\left(s',s\right),1+\varepsilon\right\}\right) Q\left(ds',s\right) \\ = & \lim_{k\to\infty} \int \log\left(\min\left\{x_k\left(s',s\right),1+\varepsilon\right\}\right) Q\left(ds',s\right) \\ \leq & \lim_{k\to\infty} \int \log\left(x_k\left(s',s\right)\right) Q\left(ds',s\right) \leq 0. \end{aligned}$$

Hence, it suffices to show that

$$\int \lim_{k\to\infty} \log\left(\min\left\{x_k\left(s',s\right),1+\varepsilon\right\}\right) Q\left(ds',s\right) = 0.$$

Under (i) or (ii),

$$\begin{aligned} \lim_{k\to\infty} x_k\left(s',s\right) &= \frac{\lim_{k\to\infty}\left(T^{k-1}f\right)\left(s'\right)}{\lim_{k\to\infty}\int\left(T^{k-1}\right)f\left(\hat{s}\right)Q\left(d\hat{s},s\right)} = \frac{\lim_{k\to\infty}\left(T^{k-1}f\right)\left(s'\right)}{\lim_{k\to\infty}\left(T^{k}f\right)\left(s\right)} \\ &= 1. \end{aligned}$$

Thus, because $\log\left(\min\left\{x_k\left(s',s\right),1+\varepsilon\right\}\right)$ is bounded from below,

$$\begin{aligned} & \lim_{k\to\infty} \log\left(\min\left\{x_k\left(s',s\right),1+\varepsilon\right\}\right) \\ = & \log \lim_{k\to\infty} \left(\min\left\{x_k\left(s',s\right),1+\varepsilon\right\}\right) = \\ & \log\left(\min\left\{\lim_{k\to\infty} x_k\left(s',s\right),1+\varepsilon\right\}\right) \\ = & 0. \end{aligned}$$

Proposition 2.13. Properties (a) and (b) define setwise convergence, and setwise convergence implies weak convergence. That is, with $f(.)$ bounded and continuous, expected values converge.

Proposition 2.14. First, we show a lemma that consumption equity prices and consumption equity dividend-price ratios are iid. Then we use the lemma to show that the kernel has permanent innovations.

Lemma A.2. Assume that c_t is iid with cdf F and that $\eta < 1$, where

$$\eta \equiv \max_{c\in[\underline{c},\bar{c}]} \beta\tau^{1-\rho}\left\{\int\left(\frac{c'}{c}\right)^{1-\gamma} dF\left(c'\right)\right\}^{1/\theta}.$$

Then the price of consumption equity, $V_t^c/C_t = f^*(c_t)$, where the function $f^*$is the unique solution to

$$\begin{aligned} T^* f^* &= f^*, \\ f^*(c) &= \psi\, c^{\gamma-1} \end{aligned}$$

for some constant $\psi > 0$ and the operator T is defined as

$$(Tf)(c) = \beta\tau^{1-\rho} \left\{ \int \left(\frac{c'}{c}\right)^{1-\gamma} [f(c') + 1]^{\theta}\, dF(c') \right\}^{1/\theta}.$$

Moreover,

$$V_t^c = \tau^t v(c_t) \equiv f(c_t) \cdot C_t.$$

Proof. Using the pricing kernel (2.17), we obtain that consumption equity must satisfy

$$1 = E_t \left[\left[\beta \left(\frac{C_{t+1}}{C_t} \right)^{-\rho} \right]^{\theta} \left[\frac{V_{t+1}^c + C_{t+1}}{V_t^c} \right]^{\theta} \right]$$

or

$$\begin{aligned} [V_t^c]^{\theta} &= E_t \left[\left[\beta \left(\frac{C_{t+1}}{C_t} \right)^{-\rho} \right]^{\theta} \left[V_{t+1}^c + C_{t+1} \right]^{\theta} \right] \\ &= E_t \left[\left[\beta \left(\frac{\tau c_{t+1}}{c_t} \right)^{-\rho} \right]^{\theta} \left[V_{t+1}^c + \tau^{t+1} c_{t+1} \right]^{\theta} \right]. \end{aligned}$$

Guessing that $V_t^c = v_t \tau^t$, we obtain

$$v_t^{\theta} = E_t \left[\left[(\tau\beta) \left(\frac{\tau c_{t+1}}{c_t} \right)^{-\rho} \right]^{\theta} [v_{t+1} + c_{t+1}]^{\theta} \right]$$

or

$$v_t = \left\{ E_t \left[\left[\tau\beta \left(\frac{\tau c_{t+1}}{c_t} \right)^{-\rho} \right]^{\theta} [v_{t+1} + c_{t+1}]^{\theta} \right] \right\}^{1/\theta},$$

and dividing by c_t on both sides, we obtain

$$\frac{v_t}{c_t} = \left\{ E_t \left[\left[\lambda\beta \left(\frac{\tau c_{t+1}}{c_t} \right)^{1-\rho} \right]^{\theta} \left[\frac{v_{t+1}}{c_{t+1}} + 1 \right]^{\theta} \right] \right\}^{1/\theta}.$$

Now we can write

$$[Tf](c) = \beta\tau^{1-\rho}\left\{\int\left(\frac{c'}{c}\right)^{(1-\gamma)}[f(c')+1]^{\theta}\,dF(c')\right\}^{1/\theta},$$

where f is the price-dividend ratio of the consumption equity:

$$f(c) = \frac{v(c)}{c}.$$

The operator T can be shown to be a contraction: hence, it has a unique fixed point. Moreover, ψ is given by

$$\Psi = \beta\tau^{1-\rho}\left\{\int c'^{(1-\gamma)}[f^*(c')+1]^{\theta}\,dF(c')\right\}^{1/\theta},$$

where f^* satisfies $Tf^* = f^*$. ■

Using Lemma A.2, we can write the return on the consumption equity as

$$R^c_{t+1} = \tau\frac{v(c_{t+1})+c_{t+1}}{v(c_t)} \tag{A.3}$$

Then using (2.17) and (2.20), we get

$$\begin{aligned} x_{t+1,k} &= \frac{E_{t+1}M_{t+k}}{E_tM_{t+k}} = \frac{E_{t+1}\left[\beta^{\theta(t+1)}\,C_{t+1}^{-\rho\theta}\,Y_{t+1}^{\theta-1}\right]}{E_t\left[\beta^{\theta(t+1)}\,C_{t+1}^{-\rho\theta}\,Y_{t+1}^{\theta-1}\right]} \\ &= \left(\tau\frac{v(c_{t+1})+c_{t+1}}{v(c_t)}\right)^{\theta-1}\frac{E_{t+1}\left[c_{t+1}^{-\rho\theta}\left(R^c_{t+2}R^c_{t+3}\cdots R^c_{t+k}\right)^{\theta-1}\right]}{E_t\left[c_{t+1}^{-\rho\theta}\left(R^c_{t+1}R^c_{t+2}\cdots R^c_{t+k}\right)^{\theta-1}\right]}. \end{aligned}$$

Using (A.3) we obtain

$$\begin{aligned} & R^c_{t+2}R^c_{t+3}\cdots R^c_{t+k} \\ = {} & \tau^{k-1}\left(\frac{v(c_{t+2})+c_{t+2}}{v(c_{t+1})}\right)\left(\frac{v(c_{t+3})+c_{t+3}}{v(c_{t+2})}\right)\cdots\left(\frac{v(c_{t+k})+c_{t+k}}{v(c_{t+k-1})}\right) \\ = {} & \tau^{k-1}\frac{1}{v(c_{t+1})}\left(\frac{v(c_{t+2})+c_{t+2}}{v(c_{t+2})}\right)\left(\frac{v(c_{t+3})+c_{t+3}}{v(c_{t+3})}\right)\cdots \\ & \left(\frac{v(c_{t+k-1})+c_{t+k-1}}{v(c_{t+k-1})}\right)(v(c_{t+k})+c_{t+k}), \end{aligned}$$

and likewise for $R^c_{t+1}R^c_{t+3}\cdots R^c_{t+k}$. Placing $E_{t+1}M_{t+k}$ and E_tM_{t+k} into the expression for $x_{t,k}$, we get

$$\begin{aligned}
x_{t+1,k} &= \left(\tau\frac{v(c_{t+1})+c_{t+1}}{v(c_t)}\right)^{\theta-1}\frac{\frac{\tau^{k-1}}{v(c_{t+1})^{1-\theta}}E_t\left[\left(\frac{v(c)+c}{v(c)}\right)^{\theta-1}\right]^{k-2}E_t\left[(v(c)+c)^{\theta-1}c^{-\rho\theta}\right]}{\frac{\tau^k}{v(c_t)^{1-\theta}}E_t\left[\left(\frac{v(c)+c}{v(c)}\right)^{\theta-1}\right]^{k-1}E_t\left[(v(c)+c)^{\theta-1}c^{-\rho\theta}\right]} \\
&= \left(\frac{v(c_{t+1})+c_{t+1}}{v(c_t)}\right)^{\theta-1}\left(\frac{v(c_t)}{v(c_{t+1})^{1-\theta}}\right)^{\theta-1}\frac{1}{E_t\left[\left(\frac{v(c)+c}{v(c)}\right)^{\theta-1}\right]} \\
&= \left(\frac{v(c_{t+1})+c_{t+1}}{v(c_{t+1})}\right)^{\theta-1}/E_t\left[\left(\frac{v(c)+c}{v(c)}\right)^{\theta-1}\right] \\
&= \left(1+\frac{1}{\psi}c_{t+1}^{\gamma-1}\right)^{\theta-1}/E_t\left[\left(1+\frac{1}{\psi}c_{t+1}^{\gamma-1}\right)^{\theta-1}\right],
\end{aligned}$$

which finishes the proof of the proposition.

Proposition 3.1. Define $h_t(k)$ and $y_t(k)$ as

$$\begin{aligned}
h_t(k) &\equiv E_t\left\{\log E_{t+1}\left[\frac{X_{t+k}}{X_{t+1}}\right]-\log E_t\frac{X_{t+k}}{X_t}\right\}+\log E_t\frac{X_{t+1}}{X_t} \\
y_t(k) &\equiv -\left(\frac{1}{k}\right)\log E_t\left[\frac{X_{t+k}}{X_t}\right]+\log E_t\left[\frac{X_{t+1}}{X_t}\right].
\end{aligned}$$

We will need to show that

$$J\left(\frac{X^P_{t+1}}{X^P_t}\right)=\lim_{k\to\infty}\left(\frac{1}{k}\right)J\left(\frac{M_{t+k}}{M_t}\right).$$

In step 1, we obtain that

$$J\left(\frac{X^P_{t+1}}{X^P_t}\right)=E\left[\log E_t\left[\frac{X_{t+1}}{X_t}\right]-E_t\left[\log\frac{X_{t+1}}{X_t}\right]\right]-\lim_{k\to\infty}E\left[h_t(k)\right].$$

In step 2, we obtain that

$$\begin{aligned}
\lim_{k\to\infty}\left(\frac{1}{k}\right)J\left(\frac{M_{t+k}}{M_t}\right) &= E\left[\log E_t\left[\frac{X_{t+1}}{X_t}\right]-E_t\left[\log\left(\frac{X_{t+1}}{X_t}\right)\right]\right]-\lim_{k\to\infty}E\left[y_t(k)\right] \\
&\quad+\lim_{k\to\infty}\left(\frac{1}{k}\right)J\left(E_t\left[\frac{X_{t+k}}{X_t}\right]\right).
\end{aligned}$$

In step 3, we obtain that

$$\lim_{k\to\infty} E\left[h_t\left(k\right)\right] = \lim_{k\to\infty} E\left[y_t\left(k\right)\right].$$

Thus, using (d), we conclude the proof.

Step 1. By definition of J and J_t and assumption (ii),

$$J\left(\frac{X^P_{t+1}}{X^P_t}\right) = E\left[J_t\left(\frac{X^P_{t+1}}{X^P_t}\right)\right] + J\left[E_t\left(\frac{X^P_{t+1}}{X^P_t}\right)\right] = E\left[J_t\left(\frac{X^P_{t+1}}{X^P_t}\right)\right].$$

Then

$$\begin{aligned}
-\lim_{k\to\infty} E\log\frac{E_{t+1}X_{t+k}}{E_tX_{t+k}} &= -\lim_{k\to\infty} E\log\frac{E_{t+1}X^P_{t+k}}{E_tX^P_{t+k}} = -E\left[E_t\log\frac{X^P_{t+1}}{X^P_t}\right] \\
&= E\left[\log E_t\frac{X^P_{t+1}}{X^P_t}\right] - E\left[E_t\log\frac{X^P_{t+1}}{X^P_t}\right] = E\left[J_t\left(\frac{X^P_{t+1}}{X^P_t}\right)\right],
\end{aligned}$$

where the first equality follows from (i), (iii), and (a), the second and third from (ii), the fourth from (ii), and

$$E\left[\lim_{k\to\infty}\log\frac{E_{t+1}X_{t+k}}{E_tX_{t+k}}\right] = -J\left(\frac{X^P_{t+1}}{X^P_t}\right).$$

Using the definition of h_t, we can write

$$\begin{aligned}
h_t\left(k\right) &= E_t\left[\log\frac{E_{t+1}\left[X_{t+k}\right]}{E_t\left[X_{t+k}\right]} + \log\frac{X_t}{X_{t+1}}\right] + \log E_t\frac{X_{t+1}}{X_t} \\
&= E_t\left[\log\frac{E_{t+1}\left[X_{t+k}\right]}{E_t\left[X_{t+k}\right]}\right] + \log E_t\frac{X_{t+1}}{X_t} - E_t\left[\log\frac{X_{t+1}}{X_t}\right].
\end{aligned}$$

Taking unconditional expectation and limits, we have that

$$\lim_{k\to\infty} E\left[h_t\left(k\right)\right] = \lim_{k\to\infty} E\left[\log\frac{E_{t+1}\left[X_{t+k}\right]}{E_t\left[X_{t+k}\right]}\right] + E\left[\log E_t\left[\frac{X_{t+1}}{X_t}\right] - E_t\left[\log\frac{X_{t+1}}{X_t}\right]\right].$$

Using our previous result

$$J\left(\frac{X^P_{t+1}}{X^P_t}\right) = E\left[\log E_t\left[\frac{X_{t+1}}{X_t}\right] - E_t\left[\log\frac{X_{t+1}}{X_t}\right]\right] - \lim_{k\to\infty} E\left[h_t\left(k\right)\right].$$

Step 2.

$$
\begin{aligned}
\left(\frac{1}{k}\right) J\left(\frac{X_{t+k}}{X_t}\right) &= \left(\frac{1}{k}\right)\left\{\log E\left[\frac{X_{t+k}}{X_t}\right] - E\left[\log\left(\frac{X_{t+k}}{X_t}\right)\right]\right\} \\
&= \left(\frac{1}{k}\right)\left\{\log E\left[\frac{X_{t+k}}{X_t}\right] - \sum_{s=0}^{k-1} E\left[\log\left(\frac{X_{t+s+1}}{X_{t+s}}\right)\right]\right\} \\
&= \left(\frac{1}{k}\right)\log E\left[\frac{X_{t+k}}{X_t}\right] - E\left[\log\left(\frac{X_{t+1}}{X_t}\right)\right] \\
&= \left(\frac{1}{k}\right)\left(\log E\left[E_t\left[\frac{X_{t+k}}{X_t}\right]\right] - E\log E_t\left[\frac{X_{t+k}}{X_t}\right]\right) \\
&\quad \left(\frac{1}{k}\right) E\log E_t\left[\frac{X_{t+k}}{X_t}\right] - E\log E_t\left[\frac{X_{t+1}}{X_t}\right] \\
&\quad +E\log E_t\left[\frac{X_{t+1}}{X_t}\right] - E\left[\log\left(\frac{X_{t+1}}{X_t}\right)\right] \\
&= \left(\frac{1}{k}\right) J\left(E_t\left[\frac{X_{t+k}}{X_t}\right]\right) - E\left[y_t\left(k\right)\right] + \\
&\quad +E\left[\log E_t\left[\frac{X_{t+1}}{X_t}\right] - E_t\left[\log\left(\frac{X_{t+1}}{X_t}\right)\right]\right],
\end{aligned}
$$

where the first equality is by definition, the second by properties of logs, the third by (b), the fourth by adding and substracting the relevant quantities, and the fifth by definition of $y_t\left(k\right)$ and $J\left(\cdot\right)$. By taking limits, we obtain the desired result:

$$
\begin{aligned}
\lim_{k\to\infty}\left(\frac{1}{k}\right) J\left(\frac{X_{t+k}}{X_t}\right) &= E\left[\log E_t\left[\frac{X_{t+1}}{X_t}\right] - E_t\left[\log\left(\frac{X_{t+1}}{X_t}\right)\right]\right] - \lim_{k} E\left[y_t\left(k\right)\right] \\
&\quad + \lim_{k\to\infty}\left(\frac{1}{k}\right) J\left(E_t\left[\frac{X_{t+k}}{X_t}\right]\right).
\end{aligned}
$$

Step 3. For any $k,\ y_t\left(k\right)$ can be written as

$$
E\left[y_t\left(k\right)\right] \equiv E\left[-\frac{1}{k}\left[\log E_t\left[\frac{X_{t+k}}{X_t}\right]\right]\right] = -\frac{1}{k}E\left[\sum_{s=0}^{k-1}\log E_t\left[\frac{X_{t+s+1}}{X_t}\right] - \log E_t\left[\frac{X_{t+s}}{X_t}\right]\right],
$$

and by assumption (c),

$$
\lim_{k\to\infty} E\left[y_t\left(k\right)\right] = \lim_{k\to\infty} E\left[\log E_t\left[\frac{X_{t+k}}{X_t}\right] - \log E_t\left[\frac{X_{t+k-1}}{X_t}\right]\right].
$$

Notice that by (b),

$$E_t\left[\frac{X_{t+k-1}}{X_t}\right] = E_{t+1}\left[\frac{X_{t+k}}{X_{t+1}}\right].$$

Then for any k,

$$\begin{aligned} & -E\left[\log E_t\left[\frac{X_{t+k}}{X_t}\right] - \log E_t\left[\frac{X_{t+k-1}}{X_t}\right]\right] \\ = \quad & E\left[h_t\left(k\right)\right] \equiv E\left[E_t\left[\log E_{t+1}\left[\frac{X_{t+k}}{X_{t+1}}\right] - \log E_t\left[\frac{X_{t+k}}{X_t}\right]\right]\right] \end{aligned}$$

and, hence,

$$\lim_{k\to\infty} E\left[y_t\left(k\right)\right] = \lim_{k\to\infty} E\left[h_t\left(k\right)\right].$$

Appendix B: Data

For Table 1, the data on monthly yields of zero-coupon bonds from 1946:12 to 1985:12 comes from McCulloch and Kwon (1990, 1993), who use a cubic spline to approximate the discount function of zero-coupon bonds using the price of coupon bonds. They make some adjustments based on tax effects and for the callable feature of some of the long-term bonds. The data for 1986:1 to 1999:12 are from Bliss (1997). From the four methods available, we use the method proposed by McCulloch and Kwon (1990, 1993). The second part of the sample does not use callable bonds and does not adjust for tax effects. Forward rates and holding periods returns are calculated from the yields of zero-coupon bonds. The one-month short rate is the yield on a one-month zero coupon bond. Yields are available for bonds of maturities going from 1 to 30 years, although for longer maturities, yields are not available for all years. For maturities shorter than 13 years yields are available for all years, for a maturity of 29 years, they are available for approximately half of the sample. The unconditional expectations and standard errors are estimated for each maturity, with all the data available for that maturity, even if the data are discontinuous.

For Table 3, for the United States, equity returns are from Shiller (1998); short-term rates are from Shiller (1998) before 1926, and from Ibbotson Associates (2000) after 1926; and long-term rates are from Campbell (1996) before 1926, from Ibbotson Associates (2000) after 1926.

Ibbotson Associates' (2000) short-term rate is based on the total monthly holding return for the shortest bill not having less than one month maturity. Shiller (1998), for equity returns, used the Standard and Poor Composite Stock Price Index. The short-term rate is the total return to investing for six months at 4-6 month prime commercial paper rates. To adjust for a default premium, we subtract 0.92% from this rate. This is the average difference between T-Bills from Ibbotson Associates (2000) and Shiller's (1998) commercial paper rates for 1926–98.

The data for the United Kingdom is from the Global Financial Data-base. Specifically, the bill index uses the three-month yield on commercial bills from 1800 through 1899 and the yield on treasury bills from 1900 on. The stock index uses Bank of England shares exclusively through 1917. The stock price index uses the Banker's Index from 1917 until 1932 and the Actuaries General/All-Share Index from 1932 on. To adjust for a default premium, we have subtracted 0.037% from the short rate for 1801–99. This is the average difference between the rates on commercial bills and treasury bills for 1900–98.

For Table 5, the inflation rates are computed using a price index from January to December of each year. Until 1926, the price index is the PPI; afterwards, the CPI index from Ibbotson Associates (2000).

For Table 6, the aggregate equity index is from Global Financial Data, further described above. Inflation is based on the CPI, given by Global Financial Data. The Bank of England publishes estimates of nominal and real term structures for forward rates and yields. We use the series indexed by V, corresponding to the Svensson method, because these are available for the whole sample period, 1982–2000. See, http://www.bankofengland.co.uk/ and Anderson and Sleath (1999) for details.

Appendix C: Small Sample Bias

We derive here an estimate of the size of the small sample bias in our estimates in Table 1. For notational convenience, define

$$\frac{a}{b} \equiv \frac{E\left[\log \frac{R_{t+1}}{R_{t+1,1}}\right] - E\left[h_t\left(\infty\right)\right]}{E\left[\log \frac{R_{t+1}}{R_{t+1,1}}\right] + J\left(1/R_{t+1,1}\right)}.$$

In Table 1, we estimate this ratio as the ratio of the estimates $\widehat{a}/\widehat{b} \equiv f\left(\widehat{a}, \widehat{b}\right)$. Using a second-order Taylor series approximation around the true values and considering that $\widehat{a}$ is an unbiased estimator of a, we can write

$$\begin{aligned} E\left[\frac{\widehat{a}}{\widehat{b}}\right] &\simeq \frac{a}{b} + \left[\left(\frac{1}{b^2}\right)\left(\frac{a}{b} var\left(\widehat{b}\right) - cov\left(\widehat{a}, \widehat{b}\right)\right)\right] + \left[-\frac{a}{b^2} E\left(\widehat{b} - b\right)\right] \\ &\simeq \frac{a}{b} + bias_1 + bias_2. \end{aligned}$$

We estimate $bias_1$ directly from the point estimates and the variance-covariance matrix of the underlying sample means. We estimate $bias_2$ by $\frac{1}{2}\frac{\hat{a}}{\hat{b}^2}\frac{1}{\hat{c}^2} Var\left(\hat{c}\right)$, with $\hat{c}$ the sample mean of $1/R_{t,t+1}$. For forward rates, we estimate the size of the overall bias, $bias_1 + bias_2$, as $[-0.004, 0.0073, -0.0012]$ for the three maturities in panel A of Table 1, where a negative number means that our estimate should be increased by that amount. Corresponding values for Panel B,C, and D are $[0.006, 0.0132, 0.0484]$, $[-0.0072, -0.0079, -0.0115]$, and $[-0.0132, -0.0163, -0.0207]$.

Table 1
Size of Permanent Component Based on Aggregate Equity and Zero-Coupon Bonds

Maturity	(1) Equity Premium $E[\log(R/R_1)]$	(2) Term Premium $E[\log(R_k/R_1)]$	(3) J(1/R1) Adjustment for volatility of short rate	(4) Size of Permanent Component J(P)/J	(5) (1)-(2) $E[\log(R/R_1)]$ $-E[\log(R_k/R_1)]$	(6) P[(5) < 0]
A. Forward Rates		E[f(k)]	Holding Period is 1 Year			
25 years	0.0664	-0.0004	0.0005	**0.9996**	0.0669	0.0003
	(0.0182)	(0.0049)	(0.0002)	(0.0710)	(0.0195)	
29 years		-0.0040		**1.0520**	0.0704	0.0033
		(0.0070)		(0.1053)	(0.0259)	
B. Holding Returns		E[h(k)]	Holding Period is 1 Year			
25 years	0.0664	-0.0083	0.0005	**1.1164**	0.0747	0.0124
	(0.0182)	(0.0257)	(0.0002)	(0.3928)	(0.0332)	
29 years		-0.0199		**1.2899**	0.0863	0.0206
		(0.0353)		(0.5611)	(0.0423)	
C. Yields		E[y(k)]	Holding Period is 1 Year			
25 years	0.0664	0.0082	0.0005	**0.8701**	0.0582	0.0017
	(0.0182)	(0.0033)	(0.0002)	(0.0541)	(0.0199)	
29 years		0.0082		**0.8706**	0.0582	0.0055
		(0.0036)		(0.0610)	(0.0229)	
D. Yields		E[y(k)]	Holding Period is 1 Month			
25 years	0.0763	0.0174	0.0004	**0.7673**	0.0588	0.0028
	(0.0190)	(0.0031)	(0.0001)	(0.0717)	(0.0212)	
29 years		0.0168		**0.7755**	0.0595	0.0066
		(0.0033)		(0.0796)	(0.0240)	

For A., term premia (2) are given by one-year forward rates for each maturity minus one-year yields for each month. For B., term premia (2) are given by overlapping holding returns minus one-year yields for each month. For C., term premia (2) are given by yields for each maturity minus one-year yields for each month. For A., B., and C., equity excess returns are overlapping total returns on NYSE, Amex, and Nasdaq minus one year yields for each month. For D., short rates are monthly rates. Newey-West asymptotic standard errors using 36 lags are shown in parentheses. P values in (6) are based on asymptotic distributions. The data are monthly from 1946:12 to 1999:12. See Appendix B for more details.

Table 2
Size of Permanent Component Based on Growth-Optimal Portfolios and 25-Year Zero-Coupon Bonds

	(1) Growth Optimal $E[\log(R/R_1)]$	(2) Term Premium $E[\log(R_k/R_1)]$	(3) J(1/R1) Adjustment for volatility of short rate	(4) Size of Permanent Component J(P)/J	(5) (1)-(2) $E[\log(R/R_1)]$ $-E[\log(R_k/R_1)]$	(6) P[(5) < 0]
A. Market Portfolio						
One-year holding period						
Forward rates	0.0664	-0.0004	0.0005	**0.9996**	0.0681	0.0003
	(0.0182)	(0.0049)	(0.0002)	(0.0710)	(0.0195)	
Holding return		-0.0083		**1.1164**	0.0759	0.0124
		(0.0257)		(0.3928)	(0.0326)	
Yields		0.0082		**0.8701**	0.0595	0.0017
		(0.0033)		(0.0541)	(0.0198)	
One-month holding period						
Yields	0.0763	0.0174	0.0004	**0.7673**	0.0601	0.0028
	(0.0190)	(0.0031)	(0.0001)	(0.0717)	(0.0212)	
B. Growth-Optimal Leveraged Market Portfolio, (Portfolio weight: 3.47 for monthly holding period, 2.14 for yearly)						
One-year holding period						
Forward rates	0.1095	-0.0004	0.0005	**0.9998**	0.11	0.01
	(0.0486)	(0.0049)	(0.0002)	(0.0431)	(0.0473)	
Holding return		-0.0083		**1.0708**	0.1178	0.0163
		(0.0257)		(0.2435)	(0.0551)	
Yields		0.0082		**0.9210**	0.1013	0.0169
		(0.0033)		(0.0386)	(0.0477)	
One-month holding period						
Yields	0.1689	0.0174	0.0004	**0.8946**	0.1515	0.0315
	(0.0818)	(0.0031)	(0.0002)	(0.0518)	(0.0814)	
C. Growth-Optimal Portfolio Based on the 10 CRSP Size-Decile Portfolios						
One-year holding period						
Forward rates	0.1692	-0.0004	0.0005	**0.9999**	0.1697	0.0006
	(0.0528)	(0.0049)	(0.0002)	(0.028)	(0.0525)	
Holding return		-0.0083		**1.0459**	0.1775	0.0021
		(0.0257)		(0.1551)	(0.0621)	
Yields		0.0082		**0.9488**	0.161	0.0009
		(0.0033)		(0.0202)	(0.0518)	
One-month holding period						
Yields	0.2251	0.0174	0.0004	**0.9209**	0.2076	0.0086
	(0.0876)	(0.0031)	(0.0002)	(0.0318)	(0.0872)	

Table 3
Size of Permanent Component Based on Aggregate Equity and Coupon Bonds

		(1) E[logR/R_1] Equity Premium	(2) E[y] Term Premium	E[h]	(3) J(1/R_1) Adjustment	(4) J(P)/J Size of Permanent Component	(5) (1)-(2)	P[(5) < 0]
US	1872-1999	0.0494	0.0034		0.0003	**0.9265**	0.0461	0.0003
		(0.0142)	(0.0028)		(0.0001)	(0.054)	(0.0136)	
				0.0043		**0.9077**	0.0452	0.0006
				(0.0064)		(0.1235)	(0.0139)	
	1926-99	0.0652	0.014		0.0005	**0.7792**	0.0511	0.0049
		(0.0202)	(0.0023)		(0.0001)	(0.0691)	(0.0198)	
				0.0136		**0.7855**	0.0516	0.0061
				(0.0101)		(0.1544)	(0.0206)	
	1946-99	0.0715	0.0122		0.0004	**0.8245**	0.0593	0.0007
		(0.0193)	(0.0025)		(0.0001)	(0.0462)	(0.0185)	
				0.006		**0.9113**	0.0656	0.0004
				(0.0129)		(0.1728)	(0.0196)	

		(1) E[logR/R_1] Equity Premium	(2) E[y] Term Premium	E[h]	(3) J(1/R_1) Adjustment	(4) J(P)/J Size of Permanent Component	(5) (1)-(2)	P[(5) < 0]
UK	1801-1998	0.0239	0.0002		0.0003	**0.9781**	0.0237	0.0014
		(0.0083)	(0.0020)		(0.0001)	(0.0808)	(0.0079)	
				0.0036		**0.8361**	0.0202	0.0053
				(0.0058)		(0.2228)	(0.0079)	
	1926-98	0.0550	0.0111		0.0008	**0.7870**	0.0439	0.0070
		(0.0173)	0.0031		(0.0002)	(0.0899)	(0.0179)	
				0.0131		**0.7516**	0.0419	0.0091
				0.0130		(0.2189)	(0.0177)	
	1946-98	0.0604	0.0092		0.0007	**0.8370**	0.0511	0.0074
		(0.0198)	(0.0038)		(0.0002)	(0.0904)	(0.0210)	
				0.0018		**0.9583**	0.0585	0.0006
				(0.0143)		(0.2289)	(0.0181)	

(1) Average annual log return on equity minus average short rate for the year.
(2) Average yield on long-term government coupon bond minus average short rate for the year.
(3) Average annual holding period return on long-term government coupon bond minus average short rate for the year.
Newey-West asymptotic standard errors with 5 lags are shown in parentheses. See Appendix B for more details.

Table 4
Required Persistence for Bonds with Finite Maturities

Maturity (years)	Term spread 0	0.50%	1%	1.50%
10	1.0000	0.9986	0.9972	0.9957
20	1.0000	0.9993	0.9987	0.9980
30	1.0000	0.9996	0.9991	0.9987

Table 5
The Size of the Permanent Component due to Inflation

1947-99		AR(1)	AR(2)	σ^2	Size of permanent component	
AR1		0.66		0.0005	**0.0021**	(0.0009)
AR2		0.87	-0.24	0.0004	**0.0015**	(0.0006)
(1/2k) var(log P_{t+k}/P_t)	k=20				**0.0043**	(0.0031)
	k=30				**0.0030**	(0.0027)
J(P_{t+k}/P_t) / var(log P_{t+k}/P_t)		(k=20)	0.46			
		(k=30)	0.45			

1870-1999		AR(1)	AR(2)	σ^2	Size of permanent component	
AR1		0.28		0.0052	**0.0049**	(0.0013)
AR2		0.27	0.00	0.0052	**0.0050**	(0.0006)
(1/2k) var(log P_{t+k}/P_t)	k=20				**0.0077**	(0.0035)
	k=30				**0.0067**	(0.0038)
J(P_{t+k}/P_t) / var(log P_{t+k}/P_t)		(k=20)	0.47			
		(k=30)	0.48			

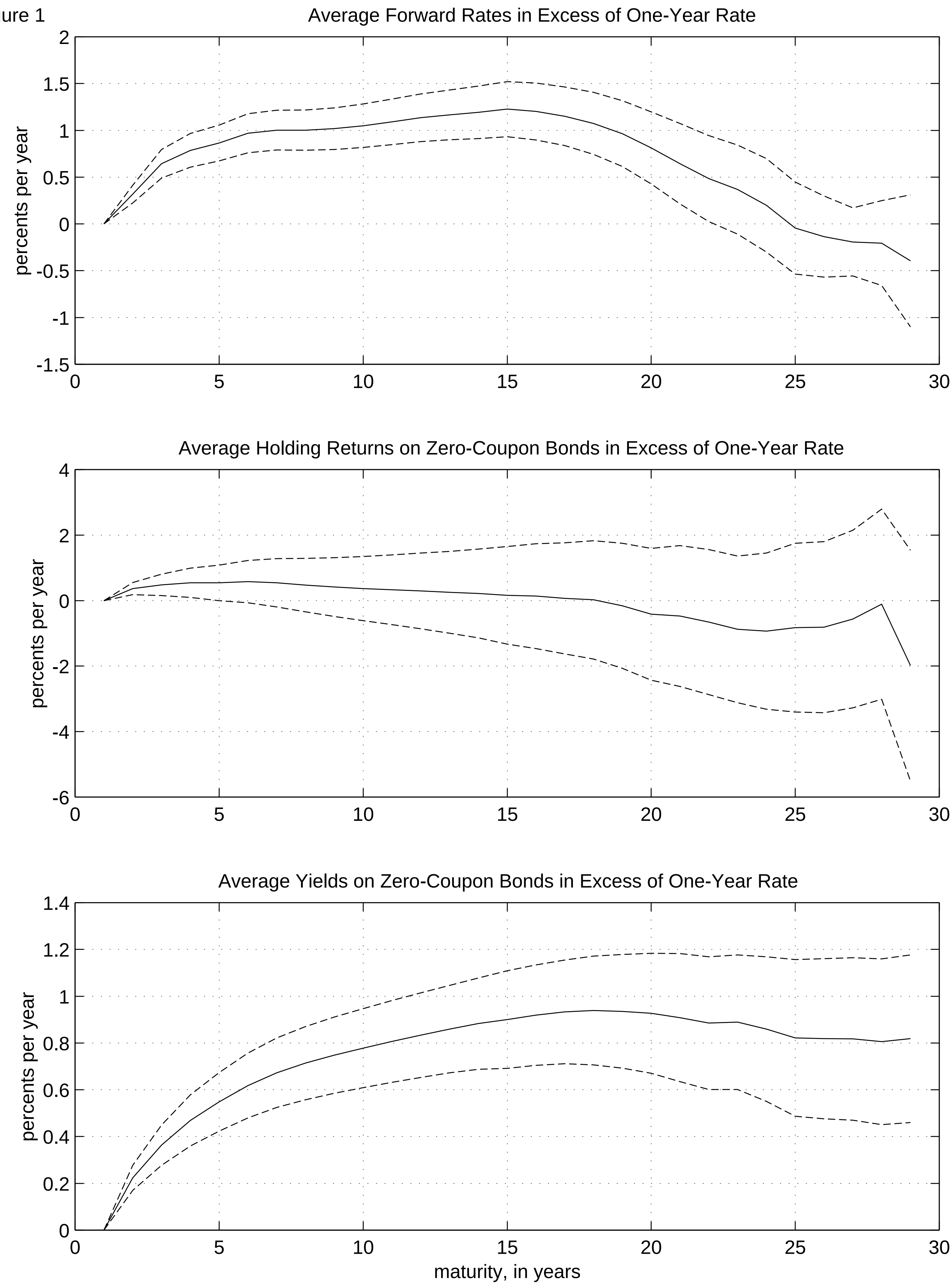
For the AR(1) and AR(2) cases, the size of the permanent component is computed as one-half of the spectral density at frequency zero. The numbers in parentheses are standard errors obtained through Monte Carlo simulations. For (1/2k) var(log P_{t+k}/P_t), we have used the methods proposed by Cochrane (1988) for small sample corrections and standard errors. See our discussion in the text for more details.

Table 6
Inflation-Indexed Bonds and the Size of the Permanent Component of Pricing Kernels, U.K. 1982-99

		Nominal Kernel			Real Kernel			
	(1)	(2)		(3) (1)-(2)	(4)	(5)		(6) (1)-(4)-(5)
Maturity years	Equity	Forward	Yield	Size of Permanent Component	Inflation Rate	Forward	Yield	Size of Permanent Component
	E[log(R)]	E[log(F)]	E[log(Y)]	J(P)	E[log(π)]	E[log(F)]	E[log(Y)]	J(P)
20	0.1706 (0.0197)	0.0781 (0.0038)		**0.0924** (0.0206)	0.0422 (0.0063)	0.0343 (0.0022)		**0.0941** (0.0229)
			0.0836 (0.0053)	**0.0870** (0.0193)			0.0348 (0.0017)	**0.0936** (0.0223)
25		0.0762 (0.0040)		**0.0944** (0.0212)		0.0342 (0.0023)		**0.0943** (0.0230)
			0.0815 (0.0046)	**0.089** (0.0200)			0.0347 (0.0018)	**0.0937** (0.0224)

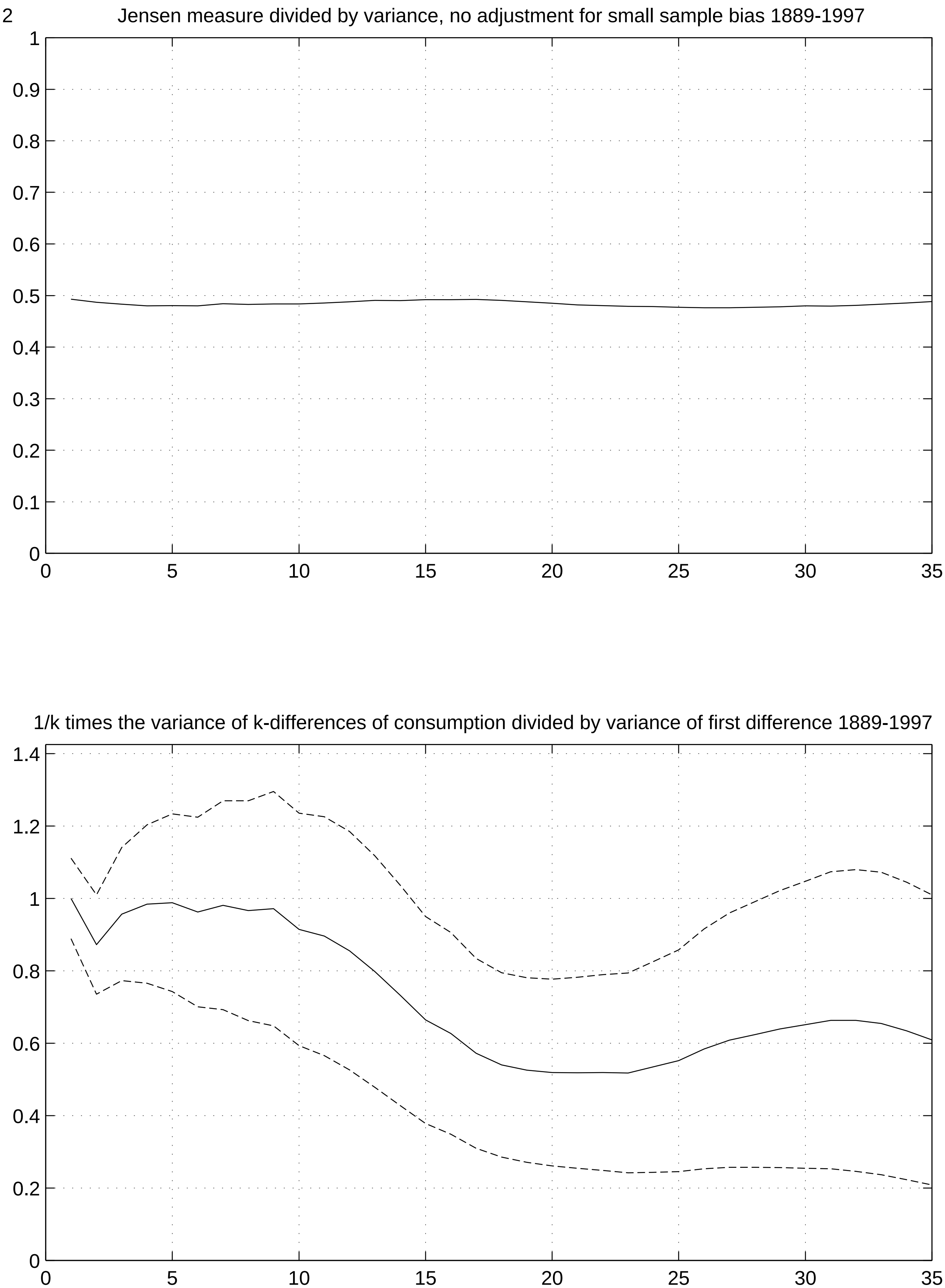
Real and nominal forward rates and yields are from the Bank of England. Stock returns and inflation rates are from Global Financial Data. Asymptotic standard errors, given in parenthesis, are computed with the Newey-West method with 3 years of lags and leads.

Figure 1



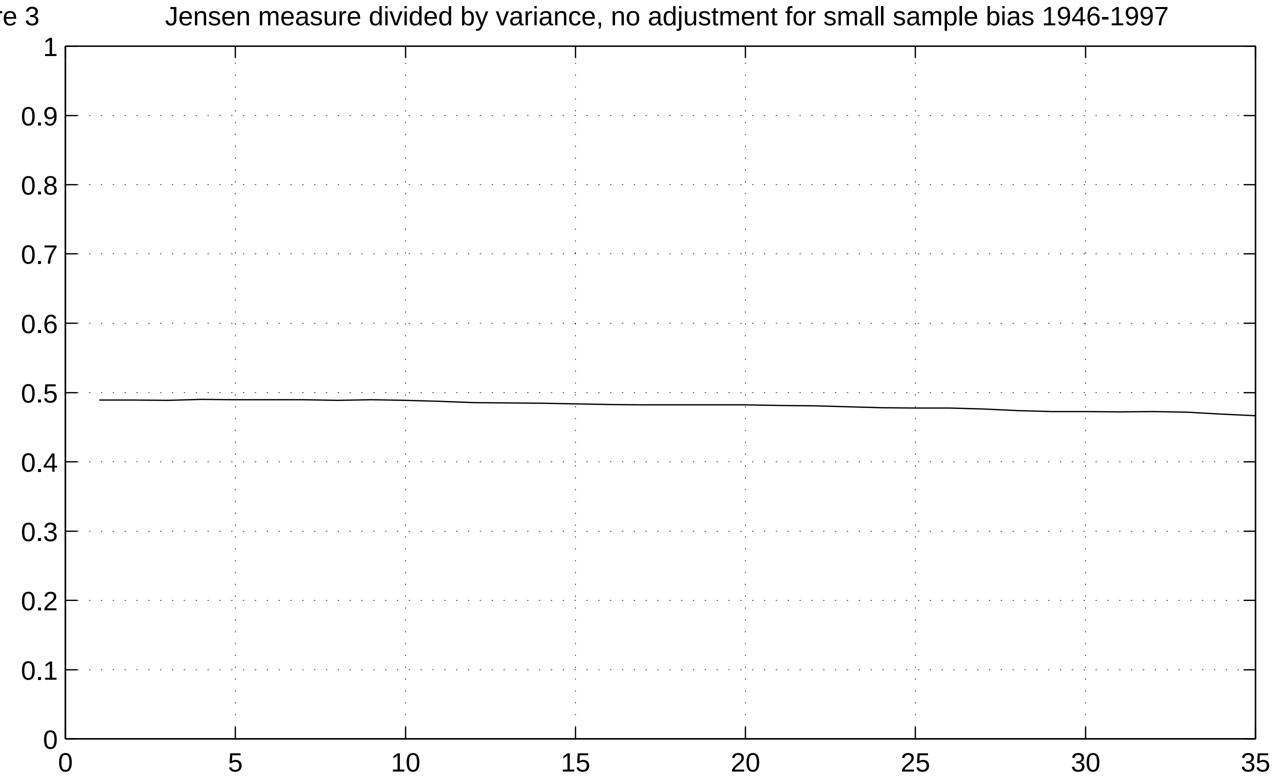
U.S., 1946:12 - 1999:12. Bands showing 1 asymptotic standard error.

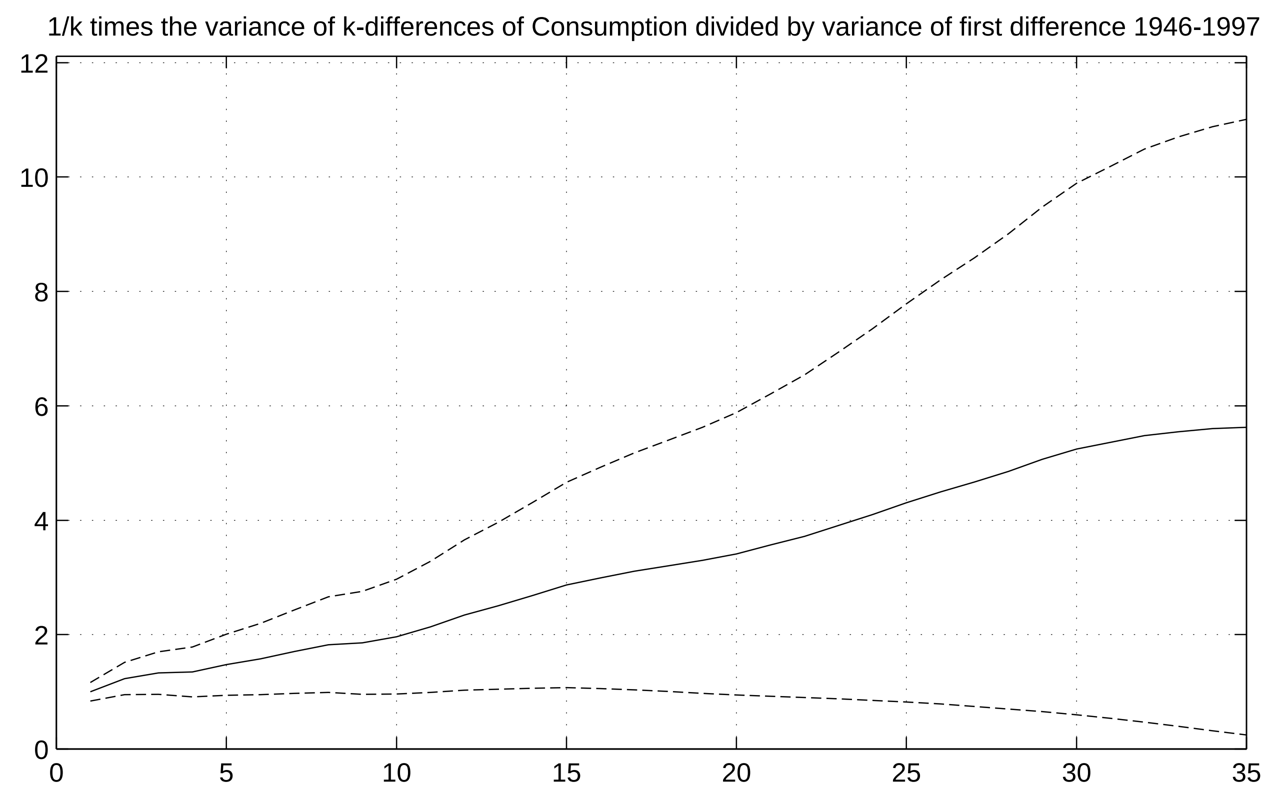
Figure 2



Bands showing 1 asymptotic standard error.

Figure 3


Jensen measure divided by variance, no adjustment for small sample bias 1946-1997
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
5
10
15
20
25
30
35


1/k times the variance of k-differences of Consumption divided by variance of first difference 1946-1997
12
10
8
6
4
2
0
0
5
10
15
20
25
30
35