USING ASSET PRICES TO MEASURE THE PERSISTENCE OF THE MARGINAL UTILITY OF WEALTH

By Fernando Alvarez and Urban J. Jermann[1]

We derive a lower bound for the volatility of the permanent component of investors' marginal utility of wealth, or more generally, 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 volatile; its volatility is about at least as large as the volatility of the stochastic discount factor. A related measure for the transitory component suggest it to be considerably less important. We also show that, for many cases where the pricing kernel is a function of consumption, innovations to consumption need to have permanent effects.

Keywords: Pricing kernel, stochastic discount factor, permanent component, unit roots

1 Introduction

The absence of arbitrage opportunities implies the existence of a *pricing kernel*, that is, a stochastic process that assigns values to state-contingent payments. As is well known, asset pricing kernels can be thought of as investors' marginal utility of wealth in frictionless markets. Since the properties of such processes are important for asset pricing, they have been the subject of much recent research.[2] Our focus is on the persistence properties of pricing kernels, these are key determinants of the prices of long-lived securities.

The main result of this paper is to derive and estimate a lower bound for the volatility 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 pricing kernels to be very volatile; its volatility is about at least as large as the volatility of the stochastic discount factor. A related bound that measures the volatility of the transitory component suggests it to be considerably less important than the permanent component.

Our results complement the seminal 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. They find that, to be

consistent with the high Sharpe ratios in the data, stochastic discount factors have to be very volatile. We find that, to be consistent with the low returns on long-term bonds relative to equity, the permanent component of pricing kernels have to be very large. This property is important, because the low frequency components of pricing kernels are important determinants of the prices of long-lived securities such as stocks. Recent work on asset pricing has highlighted the need for a better understanding of these low frequency components, see for instance Bansal and Yaron (2003), and Hansen, Heaton and Li (2004). Our results are also related to Hansen and Scheinkman (2003), where they present a general framework for linking the short and long run properties of asset prices.

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. On this point, 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 a 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 issues, the persistence properties of the processes specifying economic variables can be very important. For instance, on the issue of the welfare costs of economic uncertainty, see Dolmas (1998) ; 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). The lesson from our analysis for these cases and many related studies of dynamic general equilibrium models is that models should be calibrated so as to generate macroeconomic time-series with important permanent components.

Following Nelson and Plosser (1982) a large body of literature has tested macroeconomic time-series for stationarity versus unit roots.[3] More recently, a large and growing literature on structural VARs is using identifying assumptions based on restricting the origin of permanent fluctuations in macroeconomic variables to certain types of shocks. The relationship between such structural shocks and macroeconomic variables is then compared to the implications of different classes of macroeconomic models. See for instance Shapiro and Watson (1988), Blanchard and Quah (1989), and more recently

Gali (1999), Fisher (2002), Christiano, Eichenbaum and Vigfusson (2002). The identification strategies used in this literature hinge critically on the presence of unit roots in the key macroeconomic time series. The results in our paper provide validation for this approach by presenting new evidence about the importance of permanent fluctuations. We introduce new information about persistence from the prices of long-term bonds. Prices of long-term bonds are particularly informative about the persistence of pricing kernels because they are the market's forecast of the long-term changes in the pricing kernel.

As a second set of results, we measure the volatility of the permanent component in consumption directly, and compare it to the volatility of the permanent component of pricing kernels. This can provide guidance for the specification of functional forms of the marginal utility of wealth.[4] Specifically, we find the volatility 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 a preview of the main results. Section 3 presents theoretical results. Section 4 presents empirical evidence. Section 5 relates pricing kernels and

aggregate consumption. Section 6 concludes. Proofs are in Appendix A. Appendix B describes the data sources. Appendix C addresses a small sample bias.

2 Definitions and Preview of the Main Result

In this section, we start with some key definitions and assumptions. Then, to preview the main theoretical result of the paper, we state without derivation an expression for the lower bound of the permanent component of the stochastic discount factor. We compute this lower bound for two benchmark cases: one with only permanent movements, and one with only transitory movements.

Let D_{t+k} be a state-contingent dividend to be paid at time $t+k$ and let $V_t(D_{t+k})$ 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

$$V_t(D_{t+k}) = \frac{E_t(M_{t+k} \cdot D_{t+k})}{M_t}. \quad (1)$$

[5]

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.[6] We use R_{t+1} for the gross return on a generic portfolio held from t to $t+1$; hence,(1) implies that

$$1 = E_t\left(\frac{M_{t+1}}{M_t} \cdot R_{t+1}\right). \tag{2}$$

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 $t+k$,

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

The holding return on this discount bond is the ratio of the price at which the bond is sold, $V_{t+1}\left(1_{t+k}\right)$, to the price at which it was bought, $V_t\left(1_{t+k}\right)$. With this convention, $V_t\left(1_t\right) \equiv 1$. Thus, for $k \geq 2$ the return consists solely of capital gains; for $k = 1$, the return is risk free. In this paper we focus on the limiting long term bond, which has return $R_{t+1,\infty} \equiv \lim_{k\to\infty} R_{t+1,k}$.

Throughout the paper we maintain the assumption that stochastic discount factors M_{t+1}/M_t and returns R_{t+1} are jointly stationary and ergodic. An immediate implication of the stationarity of stochastic discount factors is that all bond returns are stationary. The assumption of stationarity of returns is standard in the asset pricing literature. In section 4 we review some of the evidence on the stationarity of interest rates. Under our maintained assumption about stationarity we find that pricing kernels M_t have a

large permanent component. Alternatively, if we were to consider M_{t+1}/M_t as non-stationary, then M_t would not be stationary either. To use a time series analogy, if $\log M_{t+1} - \log M_t$ were to have unit root, then $\log M_t$ would be integrated at least of order 2.

Below we decompose the pricing kernel M_t into two components:

$$M_t = M_t^P \; M_t^T$$

where M_t^P is a martingale, so it captures the permanent part of M_t, and M_t^T is the transitory component of M_t. The main result of the paper is that the volatility of the growth rate of the permanent component, M_{t+1}^P/M_t^P, relative to the volatility of the stochastic discount factor, M_{t+1}/M_t, is at least as large as

$$\frac{E \log R_{t+1}/R_{t+1,1} - E \log R_{t+1,\infty}/R_{t+1,1}}{E \log R_{t+1}/R_{t+1,1} + L\left(1/R_{t+1,1}\right)} \tag{3}$$

where R_{t+1} is the return of any asset. $L\left(1/R_{t+1,1}\right)$ is a measure of the volatility of the short term interest rate to be described in detail below. For this preliminary discussion note that $L = 0$ if interest rates have zero variance and otherwise $L > 0$. The numerator of this expression is the difference between two (log) excess returns, or two risk premiums. As is easily seen, if the term premium for the bond with infinitely long maturity is positive,

$E \log R_{t+1,\infty}/R_{t+1,1} > 0$, this expression is maximized by selecting the asset with the highest expected log excess return $E \log R_{t+1}/R_{t+1,1}$.

We now compute the lower bound for two examples for which it is obvious what the volatility of the permanent component of the pricing kernel is. Consider an investor with time separable expected utility, and consider two consumption processes: iid consumption growth and iid consumption level. The pricing kernel is

$$M_{t+1} = \left(\frac{1}{1+\rho}\right)^t U'(c_t) = \left(\frac{1}{1+\rho}\right)^t c_t^{-\gamma}$$

where U has CRRA γ.

Example 1. Assume that c_{t+1}/c_t is iid. Clearly M_t has only permanent shocks. In this case, it is easy to verify that interest rates $R_{t+1,1}$ are constant, which implies that $L(1/R_{t+1,1}) = 0$, and that

$$\log(R_{t+1,k}/R_{t+1,1}) = 0,$$

so that all term premiums are zero. With these values, expression (3) is equal to 1, so that the volatility of the permanent component of the stochastic discount factor is, indeed, at least as large as the volatility of the stochastic discount factor.

Example 2. Assume that c_{t+1} is iid. Clearly M_t has no permanent com-

ponent. In this case, neither short term interest rates nor returns on long term bonds are constant in general. Indeed,

$$\begin{aligned} R_{t+1,1} &= (1+\rho)\frac{U'(c_t)}{E\left[U'(c_{t+1})\right]}, \text{ and} \\ R_{t+1,k} &= (1+\rho)\frac{U'(c_t)}{U'(c_{t+1})} = M_t/M_{t+1} \text{ for } k \geq 2, \end{aligned}$$

that is, for $k > 2$, the holding return equals the inverse of the stochastic discount factor. It is now easy to show that the highest lower bound computed from expression (3) is attained by choosing the return $R_{t+1} = R_{t+1,k}$ for $k \geq 2$, and that this lower bound equals 0. Indeed, ruling out arbitrage implies that for any return R_{t+1}

$$E_t\left(\frac{M_{t+1}}{M_t}R_{t+1}\right) = 1.$$

Using Jensen's inequality

$$0 = \log E_t\left(\frac{M_{t+1}}{M_t}R_{t+1}\right) \geq E_t \log\left(\frac{M_{t+1}}{M_t}R_{t+1}\right)$$

which implies

$$E_t \log R_{t+1} \leq E_t \log \frac{M_t}{M_{t+1}},$$

with equality if R_{t+1} and M_t/M_{t+1} are proportional. Thus, because $R_{t+1,k} = M_t/M_{t+1}$, for $k \geq 2$ no log return is higher than the log return of long term

bonds. Setting $R_{t+1} = R_{t+1,k}$ for $k \geq 2$ gives the highest lower bound (3), and its value will be zero. Hence we have verified that the bound shows that, for the case where the level of consumption is iid, there is no permanent component.

3 Theoretical Results

In this section, we first show an existence result for the multiplicative decomposition of M_t into a transitory and permanent component, and we derive a lower bound for the volatility of the permanent component. We then present a related bound for the volatility of the transitory component. We also present a proposition that guarantees the applicability of our bound for the permanent component to any appropriate multiplicative decomposition under some regularity assumptions. Finally, we compare our bound to a result by Cochrane and Hansen (1992) about the conditional and unconditional volatility of stochastic discount factors.

We start with two conditions under which we can decompose the kernel into permanent and transitory components properly defined. First, (1)

assume there is a number β such that

$$0 < \lim_{k \to \infty} V_t \left(1_{t+k}\right) / \beta^k < \infty,$$

for all t. In the language of Hansen and Scheinkman (2003), the number β is the dominant eigenvalue of the pricing operator. Assumption (1) can be violated if either the limit does not exists or if it takes the values 0 or ∞. The existence of the limit imposes a regularity condition on the shape of the term structure for large k. Specifically, it requires that the yield $-(1/k) \log V_t \left(1_{t+k}\right)$ converges fast enough as $k \to \infty$. The limit can take the value 0 or ∞ if bond prices are non-stationary. For instance, consider the case where after date s there are only two possible outcomes: either the yields of bonds of all maturities are equal to $-\log \bar{\beta}$ or they are equal to $-\log \underline{\beta}$. In this case there is no β for which the limit in assumption (1) is strictly positive and finite. Since we have assumed throughout that bond prices are stationary, this possibility is ruled out.

Second, (2) assume that for each $t+1$ there is a random variable x_{t+1} such that

$$\left(M_{t+1} / \beta^{t+1}\right) V_{t+1} \left(1_{t+1+k}\right) / \beta^k \leq x_{t+1} \quad a.s.,$$

with $E_t x_{t+1}$ finite for all k. Assumption (2) strengthens assumption (1).

Instead of requiring that $V_t(1_{t+k})/\beta^k$ has a finite limit, assumption (2) requires that for each k its product with the marginal valuation be bounded by a variable that has a finite conditional expectation.

Proposition 1 *Under assumptions (1) and (2), there is a unique decomposition*

$$M_t = M_t^T \; M_t^P$$

with $E_t M_{t+1}^P = M_t^P$ and

$$\begin{aligned} M_t^P &= \lim_{k\to\infty} E_t M_{t+k}/\beta^{t+k} \\ M_t^T &= \lim_{k\to\infty} \beta^{t+k}/V_t(1_{t+k}). \end{aligned}$$

Due to assumption (1) M_t^P is well defined, strictly positive and finite. Assumption (2) is used to establish that M_t^P as defined above is a martingale. The decomposition obtained through Proposition 1 is unique given its constructive nature.

The value of the permanent component is the expected value of the process M in the long run, relative to its long-term drift β. We call M_t^P the permanent component because it is unaffected by information at t that does not lead to revisions of the expected value of M in the long run. The decomposition in Proposition 1 is analogous to the one by Beveridge and Nelson

(1981). Beveridge and Nelson's decomposition is additive instead of multiplicative and their permanent component is a random walk while in our decomposition the permanent component is a martingale.

The component M_t^T is a scaled long term interest rate. Given our stationarity assumption for the stochastic discount factor, interest rates inherit this property, and interpreting M_t^T as containing only transitory components follows naturally. This stationarity property for M^T is again linked to a related property in the Beveridge and Nelson decomposition. Interest rates are a function of the expected growth rate of the pricing kernel. Thus, assuming stationarity for interest rates is similar to the assumption behind the Beveridge and Nelson decomposition that growth rates are stationary while levels are not.

Nothing in Proposition 1 rules out the possibility that there exist other decompositions of M into two parts with one being a martingale and the other containing transitory components. Such alternative decompositions could exist independently of whether assumption (1) and (2) apply. With assumptions (1) and (2) holding, it might still be possible to construct a decomposition in another way. Alternatively, Proposition 1 has nothing to say for the case where assumptions (1) and (2) would not hold. However, as

we will show later in this section, our volatility bounds also apply to such decompositions more generally.

In order to characterize the importance of permanent and transitory components we use $L_t(x_{t+1}) \equiv \log E_t x_{t+1} - E_t \log x_{t+1}$, and $L(x_{t+1}) \equiv \log E x_{t+1} - E \log x_{t+1}$ as measures of conditional and unconditional volatility of x_{t+1}. Throughout the rest of the paper we refer to the expected values of different random variables without stating explicitly the assumption that these random variables are integrable. The following result can then be shown.

Proposition 2 *Assume that assumptions (1) and (2) hold, then (i) the conditional volatility of the permanent component satisfies*

$$L_t\left(\frac{M^P_{t+1}}{M^P_t}\right) \geq E_t \log R_{t+1} - E_t \log R_{t+1,\infty}, \tag{4}$$

for any positive return R_{t+1}. Furthermore, (ii) the unconditional volatility of the permanent component satisfies,

$$\frac{L\left(\frac{M^P_{t+1}}{M^P_t}\right)}{L\left(\frac{M_{t+1}}{M_t}\right)} \geq \min\left\{1, \frac{E\left(\log \frac{R_{t+1}}{R_{t+1,1}}\right) - E\left(\log \frac{R_{t+1,\infty}}{R_{t+1,1}}\right)}{E\left(\log \frac{R_{t+1}}{R_{t+1,1}}\right) + L\left(1/R_{t+1,1}\right)}\right\} \tag{5}$$

for any positive R_{t+1} such that $E\left(\log \frac{R_{t+1}}{R_{t+1,1}}\right) + L\left(1/R_{t+1,1}\right) > 0$.

Inequality (4) bounds the conditional volatility of the permanent component in the same units as L by the difference of any expected log excess

return relative to the return of the asymptotic discount bond. Inequality (5) bounds the unconditional volatility of the permanent component relative to the one of the stochastic discount factor. As we further discuss below, equation (5) describes a property of the data that is closely related to Cochrane's (1988) size of the random walk component.

To better understand the measure of volatility $L(x)$, note that if $var(x) = 0$, then $L(x) = 0$; the reverse is not true, as higher-order moments than the variance also affect $L(x)$. More specifically, the variance and $L(x)$ are special cases of the general measure of volatility $f(Ex) - Ef(x)$, where $f(\cdot)$ is a concave function. The statistic $L(x)$ is obtained by making $f(x) = \log x$, while for the variance, $f(x) = -x^2$. It follows that if a random variable x_1 is more risky than x_2 in the sense of Rothschild-Stiglitz, then $L(x_1) \geq L(x_2)$ and, of course, $var(x_1) \geq var(x_2)$.[7] As a special case, if x is lognormal, then $L(x) = 1/2\, var(\log x)$. $L(x)$ has been used to measure income inequality and it is also known as Theil's second entropy measure (Theil 1967). Based on Proposition 2, Luttmer (2003) has worked out a continuous-time version of our volatility bound and shown its relationship to Hansen and Jagannathan's volatility bound for stochastic discount factors.

The following proposition characterizes the transitory component, an up-

per bound to its relative volatility can then be easily obtained along the lines of Proposition 2.

Proposition 3 *Under assumptions (1) and (2),* $R_{t+1,\infty} = M_t^T / M_{t+1}^T$, *and*

$$\frac{L\left(M_{t+1}^T / M_t^T\right)}{L\left(M_{t+1}/M_t\right)} \leq \frac{L\left(1/R_{t+1,\infty}\right)}{E\log\left(R_{t+1}/R_{t+1,1}\right) + L\left(1/R_{t+1,1}\right)}$$

for any positive R_{t+1} *such that* $E\left[\log \frac{R_{t+1}}{R_{t+1,1}}\right] + L\left(1/R_{t+1,1}\right) > 0$.

Our decomposition does not require the permanent and transitory components to be independent. Thus, knowing the amount of transitory volatility relative to the overall volatility of the stochastic discount factors adds independent information in addition to knowing the volatility of the permanent component relative to the volatility of the stochastic discount factor. As we will see below, given data availability reasons, we will be able to learn more about the volatility of the permanent component than about the volatility of the transitory one. Kazemi (1992), in a related result, has shown that in a Markov economy with a limiting stationary distribution, $R_{t,t+\infty}$ equals M_t/M_{t+1}.

As we mentioned above, the decomposition derived in Proposition 1 is not necessarily the only one yielding a martingale and a transitory component, and thus the bounds derived above might not necessarily apply to other cases.

To strengthen our results, we show here that the volatility bounds derived in Proposition 2 are valid for any decomposition of the pricing kernel into a martingale and a transitory component under some regularity assumptions. In order to do this, we need a definition for the transitory component, which we describe as having no permanent innovations.

Definition. We say that a random variable indexed by time, X_t, has *no permanent innovations* if

$$\lim_{k\to\infty} \frac{E_{t+1}\left(X_{t+k}\right)}{E_t\left(X_{t+k}\right)} = 1, \text{ a.s. for all } t. \tag{6}$$

We say that there are no permanent innovations because, as the forecasting horizon k becomes longer, information arriving at $t+1$ will not lead to revisions of the forecasts made with current period t information. Alternatively, condition (6) says that innovations in the forecasts of X_{t+k} have limited persistence, since their effect vanishes for large k. As can easily be seen, a linear process that is covariance-stationary, has no permanent innovations.

Proposition 4 *Assume that the kernel has a component with transitory innovations M_t^T, that is one for which (6) holds, and a component with permanent innovations M_t^P that is a martingale, so that*

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

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

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

assume that

$$\lim_{k\to\infty} \frac{(1+v_{t+1,t+k})}{(1+v_{t,t+k})} = 1 \ a.s..$$

Then the bounds in equations (4) and (5) apply.

For an example illustrating this result see the supplementary material to this article.

Following Cochrane and Hansen (1992, pp 134-137) one can derive the following lower bound for the fraction of the variance of the stochastic discount factor accounted for by its innovations:

$$\frac{E\left[var_t\left(\frac{M_{t+1}}{M_t}\right)\right]}{var\left(\frac{M_{t+1}}{M_t}\right)} \geq 1 - \frac{1}{\left(\frac{E|R_{t+1}-R_{t+1,1}|}{\sigma(R_{t+1})}\right)^2} \frac{var\left[V_t\left(1_{t+1}\right)\right]}{\left(E\left[V_t\left(1_{t+1}\right)\right]\right)^2},$$

where R_{t+1} stands for any return. This lower bound takes a value of about 0.99 when R_{t+1} is an asset with a Sharpe ratio of 0.5 and one-period interest volatility is low, such as $var\left[V_t\left(1_{t+1}\right)\right] = 0.05^2$. A natural interpretation of this result is in terms of a persistent and transitory components, and the conclusion would be in line with our main result. However, such an interpretation is not necessarily correct. Indeed, one can easily construct

examples of pricing kernels with one period interest rates that are arbitrarily smooth and that have no permanent innovations. The example we use in section 4.3 below is of this type. Nevertheless, our results confirm such a natural interpretation of the findings of Cochrane and Hansen. We learn from our analysis that the reason the two results can have a similar interpretation is because the term premiums for long term bonds are very small.

3.1 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 this purpose, we show in this section 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 maturing k periods from now relative to a one-period deposit now:

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

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. Assuming that bond prices have means that are independent of calendar time, so that $EV_t(1_{t+k}) = EV_\tau(1_{\tau+k})$ for every t and k, then, it is immediate that $E[f_t(k)] = E[h_t(k)]$; with $h_t(k) \equiv \log(R_{t+1,k}/R_{t+1,1})$, the log excess holding return.

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

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

Concerning holding returns, for empirical implementation, we assume enough regularity so that

$$E_t \log \lim_{k\to\infty}(R_{t+1,k}/R_{t+1,1}) = \lim_{k\to\infty} E_t \log(R_{t+1,k}/R_{t+1,1}) \equiv h_t(\infty).$$

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

Proposition 5 *If the limits of $h_t(k)$, $f_t(k)$, and $y_t(k)$ exist, the unconditional expectations of holding returns are independent of calendar time; that*

is,

$$E\left(\log R_{t+1,k}\right) = E\left(\log R_{\tau+1,k}\right) \quad \textit{for all } t, \tau, k,$$

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

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

In practice, these three measures may not be equally convenient 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. To the extent that a term spread measure converge more rapidly to the asymptotic value, it will be preferred. In the cases considered here, 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 differ-

entials, y, might be slightly less informative for k finite than the term spreads estimated from forward rates and holding returns.

4 Empirical Evidence

The main objective of this section is to estimate a lower bound for the volatility of the permanent component of pricing kernels, as well as the related upper bound for the transitory component. We address these two points in subsection 4.1 and 4.2 below. We also present two sets of additional results that help interpret these estimates. First, we consider a simple process for the pricing kernel that corresponds to the specification implied by many studies of dynamic general equilibrium models. We show how our main findings can provide guidance for the degree of persistence that such models should reasonably display. Second, we measure the part of the permanent component that is due to inflation. As is well known, price levels are typically nonstationary. We document to what extent our findings provide information about the permanent components of real variables over and above the permanent components in price levels.

4.1 The volatility of the permanent component

Tables 1, 2, and 3 present estimates of the lower bound to the volatility of the permanent component of pricing kernels derived in Proposition 2. Specifically, we report estimates of

$$\frac{E\left(\log \frac{R_{t+1}}{R_{t+1,1}}\right) - E\left(\log \frac{R_{t+1,\infty}}{R_{t+1,1}}\right)}{E\left(\log \frac{R_{t+1}}{R_{t+1,1}}\right) + L\left(1/R_{t+1,1}\right)} \tag{7}$$

obtained by replacing each expected value with its sample analog for different data sets.

In Table 1, we report estimates of the lower bound given in equation (7) and of each of the three quantities entering into it, as well as the asymptotic normal probability that the numerator is negative. We present estimates using zero-coupon bonds for maturities 25 and 29 years, for various measures of the term spread (based on yields, forward rates and holding returns), and for holding periods of one year and one month. As return R_{t+1} we use the CRSP value-weighted index covering the NYSE, Amex and NASDAQ. The data is monthly, from 1946:12 to 1999:12. 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.[8]

Based on the asymptotic (normal) distribution, the probability that the term spread is larger than the log equity premium is very small, in most cases well below 1%. Hence, the hypothesis that the pricing kernel has no permanent innovation is clearly rejected. Not only is there a permanent component, it is very volatile. We find that the lower bound of the volatility of the permanent component is about 100%; none of our estimates are below 75%. The estimates are precise, standard errors are below 10%, except for holding returns.

Two points about the result in Table 1 are noteworthy. First, the choice of the holding period, and hence the level of the risk-free rate, has some effects on our estimates. For instance, using yields with a yearly holding period the size of the permanent component is estimated to be about 87%. Instead, using yields and a monthly holding period we estimate it to be 77%. This difference is due to the fact that monthly yields are about 1% below annual yields, affecting the estimate of the denominator of the lower bound.[9] Second, by estimating the right-hand side of equation (7) as the ratio of sample means, our estimates are 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 for forward rates and yields, on the order of about 1% in absolute value terms. Estimates of the bias are somewhat larger for holding returns.

Since (7) holds for any return R_{t+1}, we select portfolios with high $E\left(\log \frac{R_{t+1}}{R_{t+1,1}}\right)$ in Table 2 to sharpen the bounds based on the equity premium in Table 1. Table 2 contains the same information as Table 1, except that Table 2 covers only bonds with 25 years of maturity. We find estimates of $E\left(\log \frac{R_{t+1}}{R_{t+1,1}}\right)$ of up to 22.5% compared to 7.6% in Table 1. The smallest estimate of the lower bound in Table 2 is 89% as opposed to 77% in Table 1.

In panel A we let R_{t+1} be a fixed weight portfolio of aggregate equity and the risk-free rate that maximizes $E\left(\log \frac{R_{t+1}}{R_{t+1,1}}\right)$, that is, we are deriving the so-called "growth optimal" portfolio (see Bansal and Lehmann, 1997). Depending on the choice of the holding period, $E\left(\log \frac{R_{t+1}}{R_{t+1,1}}\right)$ is up to 9% larger than the premium presented in Table 1, with a share of equity of 2.14 or 3.46. In panel B of Table 2, we choose a fixed-weight portfolio from the menu of the 10 CRSP size decile portfolios. This leads to an average log excess return of up to 22.5%.

Table 3 extends the sample period to over 100 years and adds an addi-

tional country, the U.K. For the U.S., given data availability, we use coupon bonds with about 20 years of maturity. For the U.K., we use consols. For the U.S., 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 U.K. are similar to those for the U.S.

A natural concern is whether 25- or 29-year bonds allow for good approximations of the limiting term spread. From Figure 1, which plots term structures for three definitions of term spreads, we take that the long end of the term structure is either flat or decreasing. Extrapolating from these pictures, 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 returns. The other is that for longer maturities, as discussed before, our data set is smaller.

Note that for the bound in Equation (7) to be well defined, specifically for $L\left(1/R_{t+1,1}\right)$ to be finite, we have assumed that interest rates are stationary.[10] While the assumption of stationary interest rates is confirmed by many stud-

ies (for instance, Ait-Sahalia (1996)), others report the inability to reject unit roots (for instance, Hall, Anderson, and Granger (1992)). Cochrane (2005, page 199) sums up the issue eloquently: "the level of nominal interest rates is surely a stationary variable in a fundamental sense: we have observations near 6% as far back as ancient Babylon, and it is about 6% again today."Also, consistent with the idea that interest rates are stationary and therefore $L\left(1/R_{t+1,1}\right)$ finite, Table 3 shows lower estimates for the very long samples than for the postwar period.

4.2 The volatility of the transitory component

We now report on estimates for volatility of the transitory component and the related upper bound for the volatility of the transitory component relative to the volatility of the stochastic discount factor. As shown in Figure 2, $L\left(1/R_{\infty}\right)$ goes up to 0.04 for 29 year maturity, while being about 0.015 for 20 years of maturity. The corresponding upper bound for the volatility relative to the overall volatility $L\left(1/R_{\infty}\right)/L\left(M'/M\right)$ reaches a maximum of 23% at the 29 year maturity, while being about 9% for 20 year maturity. This upper bound is based on the CRSP decile portfolios as reported in Table 2. Unfortunately, these estimates are somewhat difficult to interpret because

there is no apparent convergence for the available maturities. Moreover, the lack of a complete data set for all maturities seems to result in a substantial upward bias of the estimates of $L\left(1/R_k\right)$ for maturities $k \geq 20$ years. Figure 3 shows that the data for the longest maturities is concentrated in the part of the sample characterized by high volatility. A simple way to adjust for this sample bias would be to assume that the ratio of the volatilities for different maturities is constant across the entire sample. We can then consider the volatility for the 13 year bond, the longest for which we have a complete sample, as a benchmark. The ratio of the volatilities of the 13 year bond for the entire sample over that for the sample covered by the longest available maturity, 29 years, is about 0.8 so that the relative upper bound would be adjusted to about 18%, down from 23%.

Concerning the measurement of the permanent component, note that the average term spread for the 13 year bond is actually larger for the shorter sample covered by the 29 year bond, although by only 20 basis points. Thus, any adjustment would, if anything, further increase the estimates of the volatility of the permanent component in equation 7.

4.3 An example of a pricing kernel

We present here an example that illustrates the power of bond data to distinguish between similar levels of persistence. In particular, the example shows 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. The example is relevant, because many studies of dynamic general equilibrium models imply stationary pricing kernels.

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

$$h_t(k) = \frac{\sigma_\varepsilon^2}{2}\left(1 - \rho^{2(k-1)}\right). \tag{8}$$

This expression suggests that if the volatility of the innovation of the pricing kernel, σ_ε^2, is large, then values of ρ only slightly 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. As is clear from Table 4, ρ has to be extremely close to 1.

For this calculation we have set $\sigma_\varepsilon^2 = 0.4$, for the following reasons. Based on Proposition 2 and assuming lognormality, we get

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

where R_{t+1} can be any risky return. Based on our estimates in Table 2, the growth optimal excess return should be at least 20%, so that $var\left(\log \frac{M_{t+1}}{M_t}\right) \geq 0.4$. Finally, for ρ close to 1 we can write

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

4.4 Nominal versus real pricing kernels

Because we have so far used bond data for 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 volatility 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 $M_t = \left(\frac{1}{P_t}\right) \widetilde{M}_t^T$, where P_t is the aggregate price level. Thus $1/P_t$ converts nominal payouts into real payouts and $\widetilde{M}_t^T$ prices real payouts. Because, $1/P_t$ is directly observable, we can measure the volatility of its permanent component directly and then compare it to the estimated volatility of the permanent component of pricing kernels reported in Tables 1, 2, and 3. It turns out that the volatility of the permanent component in $1/P_t$ is estimated at up to 100 times smaller than the lower bound of the volatility of the permanent component in pricing kernels estimated above. 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. It should be noted that this interpretation is only valid to the extent that the behavior of the official consumer price index accurately reflects the properties of the price level faced by asset market participants.

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

Proposition 6 *Assume that the process X_t satisfies assumptions (1) and (2) and that the following regularity conditions are satisfied: (a) $\frac{X_{t+1}}{X_t}$ is strictly*

stationary, and (b) $\lim_{k\to\infty} \frac{1}{k} L\left(\frac{E_t X_{t+k}}{X_t}\right) = 0$. *Then*

$$L\left(\frac{X_{t+1}^P}{X_t^P}\right) = \lim_{k\to\infty} \frac{1}{k} L\left(\frac{X_{t+k}}{X_t}\right). \tag{9}$$

The usefulness of this proposition is that $L\left(X_{t+1}^P / X_t^P\right)$ is a natural measure for the volatility 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. The quantity $\lim_{k\to\infty} \frac{1}{k} L\left(X_{t+k}/X_t\right)$ can be estimated with knowledge of only X_t. This result is analogous to a result in Cochrane (1988), with a main 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 $L(.)$ estimates, which would suggest that departures from lognormality are small. Overall, we estimate the volatility 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) volatility of

the permanent component of the pricing kernel,

$$L\left(\frac{M_{t+1}^{P}}{M_{t}^{P}}\right) \geq E\left[\log R_{t+1} - \log R_{t+1,\infty}\right], \tag{10}$$

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. The first two rows display results based on estimating an AR1 or AR2 for inflation and then computing the volatility 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 $\lim_{k\to\infty}(1/k)\,var\left(\log X_{t+k}/X_{t}\right)$. For the postwar period, the volatility of the permanent component is 0.43% or 0.30%, depending on whether $k = 20$ or 30.[11] The table also shows that $L\left(X_{t+k}/X_{t}\right)/var\left(\log X_{t+k}/X_{t}\right)$ is approximately 0.5. Note that the roots of the process for inflation reported in Table 5 are far from one, supporting our implicit assumption that inflation rates are stationary.

A second view about the volatility of the permanent component can be obtained from inflation-indexed bonds. Such bonds have been traded in the U.K. 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. 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 $L\left(1/R_{t+1,1}\right)$ in the definition of the volatility of the permanent component as given in equation (5). Because of this, we focus on the bound for the *absolute* volatility of the pricing kernel as given in equation (10). For the nominal kernel, we use average nominal equity returns for $E \log R_{t+1}$, and for $E \log R_{t+1,\infty}$, 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 (10). 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. The right-hand side of (10) differs for nominal and real pricing kernels only if there is an inflation risk premium for long-term nominal bonds. If long-term nominal bonds have a positive inflation risk premium then the

lower bound for the permanent component for 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 volatility of the permanent component of inflation is very small, the differences in volatility of the permanent components for nominal and real kernels are very small. Comparing columns (3) and (6), for one point estimate the volatility of the permanent component of real kernels is larger than the estimate for the corresponding nominal kernels; for the second case, they are basically identical. In any case, the corresponding standard errors are larger than the differences between the results for nominal and real kernels.

5 Pricing Kernels and Aggregate 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 of 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 an example of a utility function in which the resulting pricing kernels have permanent innovations because of the persistence introduced through the utility function. Finally, we estimate the volatility of the permanent component in consumption directly and compare it to our estimates of the volatility of the permanent component of pricing kernels.

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

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 7 *Assume that there is a unique invariant measure, λ^*. In addition, if either (i) $\lim_{k\to\infty}\left(T^k f\right)(s) = \int f d\lambda^* > 0$ and finite, 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*

$$\lim_{k\to\infty} \frac{E_{t+1}(M_{t+k})}{E_t(M_{t+k})} = 1.$$

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, assume (b) that

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

for all $h(.)$ bounded and continuous.

Proposition 8 *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.*

An example covered by this proposition is CRRA utility, $\frac{1}{1-\gamma} c_t^{1-\gamma}$ with relative risk aversion γ, where $f(c_t) = c_t^{-\gamma}$, with $\overline{c} \geq c_t \geq \varepsilon > 0$. If consumption would have a unit root, then properties (a) and (b) would not be satisfied.

For the CRRA case, even with consumption satisfying properties (a) and (b), Proposition (8) could fail to be satisfied because $c_t^{-\gamma}$ is unbounded if c_t gets arbitrarily close to zero with large enough probability. It is possible to construct examples where this is the case, for instance, along the lines of the model in Aiyagari (1994). This outcome is driven by the Inada condition $u'(0) = \infty$. Note also, the bound 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 (6) defining the property of no permanent innovations is satisfied.

5.1 Examples with additional state variables

There are many examples in the literature for which marginal utility is a function of additional state variables, and for which it is straightforward to apply Proposition 8, very much like for the CRRA utility shown above. For instance, the utility functions displaying various forms of habits such as those used by Ferson and Constantinides (1991), Abel (1999) and Campbell and Cochrane (1999). On the other hand, there are cases where Proposition 8 does not apply. For instance, as we show below, for the Epstein-Zin-Weil utility function. In this case, even with consumption satisfying the conditions required for Proposition 8, the additional state variable does not have an invariant distributions. Thus, innovations to pricing kernels have always permanent effects.

Assume the representative agent has preferences represented by nonexpected utility of the following recursive form:

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

where U_t is the utility starting at time t and ϕ is an increasing concave function. 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

$$\frac{M_{t+1}}{M_t} = \left[\beta \left(\frac{C_{t+1}}{C_t}\right)^{-\rho}\right]^{\theta} \left[\frac{1}{R^c_{t+1}}\right]^{(1-\theta)} \tag{11}$$

with $\theta = (1-\gamma)/(1-\rho)$ 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 (11) reveals that a pricing kernel M_{t+1} for this model is

$$M_{t+1} = \beta^{\theta(t+1)}\, Y_{t+1}^{\theta-1}\, C_{t+1}^{-\rho\theta}, \text{ where } Y_{t+1} = R^c_{t+1} \cdot Y_t \tag{12}$$

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

$$C_t = \tau^t c_t, \tag{13}$$

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 $R^c_{t+1} = \left(V^c_{t+1} + C_{t+1}\right)/V^c_t$. We assume that agents discount the

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

$$\max_{c\in[\underline{c},\bar{c}]} \beta\tau^{1-\rho}\left\{\int\left(\frac{c'}{c}\right)^{1-\gamma} dF\left(c'\right)\right\}^{1/\theta} < 1. \tag{14}$$

Proposition 9 *Let the pricing kernel be given by (12), let the detrended consumption be iid as in (13), and assume that (14) holds. Then the price-dividend ratio for the consumption equity is given by* $V_t^c/C_t = \psi c_t^{\gamma-1}$ *for some constant* $\psi > 0$*; hence,* V_t^c/C_t *is iid. Moreover,*

$$x_{t+1,k} \equiv \frac{E_{t+1}\left(M_{t+k}\right)}{E_t\left(M_{t+k}\right)} = \frac{\left(1+\frac{1}{\psi}c_{t+1}^{(1-\gamma)}\right)^{\theta-1}}{E_t\left\{\left(1+\frac{1}{\psi}c_{t+1}^{(1-\gamma)}\right)^{\theta-1}\right\}}; \tag{15}$$

thus the pricing kernel has permanent innovations 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 (15) also makes clear that for values of θ close to one, the volatility of the permanent component is small.

5.2 The volatility of the permanent component in consumption

We present here estimates of the volatility of the permanent component of consumption, obtained directly from consumption data. We end up drawing two conclusions. One is that the volatility 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 volatility 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 importance of the permanent component in consumption.[12] The other conclusion, as noted in Cochrane (1988) for the random walk component in GDP, is that standard errors for these direct estimates are large.

As in subsection 4.4 for inflation, we use Cochrane's method based on the variance, since $L\left(X_{t+k}/X_t\right)\ /\ var\left(\log X_{t+k}/X_t\right)$ is close to 0.5. Specifically, for k up to 35, it lies between 0.47 and 0.49. Our estimates for $(1/k)\, var\left(\log X_{t+k}/X_t\right)\ /\ var\left(\log X_{t+1}/X_t\right)$, with associated standard error bands, are presented in Figures 4 and 5 for the periods 1889–1997 and 1946–97, respectively. For the period 1889–1997, shown in Figure 4, the estimates

stabilize at around 0.5 and 0.6 for k larger than 15. For the postwar period, shown in Figure 5, standard error bands are too wide to draw firm conclusions.

6 Conclusions

The main contribution of this paper is to derive and estimate a lower bound for the volatility of the permanent component of asset pricing kernels. We find that the permanent component is about at least as volatile as the stochastic discount factor itself. This result is driven by the historically low yields on long-term bonds. These yields 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. A related bound that measures the volatility of the transitory component suggests it to be considerably less important than the permanent component. We also relate the persistence of pricing kernels to the persistence of their determinants in standard models, 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 de-

termined, to a large extent, by real as opposed to nominal factors. Finally, we present some evidence that the importance 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 permanent component.

Dept. of Economics, University of Chicago, Chicago, IL, 60637, U.S.A.; and NBER; f-alvarez1@uchicago.edu

and

Dept. of Finance, The Wharton School of the University of Pennsylvania, 3620 Locust Walk, Philadelphia, PA 19104-6367, U.S.A.; and NBER; jermann@wharton.upenn.edu; http://finance.wharton.upenn.edu/~jermann.

A Proofs

Proposition 2. We show that, i) $R_{t,t+1,\infty} = M_t^T/M_{t+1}^T$, and ii)

$$L_t\left(\frac{M_{t+1}}{M_t}\right) = L_t\left(\frac{M_{t+1}^P}{M_t^P}\right) + E_t \log \frac{R_{t+1,\infty}}{R_{t+1,1}},$$

and then that this implies

$$L_t\left(\frac{M_{t+1}^P}{M_t^P}\right) \geq E_t \log \frac{R_{t+1}}{R_{t+1,1}} - E_t \log \frac{R_{t+1,\infty}}{R_{t+1,1}}.$$

i) Using assumption (1)

$$\begin{aligned} R_{t,t+1,\infty} &\equiv \lim_{k\to\infty} \frac{V_{t+1}\left(1_{t+k}\right)}{V_t\left(1_{t+k}\right)} = \lim_{k\to\infty} \frac{\frac{E_{t+1}M_{t+k}}{M_{t+1}}}{\frac{E_t M_{t+k}}{M_t}} = \lim_{k\to\infty} \frac{\frac{E_{t+1}M_{t+k}/\beta^{t+k}}{M_{t+1}}}{\frac{E_t M_{t+k}/\beta^{t+k}}{M_t}} \\ &= \frac{\lim_{k\to\infty} \frac{E_{t+1}M_{t+k}/\beta^{t+k}}{M_{t+1}}}{\lim_{k\to\infty} \frac{E_t M_{t+k}/\beta^{t+k}}{M_t}} = \frac{\frac{M_{t+1}^P}{M_{t+1}}}{\frac{M_t^P}{M_t}} = \frac{M_t^T}{M_{t+1}^T}. \end{aligned}$$

ii) By definition,

$$\begin{aligned} L_t\left(\frac{M_{t+1}}{M_t}\right) &= \log E_t \frac{M_{t+1}}{M_t} - E_t \log \frac{M_{t+1}^P M_{t+1}^T}{M_t^P M_t^T} \\ &= -\log \frac{1}{V_t\left(1_{t+1}\right)} - E_t \log \frac{M_{t+1}^T}{M_t^T} + L_t\left(\frac{M_{t+1}^P}{M_t^P}\right) \\ &= E_t \log \frac{R_{t+1,\infty}}{R_{t+1,1}} + L_t\left(\frac{M_{t+1}^P}{M_t^P}\right). \end{aligned}$$

Hence

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

because from no-arbitrage and concavity of the log

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

For an unconditional version of the bound we use that $L\left(x_{t+1}\right) = EL_t\left(x_{t+1}\right) + L\left(E_t x_{t+1}\right)$. Using this result, taking unconditional expectations

$$\begin{aligned}EL_t\left(\frac{M^P_{t+1}}{M^P_t}\right) &= EL_t\left(\frac{M_{t+1}}{M_t}\right) - EE_t \log\frac{R_{t+1,\infty}}{R_{t+1,1}}\\ L\left(\frac{M^P_{t+1}}{M^P_t}\right) &= L\left(\frac{M_{t+1}}{M_t}\right) - L\left(E_t\frac{M_{t+1}}{M_t}\right) - E\log\frac{R_{t+1,\infty}}{R_{t+1,1}}\\ &= L\left(\frac{M_{t+1}}{M_t}\right) - L\left(1/R_{t+1,1}\right) - E\log\frac{R_{t+1,\infty}}{R_{t+1,1}}\end{aligned}$$

$$L\left(\frac{M_{t+1}}{M_t}\right) \geq E\log\frac{R_{t+1}}{R_{t+1,1}} + L\left(1/R_{t+1,1}\right)$$

and forming the ratio

$$L\left(\frac{M^P_{t+1}}{M^P_t}\right)/L\left(\frac{M_{t+1}}{M_t}\right) = \frac{L\left(\frac{M_{t+1}}{M_t}\right) - L\left(1/R_{t+1,1}\right) - E\log\frac{R_{t+1,\infty}}{R_{t+1,1}}}{L\left(\frac{M_{t+1}}{M_t}\right)}.$$

So that: if $\left[-L\left(1/R_{t+1,1}\right) - E\log\frac{R_{t+1,\infty}}{R_{t+1,1}}\right] \leq 0$, and $E\log\frac{R_{t+1}}{R_{t+1,1}} + L\left(1/R_{t+1,1}\right) > 0$

$$1 \geq L\left(\frac{M^P_{t+1}}{M^P_t}\right)/L\left(\frac{M_{t+1}}{M_t}\right) \geq \frac{E\log\frac{R_{t+1}}{R_{t+1,1}} - E\log\frac{R_{t+1,\infty}}{R_{t+1,1}}}{E\log\frac{R_{t+1}}{R_{t+1,1}} + L\left(1/R_{t+1,1}\right)}$$

and if $\left[-L\left(1/R_{t+1,1}\right) - E\log\frac{R_{t+1,\infty}}{R_{t+1,1}}\right] > 0$, and $E\log\frac{R_{t+1}}{R_{t+1,1}} + L\left(1/R_{t+1,1}\right) > 0$

$$1 < L\left(\frac{M_{t+1}^P}{M_t^P}\right)/L\left(\frac{M_{t+1}}{M_t}\right) < \frac{E\log\frac{R_{t+1}}{R_{t+1,1}} - E\log\frac{R_{t+1,\infty}}{R_{t+1,1}}}{E\log\frac{R_{t+1}}{R_{t+1,1}} + L\left(1/R_{t+1,1}\right)}.$$

Q.E.D.

Proposition 4. Given the proof of Proposition 2, we only need to show that under the stated assumptions $R_{t,t+1,\infty} = M_t^T/M_{t+1}^T$. By definition

$$R_{t,t+1,\infty} \equiv \lim_{k\to\infty}\frac{V_{t+1}\left(1_{t+k}\right)}{V_t\left(1_{t+k}\right)} = \lim_{k\to\infty}\frac{E_{t+1}M_{t+k}}{E_tM_{t+k}}\frac{M_t}{M_{t+1}},$$

and by the definition of $v_{t,t+k}$, the first term equals

$$\frac{E_{t+1}\left[M_{t+k}\right]}{E_t\left[M_{t+k}\right]} = \frac{E_{t+1}\left[M_{t+k}^T\right]E_{t+1}\left[M_{t+k}^P\right]\left(1+v_{t+1,t+k}\right)}{E_t\left[M_{t+k}^T\right]E_t\left[M_{t+k}^P\right]\left(1+v_{t,t+k}\right)}.$$

Taking limits gives $\lim_{k\to\infty}\left\{E_{t+1}\left[M_{t+k}\right]/E_t\left[M_{t+k}\right]\right\} = M_{t+1}^P/M_t^P$, due to the assumption that

$$\lim_{k\to\infty}\left[\frac{\left(1+v_{t+1,t+k}\right)}{\left(1+v_{t,t+k}\right)}\right] = 1,$$

that M_{t+1}^P/M_t^P is a martingale, and due the definition of no permanent innovations. Thus $R_{t,t+1,\infty} = \left(M_{t+1}^P/M_t^P\right)\left(M_t/M_{t+1}\right) = M_t^T/M_{t+1}^T$. *Q.E.D.*

Proposition 5. By definition,

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

Taking unconditional expectations on both sides, we have that

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

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

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

Denote the limit

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

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

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

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

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

Q.E.D.

Proposition 6. Using assumption (a) that $\frac{M_{t+1}}{M_t}$ is strictly stationary, some algebra shows that

$$\frac{1}{k}L\left(\frac{M_{t+k}}{M_t}\right)=\frac{1}{k}\log E\left(\frac{M_{t+k}}{M_t}\right)+E\log\left(\frac{M_{t+1}}{M_t}\right).$$

Again using the stationarity assumption and some algebra we have

$$\begin{aligned} &\frac{1}{k}L\left(\frac{M_{t+k}}{M_t}\right) \\ =\ &\frac{1}{k}L\left(E_t\frac{M_{t+k}}{M_t}\right)-\left(\frac{1}{k}\right)\left[\sum_{j=1}^{k}E\log R_{t+1,j}\right]+EL_t\left(\frac{M_{t+1}}{M_t}\right)+E\log R_{t+1,1}. \end{aligned}$$

Going to the limit, which given assumptions (1) and (2) exists, we get

$$\lim_{k\to\infty}\frac{1}{k}L\left(\frac{M_{t+k}}{M_t}\right)=\lim_{k\to\infty}\frac{1}{k}L\left(E_t\frac{M_{t+k}}{M_t}\right)-E\frac{\log R_{t+1,\infty}}{\log R_{t+1,1}}+EL_t\left(\frac{M_{t+1}}{M_t}\right).$$

Finally, with assumption (b) we have the postulated result, given that from the proof of Proposition 2 it is easy to see that

$$EL\left(\frac{M_{t+1}^P}{M_t^P}\right)=EL_t\left(\frac{M_{t+1}}{M_t}\right)-E\log\frac{R_{t+1,\infty}}{R_{t+1,1}}.$$

Q.E.D.

Proposition 7. Using the Markov assumption under (i) and (ii) we have

$$\lim_{k\to\infty}\frac{E_{t+1}\left(M_{t+k}\right)}{E_t\left(M_{t+k}\right)}=\lim_{k\to\infty}\frac{\left(T^{k-1}f\right)\left(s'\right)}{\left(T^kf\right)\left(s\right)}=1$$

Q.E.D.

Proposition 8. Properties (a) and (b) define setwise convergence, and with $f\left(.\right)$ bounded, expected values converge. *Q.E.D.*

Proposition 9. 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.1. 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, $V_t^c/C_t=f^*\left(c_t\right)$, where the function $f^*$is the unique solution to

$$T^*f^*=f^*,\ f^*\left(c\right)=\psi\,c^{\gamma-1}$$

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

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

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

Proof. Using the pricing kernel (12), we obtain that consumption equity must satisfy

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

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

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

and dividing by c_t on both sides, we can write

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

where f is the price-dividend ratio of the consumption equity: $f(c) = v(c)/c$. The operator T can be shown to be a contraction: hence, it has a unique fixed point. Moreover, ψ is given by

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

where f^* satisfies $Tf^* = f^*$. *Q.E.D.*

Using Lemma A.1, we can write the return on the consumption equity as

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

Then using (12), (15), and through some algebra, we get

$$\begin{aligned} x_{t+1,k} &= \frac{E_{t+1}M_{t+k}}{E_t M_{t+k}} = \frac{E_{t+1}\left[\beta^{\theta(t+1)}\, C_{t+1}^{-\rho\theta}\, Y_{t+1}^{\theta-1}\right]}{E_t\left[\beta^{\theta(t+1)}\, C_{t+1}^{-\rho\theta}\, Y_{t+1}^{\theta-1}\right]} \\ &= \left(1+\frac{1}{\psi}c_{t+1}^{\gamma-1}\right)^{\theta-1} / E_t\left[\left(1+\frac{1}{\psi}c_{t+1}^{\gamma-1}\right)^{\theta-1}\right]. \end{aligned}$$

Q.E.D.

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 (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 (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 month to 30 years, although for longer maturities, yields are not available for all years.

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 Database; http://www.globalfindata.com. 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 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.

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

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

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 population 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 $[0.0071, -0.0012]$ for the two 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 $[0.0591, 0.1277]$, $[-0.0077, -0.0112]$, and $[-0.0165, -0.0209]$.

References

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

Aiyagari, R. S. (1994): "Uninsured Idiosyncratic Risk and Aggregate Saving," *Quarterly Journal of Economics*, 109, 659-684.

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

— (2005), "Using Asset Prices to Measure the Persistence of the Marginal Utility of Wealth," *Econometrica Supplementary Material*.

Anderson, N., and J. Sleath (1999): "New Estimates of the UK Real and Nominal Yield Curves," *Quarterly Bulletin,* Bank of England, November.

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

Bansal, R., and A. Yaron (2004): "Risks for the long run: A potential solution of asset pricing puzzles," *Journal of Finance*, 59, 1481-1509.

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

Blanchard, O., and D. Quah Danny (1989): "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review,* 79, 655-673.

Beveridge, S., and C. R. Nelson (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,* 7 151–74.

Bliss, R. (1997): "Testing Term Structure Estimation Methods," in *Advances in Futures and Options Research,* Vol. 9, P. Boyle, G. Pennachi, and P. Ritchken, eds, Greenwich, Conn.: JAI Press, 197–231.

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

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

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

— (2005): *Asset Pricing*, Princeton University Press.

Cochrane, J. H., and Hansen L.P. (1992): "Asset Pricing Exploration for Macroeconomics," *NBER Macroeconomic Annual*, 152-65.

Christiano, L., M. Eichenbaum and R. Vigfusson (2002): "What Happens after a Technology Shock?" manuscript, Northwestern University, Evanston IL.

Daniel, K. D., and D. A. Marshall (2001): "Consumption and Asset Returns at Short- and Long-Horizons," Unpublished Manuscript, Northwestern University, Evanston IL.

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

Duffie, D. (1996): *Dynamic Asset Pricing Theory*, Princeton, N.J.: Princeton University Press.

Epstein, L. G., and S. E. Zin (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A

Theoretical Framework," *Econometrica,* 57, 937–69.

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

Fisher, J. D. M. (2002): "Technology Shocks Matter," Federal Reserve Bank of Chicago, WP 2002-14.

Gali, J. (1999): "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?," *American Economic Review,89*, 249-71.

Hall, A. D., H. M. Anderson, C. W. J. Granger (1992): "A Cointegration Analysis of Treasury Bill Yields," *Review of Economics and Statistics,* 74, 116–26.

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

Hansen, L. P., J. C. Heaton, and N. Li (2004): "Consumption strikes back," Unpublished Manuscript, University of Chicago.

Hansen, L. P., and R. Jagannathan (1991): "Implications of Security

Market Data for Models of Dynamic Economies," *Journal of Political Economy,* 99, 225–62.

Hansen, L. P., and J. Scheinkman (2003): "Semi Group Pricing," Unpublished manuscript University of Chicago and Princeton University.

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

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

Lettau, M. and S. Ludvigson (2004): "Understanding Trend and Cycle in Asset Values: Reevaluating the Wealth Effect on Consumption," *American Economic Review,* 94, 276-299.

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

— (2003): "Two decompositions of the local variance of state prices," Unpublished Manuscript, Unversity of Minnesota.

McCulloch, J. H., and H.-C. Kwon (1993): "U.S. Term Structure Data, 1947–1991," Working Paper 93-6, Ohio State University.

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

Newey W. K., and K. D. West (1987): "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica,* 55, 703–8.

Shapiro M., and M. Watson (1988): "Sources of Business Cycle Fluctuations," NBER Macroeconomics Annual, 111-48.

Shiller, R. (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.

Theil H. (1967): *Economics and Information Theory*, Amsterdam: North Holland Publishing Company.

Notes

[1]We thank Andy Atkeson, Erzo Luttmer, Lars Hansen, Pat Kehoe, Bob King, Narayana Kocherlakota, Stephen Leroy, Lee Ohanian, and the participants in workshops and conferences at UCLA, the University of Chicago, the Federal Reserve Banks of Minneapolis, Chicago, and Cleveland, and Duke, Boston, Ohio State, Georgetown and Yale Universities, NYU, Wharton, SED meeting in Stockholm, SITE, Minnesota workshop in macroeconomic theory and ESSFM for their comments and suggestions. We thank Robert Bliss for providing the data for U.S. zero-coupon bonds. Alvarez thanks the NSF and the Sloan Foundation for support. Earlier versions of this paper circulated as "The size of the permanent component of asset pricing kernels."

[2]A few prominent examples of research in this line are Hansen and Jagannathan (1991), Cochrane and Hansen (1992), Luttmer (1996).

[3]Asset prices have also been included in multivariate analyses of persistence of GDP and consumption, see for instance, Lettau and Ludvigson (2004).

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

[5]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 (1).

[5]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 (1).

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

[7]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.

[8]For maturities longer than 13 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 (7) are based on different 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.

[9]Our data set does not contain the information necessary to present results for monthly holding periods for forwards rates and holding returns.

[10]Equation (4), which defines a bound for the size of the permanent component in absolute terms, does not require this assumption.

[11]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 $\left(\frac{4}{3}\frac{k}{T}\right)^{0.5}\widehat{\sigma}_k^2$.

[12]This conclusion would not be valid if asset market participation is limited, unless the participants' consumption exhibits the same persistence properties as the aggregate.

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) Adjustment for volatility of short rate $L(1/R1)$	(4) Size of Permanent Component $L(P)/L$	(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
	(0.0169)	(0.0049)	(0.0002)	(0.0700)	(0.0193)	
29 years		-0.0040		**1.0520**	0.0704	0.0030
		(0.0070)		(0.1041)	(0.0256)	
B. Holding Returns		E[h(k)]	Holding Period is 1 Year			
25 years	0.0664	-0.0083	0.0005	**1.1164**	0.0747	0.0145
	(0.0169)	(0.0340)	(0.0002)	(0.5186)	(0.0342)	
29 years		-0.0199		**1.2899**	0.0863	0.0266
		(0.0469)		(0.7417)	(0.0446)	
C. Yields		E[y(k)]	Holding Period is 1 Year			
25 years	0.0664	0.0082	0.0005	**0.8701**	0.0582	0.0015
	(0.0169)	(0.0033)	(0.0002)	(0.0534)	(0.0196)	
29 years		0.0082		**0.8706**	0.0582	0.0050
		(0.0035)		(0.0602)	(0.0226)	
D. Yields		E[y(k)]	Holding Period is 1 Month			
25 years	0.0763	0.0174	0.0004	**0.7673**	0.0588	0.0028
	(0.0180)	(0.0031)	(0.0002)	(0.0717)	(0.0213)	
29 years		0.0168		**0.7755**	0.0595	0.0067
		(0.0033)		(0.0795)	(0.0241)	

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) Adjustment for volatility of short rate $L(1/R1)$	(4) Size of Permanent Component $L(P)/L$	(5) (1)-(2) $E[\log(R/R_1)]$ $-E[\log(R_k/R_1)]$	(6) P[(5) < 0]
A. Growth-Optimal Leveraged Market Portfolio, (Portfolio weight: 3.46 for monthly holding period, 2.14 for yearly)						
One-year holding period						
Forward rates	0.1095	-0.0004	0.0005	**0.9998**	0.11	0.0093
	(0.0402)	(0.0049)	(0.0002)	(0.0426)	(0.0467)	
Holding return		-0.0083		**1.0708**	0.1178	0.0092
		(0.0340)		(0.3203)	(0.050)	
Yields		0.0082		**0.9210**	0.1013	0.0159
		(0.0033)		(0.0381)	(0.0472)	
One-month holding period						
Yields	0.1689	0.0174	0.0004	**0.8946**	0.1515	0.0317
	(0.0686)	(0.0031)	(0.0002)	(0.0519)	(0.0816)	
B. 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	0.0005
	(0.0437)	(0.0049)	(0.0002)	(0.0276)	(0.0519)	
Holding return		-0.0083		**1.0459**	0.1775	0.0004
		(0.0340)		(0.2053)	(0.0628)	
Yields		0.0082		**0.9488**	0.161	0.0008
		(0.0033)		(0.0199)	(0.0512)	
One-month holding period						
Yields	0.2251	0.0174	0.0004	**0.9209**	0.2076	0.0089
	(0.0737)	(0.0031)	(0.0002)	(0.0320)	(0.0872)	

TABLE 3

Size of Permanent Component Based on Aggregate Equity and Coupon Bonds

		(1) Equity Premium $E[\log R/R_1]$	(2) Term Premium $E[y]$	$E[h]$	(3) Adjustment $L(1/R_1)$	(4) Size of Permanent Component $L(P)/L$	(5) (1)-(2)	P[(5) < 0]
US	1871-1997	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)	
	1946-97	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) Equity Premium $E[\log R/R_1]$	(2) Term Premium $E[y]$	$E[h]$	(3) Adjustment $L(1/R_1)$	(4) Size of Permanent Component $L(P)/L$	(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)	
	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)
L(P_t/P_{t+k}) / var(log P_{t+k}/P_t)		(k=20)	0.50			
		(k=30)	0.51			

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)
L(P_t/P_{t+k}) / var(log P_{t+k}/P_t)		(k=20)	0.51			
		(k=30)	0.49			

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)]	L(P)	E[log(π)]	E[log(F)]	E[log(Y)]	L(P)
25	0.1706	0.0762		**0.0944**	0.0422	0.0342		**0.0943**
	(0.0197)	(0.0040)		(0.0212)	(0.0063)	(0.0023)		(0.0230)
			0.0815	**0.089**			0.0347	**0.0937**
			(0.0046)	(0.0200)			(0.0018)	(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. See Appendix B for more details.

FIGURE 1

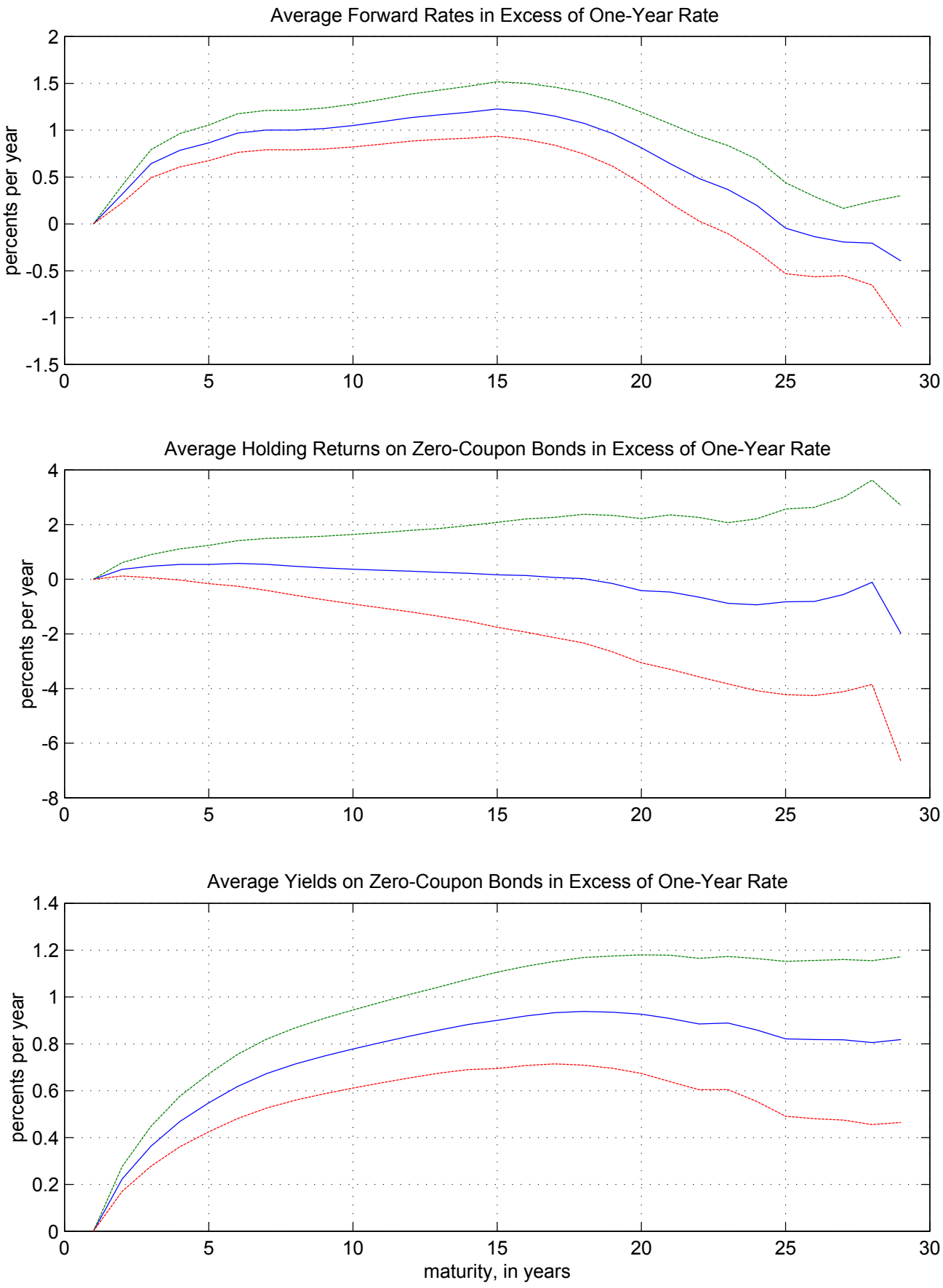
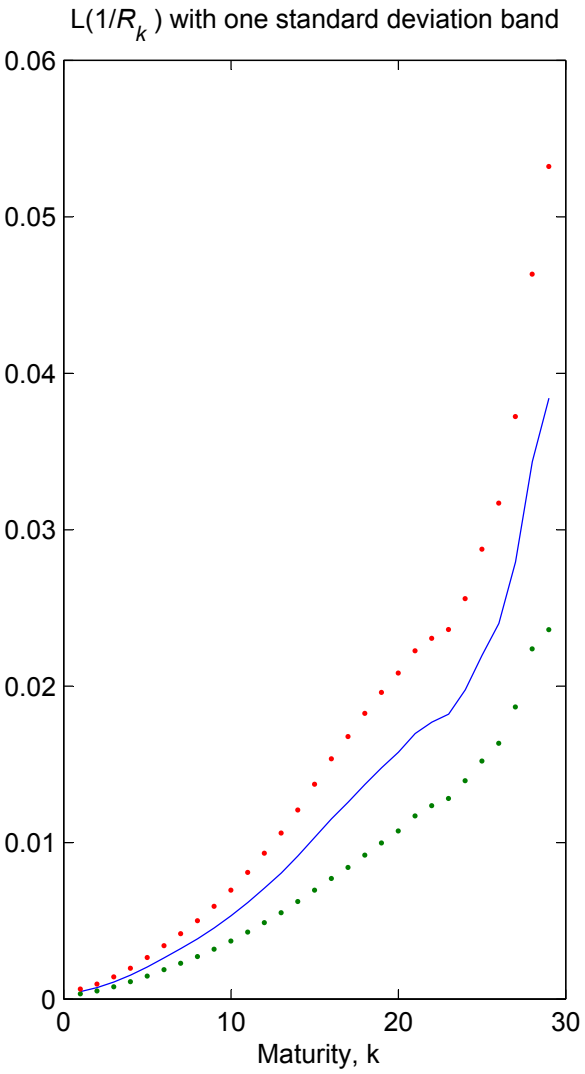

Average Forward Rates in Excess of One-Year Rate
percents per year
2
1.5
1
0.5
0
-0.5
-1
-1.5
0
5
10
15
20
25
30
Average Holding Returns on Zero-Coupon Bonds in Excess of One-Year Rate
percents per year
4
2
0
-2
-4
-6
-8
0
5
10
15
20
25
30
Average Yields on Zero-Coupon Bonds in Excess of One-Year Rate
percents per year
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
5
10
15
20
25
30
maturity, in years

FIGURE 2


L(1/R_k) with one standard deviation band
0.06
0.05
0.04
0.03
0.02
0.01
0
0
10
20
30
Maturity, k

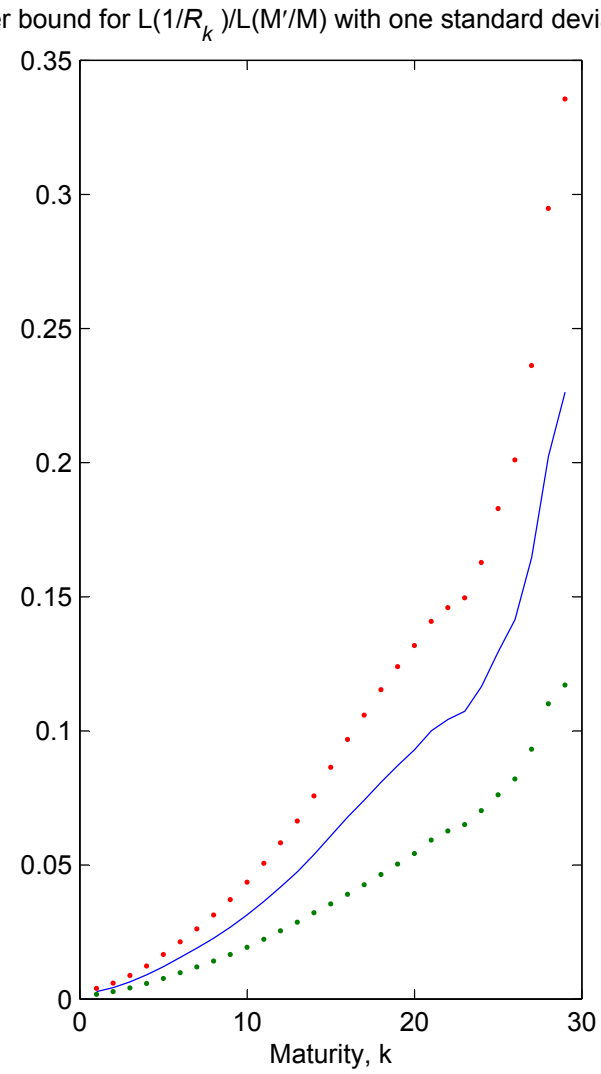

Upper bound for L(1/R_k)/L(M′/M) with one standard deviation band
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
10
20
30
Maturity, k

FIGURE 3

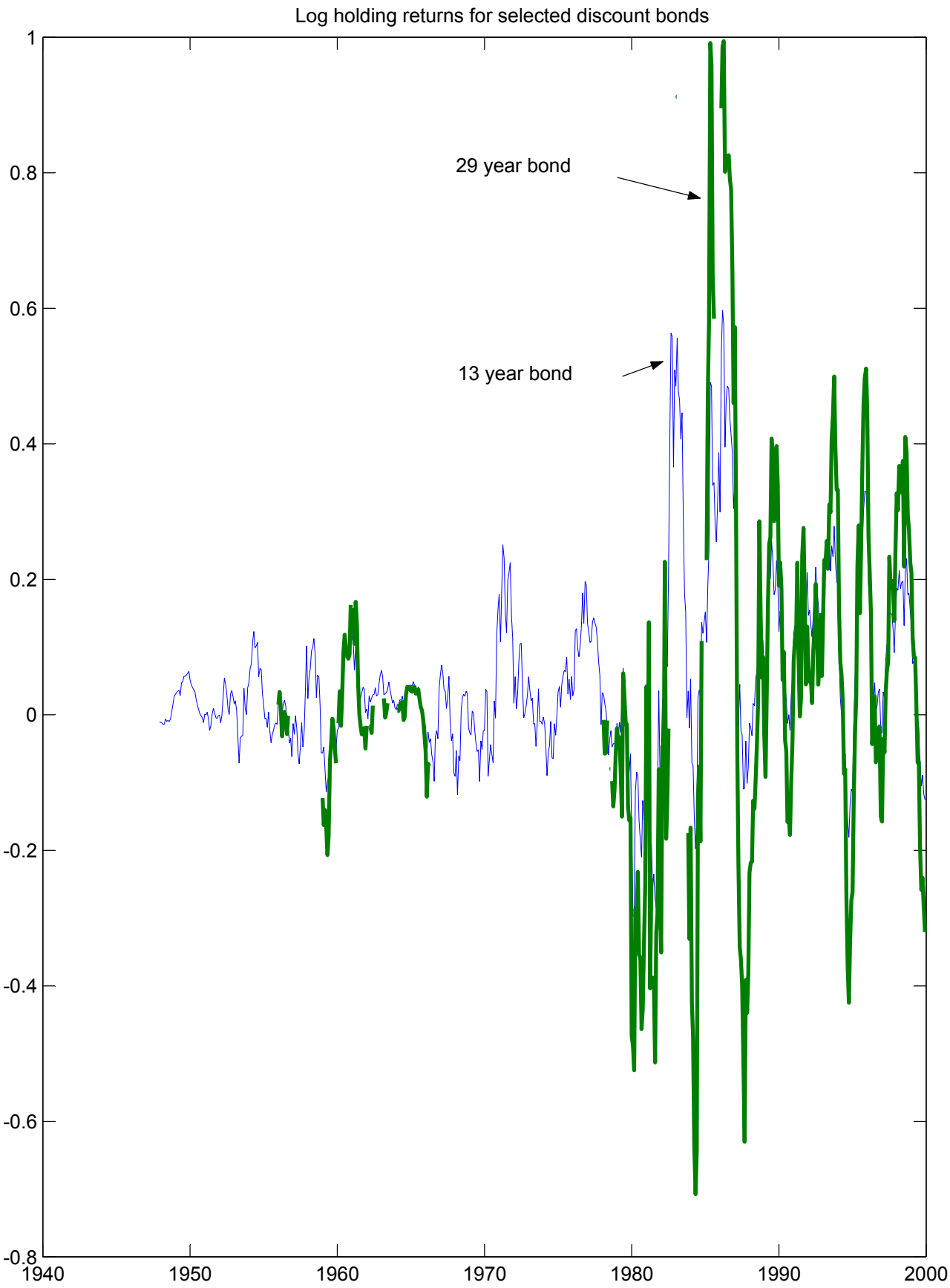

Log holding returns for selected discount bonds
29 year bond
13 year bond
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
1940
1950
1960
1970
1980
1990
2000

FIGURE 4

1/k times the variance of k-differences of Consumption divided by variance of first difference 1889-1997

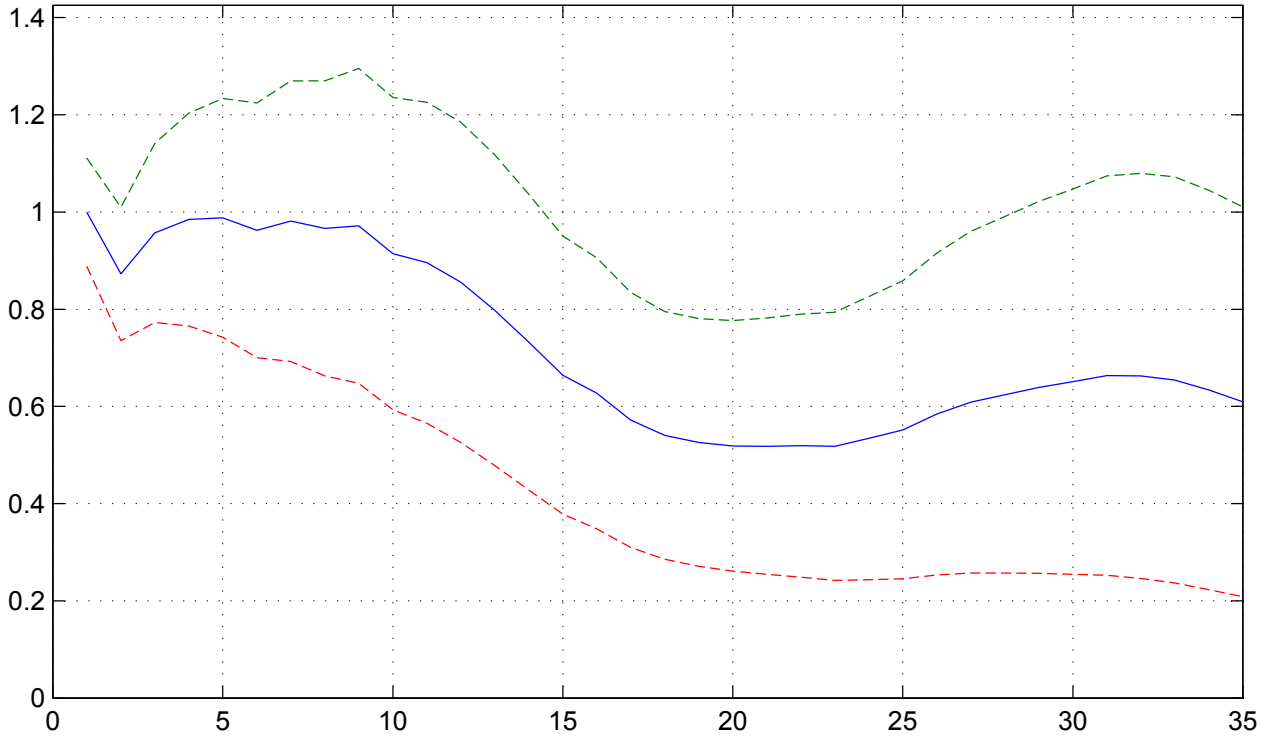
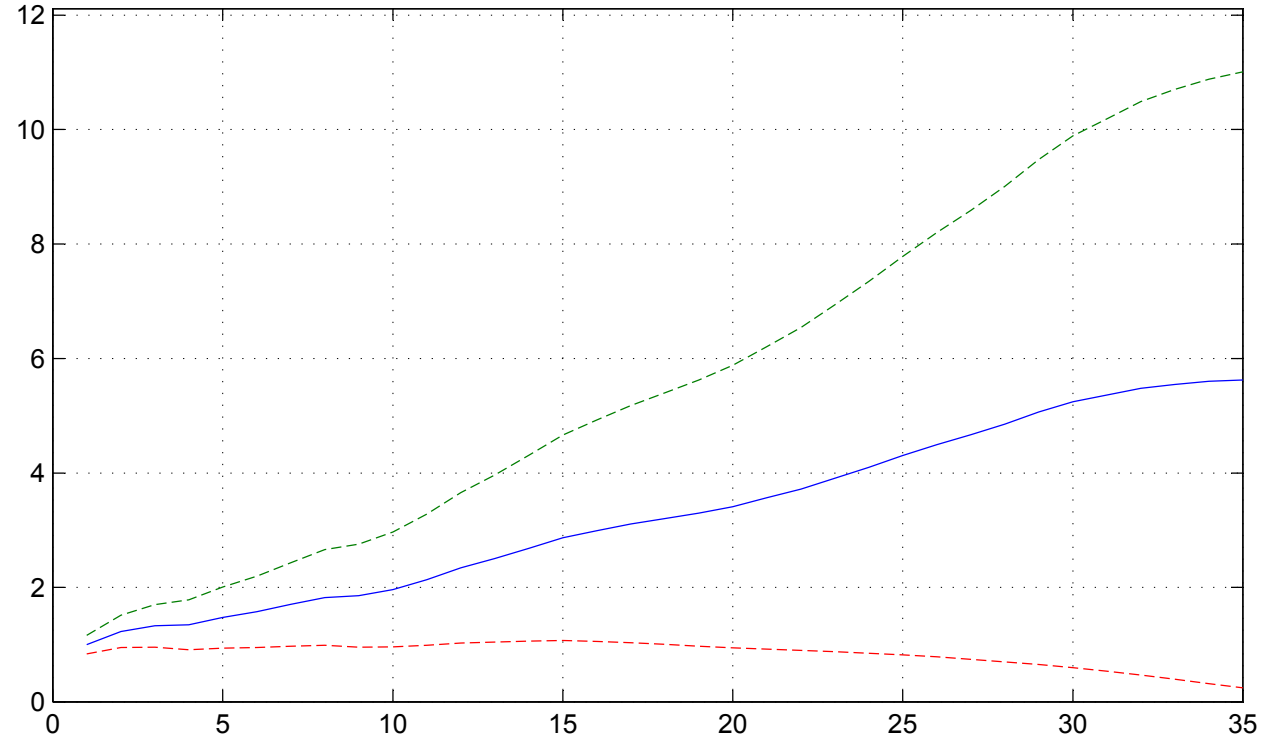


Figure 5

1/k times the variance of k-differences of Consumption divided by variance of first difference 1946-1997



Bands showing 1 asymptotic standard error; a period is one year.