Evidence on the Extent and Potential Sources of Long Memory in U.S. Treasury Security Returns and Yields

By

Robert A. Connolly**
CB3490, McColl Building
Kenan-Flagler Business School
UNC – Chapel Hill
Chapel Hill, NC
27599-3490 USA
(919) 962-0053 (O)
(919) 962-5539 (Fax)
connollr@bschool.unc.edu

Z. Nuray Güner
Department of Business Administration
Middle East Technical University
İnönü Bulvari 06531
Ankara, Turkey

90-312-210-3075 (O)
90-312-210-1107 (Fax)
nguner@ba.metu.edu.tr

Kenneth N. Hightower
Department of Economics
CB3305, Gardner Hall
UNC – Chapel Hill
Chapel Hill, NC
27599-3305 USA
(919) 966-2383 (O)
(919) 966-4986 (Fax)
Kenneth.Hightower@unc.edu

November 2002

** - Corresponding author. We gratefully acknowledge helpful comments on earlier work made by seminar participants at the University of Toronto, the University of North Carolina - Chapel Hill, the January 2000 Econometric Society meetings, and the June 2001 meetings of the Society for Computational Economics at Yale University. This work has been supported by the Turkish Academy of Sciences, in the framework of the Young Scientist Award Program (EA-TÜBA-GEBÝP/2001-1-1). Comments and suggestions from Kit Baum, Basma Bekdache, Patrick Conway, Kris Jacobs, Mark Jensen, Bill Parke, and Michael Salemi led to a much-improved paper. We are grateful to Gautam Kaul and Jennifer Conrad for making the original data set available to us and to Paisan Limratanamongkol for updating the data.

Evidence on the Extent and Potential Sources of Long Memory in U.S. Treasury Security Returns and Yields

Abstract

Unlike equity returns, many fixed-income return and volatility measures appear to display long memory. Granger and others have argued that long memory may only reflect infrequent structural breaks. We show that the extent of long memory in U.S. Treasury debt returns differs strongly for gross and excess holding period returns. We explore the impact of structural instability on tests for long memory using a version of the supLM test developed by Andrews (1993). We find only weak indications that structural instability lies behind the long memory in gross and excess returns. We also show that the evidence of long memory remains strong for yield and term premia series even after accounting for underlying structural changes.

Evidence on the Extent and Potential Sources of Long Memory in U.S. Treasury Security Returns and Yields

I. Introduction

Long memory is a widely accepted characteristic of asset return volatility. Many studies have shown that both exchange rate and stock market volatility display long memory. Bollerslev and Mikkelsen (1996), Liu (2000), and Andersen, Bollerslev, Diebold, and Labys (2001) have all proposed and applied volatility models that capture long memory properties.

By contrast, the evidence for long memory in asset returns is very spotty. For example, Lo (1991) and Lobato and Savin (1998) have both shown that equity returns display little evidence of long memory. There is some evidence that long memory may exist in exchange rates (Cheung, 1993) and gold (Booth, Kaen, and Kaveos, 1982). Nonetheless, a reasonable summary of the evidence would almost surely conclude that there is no significant case for long memory in asset returns. Indeed, Granger (1999) concludes that there is no reason to expect to see long memory in asset returns.

This paper extends the existing empirical research on long memory by documenting the long memory properties of Treasury security returns and yields. Our research is prompted by the fact that bill and bond returns display a leading characteristic of long memory: their autocorrelations are large, but die out very slowly (especially compared to equity return autocorrelations). This autocorrelation property is an important feature of fractionally-differenced time-series models, analyzed by Granger (1980), Hosking (1981), and Geweke and Porter-Hudak (1983).

The first contribution of this paper is to show that holding period returns and yields on short-term U.S. Treasury securities display strong evidence of long memory. This contrasts sharply with the results for returns on long-term U.S. Treasury securities where we find no evidence of long memory. An interesting puzzle emerging from this aspect of our work is that while there is long memory in gross holding period returns, excess holding period returns display only short-term memory. We consider potential explanations of this result later in the paper.

Why should finance researchers take an interest in long-memory models, particularly since there is so little evidence of their relevance to equity data? Perhaps the most consequential answer is that (unlike the equity literature) we find strong evidence of long memory in all of our data series except excess returns. This empirical regularity is important because Mandelbrot (1971) shows that there may be arbitrage opportunities in asset markets with long memory.

There are other considerations, too. Financial risk management systems typically use time-series representations of return behavior, but long memory does not appear to be incorporated into these products.[1] This assumption may be a reasonable approximation for short-horizon risk management, but neglected long-memory components in return and volatility phenomena may lead to inaccuracies in modeling and managing longer-horizon risks. The consequences of using an inappropriate time series model in this setting are not well known at present.

The unique long-horizon forecasting properties of long-memory models (discussed in Section II.A.) make them interesting to study, especially given the current interest in return predictability, particularly at long horizons. Andersson (1998) shows that ignoring long memory in forecasting exercises when it exists is worse than imposing long memory when it does not exist.

Finally, long memory is also important for pricing models. Backus and Zin (1993), Bollerslev and Mikkelsen (1996), and Comte and Renault (1996) are a few examples of papers that explore the consequences of long memory for pricing bonds, equity options, and interest rate options. It isn't clear yet whether there are significant gains to incorporating long memory into pricing models, but many of the earliest applications focused on equity markets where the evidence of long memory is weak.

Before reworking pricing models to include long memory components, there is a potentially important alternative interpretation of the long memory test statistics. Lobato and Savin (1998) warn that structural instability may lead to misinterpretation of long memory evidence. Granger and Hyung (1999) show that a linear process with structural breaks can mimic long memory, and they present simulation evidence that long memory in absolute S&P 500 returns (a volatility measure) is more likely due to structural breaks than an underlying I(d) process. Hightower and Parke (2001) demonstrate that certain structural stability tests and particular tests for long memory are related to one another. Each is a specific function of a common statistic based on the cumulative

---

1. Riskmetrics™ is one example. The documentation provided for the software appears to indicate clearly that ARIMA(p, d, q) models with d set either to zero or to one are standard in this risk management product.

sums of the error process. This implies that there is ambiguity in the interpretation of tests for long memory: evidence of an I(d) process may actually be structural instability in disguise.

Recent theoretical research has demonstrated that structural instability can produce the appearance of long memory in time series. Granger and Hyung (1999), Diebold and Inoue (2002), and Parke (1999) all demonstrate circumstances under which different types of structural instability can mimic the properties of long memory. Gray (1996) and Ang and Bekaert (2002a, 2002b) provide ample evidence of regime switching in short-maturity yields that can, at least in principle, generate long memory.

The second principal contribution of the paper, then, is an extended inquiry into whether the evidence for long memory in U.S. bond markets is simply an artifact of structural instability. Specifically, we reconsider the strength of the evidence for long memory using three different approaches to controlling for potential structural instability. We find only weak indications that structural instability explains our long memory evidence.

In the end, we conclude that there is still significant evidence of long memory in holding period returns and yields on short-maturity U.S. Treasury securities. In turn, this implies that efforts to price long bonds and interest rate derivatives may need to embrace models that include long memory components.

In the next section, we discuss the properties of fractionally-differenced time series and their potential use in modeling expected returns. We briefly review the properties of long memory, discuss tests for long memory, and show the connection between long memory and structural change tests. Section III describes our data and presents the basic results of our tests for long memory using U.S. Treasury security holding period returns and yields. Our empirical work uses a sample of hand-collected weekly holding period returns on seven (nearly constant-maturity) Treasury bills and bonds covering the July 1962 – May 1996 period. We analyze the structural stability issues in Section IV, and a final section summarizes the results of our study, and considers some of the implications of our findings.

Section II. Interpreting Tests of Long Memory

A. Introduction to Long Memory

Normally, only integer powers of d are considered in ARIMA(p, d, q) models, but there is no mathematical or statistical requirement that d take on only integer values (e.g., d = 1 yields a first-difference model). In a fractionally differenced model, d can take on noninteger values and the resulting time series can exhibit some particularly interesting dependencies. Granger and Joyeux (1980) and Hosking (1981) show that extending the lag operator to noninteger powers of d results in a well-defined time series that is fractionally integrated of order d.[2] The differencing operator may be written

$$(1-L)^d = \sum_{k=0}^{\infty} (-1)^k \binom{d}{k} L^k \tag{1}$$

leading to the following representation of a time series where p = q = 0:

$$(1-L)^d y_t = \sum_{k=0}^{\infty} \frac{\Gamma(k-d)}{\Gamma(-d)\Gamma(k+1)} \cdot y_{t-k} \quad . \tag{2}$$

Here, Γ is the usual gamma function.

In his excellent survey paper, Baillie (1996) reviews a number of different long-memory models. One simple model is an ARFIMA (0, d, 0) process given by

$$(1-L)^d (y_t - \mu) = \varepsilon_t \tag{3}$$

This model is studied in Granger (1980), Granger and Joyeux (1980), and Hosking (1981). Their work shows than when d < .5, the series has finite variance, but for d = .5, the series has infinite variance. The time series is stationary and invertible when -.5 < d < .5. For d = .5, standard Box-Jenkins techniques will indicate that differencing is required and provided that d < 1, differencing will produce a series whose spectrum is zero at zero frequency. This heavily used model is a special case of an ARFIMA (p, d, q) process given by

$$\phi(L)(1-L)^d (y_t - \mu) = \theta(L)\varepsilon_t \, . \tag{4}$$

where p = q = 0.

Fractionally-differenced time-series models have very interesting long-run forecasting properties. A fractional white noise series $y_t \sim I(d)$ may be represented as an MA(∞) process where the moving average coefficients decline slowly following the form $b_j \sim Aj^{d-1}$ where A is a constant. A stationary ARMA(p, q) with infinite p and q will have coefficients that decline at least exponentially: $b_j \sim A\theta^j$. One important implication of these stark differences in coefficient decay rates is that a fractionally-differenced model may provide better long-run forecasts from a very simple model compared to ARMA(p, q) models where p and q are large.[3]

2. See also Robinson (1994a) for early analysis of long-memory models.

3. For further discussion of the autocorrelation, autocovariance, and general forecasting properties of long-memory models, see Baillie's (1996) survey paper.

In principal, parameterizations of both finite-order ARMA and fractionally-differenced time series can produce dependence in a time series. The rate at which past information ceases to be useful in forecasting future values differs importantly across these models. Figure 1 provides an example of the autocorrelogram for a fractionally-differenced time series with d = .4 and an AR(1) process with $\rho = .5$.[4] The first autocorrelation for each series is nearly identical (.5 vs. .6) but the decay rates are quite different. The autocorrelations decline very slowly for the fractionally-differenced series, but fall quite rapidly to zero for the AR(1) process. This is an example of a fractionally-differenced time series displaying greater persistence than AR (or ARMA) processes, and why they may be interesting in research on debt instrument yield and return distributions.[5]

B. Testing for Long Memory

Kwiatkowski, Phillips, Schmidt, and Shin (1992) develop a test for I(0) behavior which is consistent against an I(d) alternative and can be helpful in distinguishing long memory from short memory. The null hypothesis of their test is that a time series is I(0), but under the alternative hypothesis, the time series displays I(d) behavior (with d ≤ 1). Lee and Schmidt (1996) provide further analysis of this approach to testing for long-memory effects. Their Monte Carlo evidence suggests that the KPSS test has power comparable to Lo's (1991) robust R/S statistic in distinguishing I(0) from I(d) behavior.

The first step in calculating the KPSS test statistic is to form the partial sum (S_t) of the residuals from the demeaned series.[6] The KPSS test statistic is given by

$$\eta_\mu = T^{-2}\sum_{t=1}^{T} S_t^2 / s_T^2(\ell) \tag{13}$$

where the denominator is the autocorrelation-consistent variance estimator given by

$$s^2(\ell) = T^{-1}\sum_{t=1}^{T} \hat{\varepsilon}_t^2 + 2T^{-1} + \sum_{\nu=1}^{T} \varkappa(\nu / \ell(T)) + \sum_{t=\nu+1}^{T} \hat{\varepsilon}_t \hat{\varepsilon}_{t-s}\,. \tag{14}$$

---

4. Following Lo, we simulated 5000 observations for each model. We discarded the first 3000 simulated data points and estimated the autocorrelations in the diagram from the remaining data.

5. In fact, Lo shows that a fractionally-differenced model can reproduce the general pattern of variance ratio results reported in the equity literature. In particular, he shows that a combination of an AR(1) and fractionally-differenced model with d = .25 will produce variance ratios above one at short horizons and below one at longer horizons. This suggests that fractionally-differenced models may have special importance in ongoing empirical investigations of long-range dependence in capital markets. Lo's own results suggest that there is little evidence of long-term memory in U.S. equity index returns once short-run dependencies have been accounted for in tests for long memory.

6. There is another version of the KPSS test, η_τ, which is constructed in the same way except that the residuals are derived from a regression that involves a time trend as well as an intercept term.

This robust variance estimator is based on Phillips (1987), who demonstrates its consistency under certain conditions and Newey and West (1987), who suggested the weighting scheme $\kappa(\nu/\ell(T)) = 1 - j/(\ell+1)$ to guarantee that the variance estimate is positive semi-definite.

This robust estimator is important to use. Lo (1991) shows that short-range dependence (well documented in equity prices by Lo and MacKinlay (1988), and Conrad and Kaul (1989)) may compromise inferences about the presence of long-range dependence. Since the optimal number of autocovariances is not known ex ante, we compute $\hat{\eta}_{\mu}$ using a number of different lag lengths (ℓ). The tradeoff is that using too few autocovariances produces an inadequate bias correction, but using too many leads to low power since the higher-order autocovariances are more imprecisely estimated.

Long memory tests have a very different basis than long horizon return dependence tests. Aside from some technical conditions, the null hypothesis of no long-range dependence eliminates infinite variance marginal distributions and encompasses a strong-mixing condition that requires higher-order autocorrelations to fall in size as the lag length increases. This means that the series of autocorrelations displayed by a time series under the null hypothesis decays rapidly. Included in the null hypothesis, then, are all finite-order ARMA models. The null hypothesis of no long-range dependence includes well-known models of return dependence (see Campbell, et. al. (1997), pgs. 59-64 for details). Put another way, the null hypothesis of the KPSS test excludes return behavior that is quite different from the autocorrelation properties commonly reported in earlier studies of long-horizon return dependence.

C. Interpreting Long Memory Tests

The long history covered by our time series samples of holding period returns is essential to empirical studies of long-memory processes, but it creates a potential difficulty. With a stable underlying structure, high frequency effects may be found by sampling very frequently, but over a relatively short time period (i.e., sample every 15 seconds for three business days). The long-memory phenomenon we are interested in may be measured accurately only in long samples, i.e., samples which extend over many realizations of the long-memory process. This is perhaps best achieved with samples that cover many years (say, 200 years), but where the process is not sampled with high frequency.[7] What is needed for long memory empirical studies is a long sample

---

7. Bollerslev, Cai, and Song (2000), among others, provide a clever alternative approach to identifying and estimating long-run properties of return distributions from intraday day. The relative strengths of these different approaches remain an interesting topic for further research.

realization of the process (i.e., 200 years), particularly one that is not sampled so often that short-run dependencies dominate the sample properties of the data.

The difficulty here is that the underlying structures of debt markets, instruments, and trading institutions and practices have not been stable over periods of even 50 years, let alone 100 to 200 years.[8] In our setting, the task is to provide the longest possible sample while recognizing that extending the length of the sample increases the probability of structural instability.[9]

The relationship between long memory and structural change has been the focus of increasing research in recent years. Some researchers (Diebold and Inoue (2002) and Granger and Hyung (1999)) see long memory as an artifact of processes that exhibit certain types of structural change over time. Others (Parke (1999) and Taqqu, Willinger and Sherman (1997)) propose models where many structural breaks that last for random durations give rise to time series properties that are associated with long memory (slowly decaying autocorrelations). Taken together, these lines of research point to a blurring of the differences between long memory and structural change.

Lobato and Savin (1998a) assess the fragility of evidence for long memory by splitting their sample of daily returns and squared returns into sub-samples. They recomputed their tests for, and measures of, long memory for each sub-sample, and then compared inferences from the whole sample and the sub-samples.[10]

In empirical studies, whether a process is deemed to be long memory or a structural break may come down to what types of tests are performed on the data. Hightower and Parke (2001) clarify this point by showing that structural shift tests and the KPSS test for I(d) (versus I(0)) behavior are both functions of the same term: the ratio of the partial sums of the series to a consistent variance estimate for that series. One interpretation of their theoretical results is that "large" KPSS test statistics may actually be evidence of structural shifts.

Andrews (1993) suggests a supremum test for a one-time structural change with an unknown break point as a way to account for the criticism that researchers may ``eyeball'' the most likely point for a break before running a typical test. Given a time series ε_i, residuals defined in the

8. One example is the Treasury Fed Accord of 1953 that ended the Fed's explicit policy of managing Treasury borrowing costs. A more recent example is the shift in Fed operating procedures in October 1979.

9. Ideally, instability tests such as those developed in Hansen (1992) might be used to resolve the instability issue empirically. Unfortunately, these tests are not available for our application. Hidalgo and Robinson (1996) have studied the issue of structural change in the mean with long memory for the case where the time series are Gaussian. There is substantial evidence rejecting Gaussianity for financial market return series, so this test does not seem to be particularly appropriate for the problem we are studying here.

10 . There is evidence of sub-sample instability from the equity market mean-reversion literature. In particular, see Kim and Nelson (1998) and the references therein.

usual manner as $\varepsilon_i = y_i - \overline{y}$, and defining $S_t = \sum_{i=1}^{t} \varepsilon_i$ to be the cumulative sum of squares, the LM test for a one-time change in mean at a point $\pi \in (0,1)$ can be shown to be

$$\mathrm{LM_T}(\pi) \cong \frac{\mathrm{T}^{-1}}{\pi(1-\pi)} \mathrm{S}^2_{[\mathrm{T}\pi]} / \mathrm{s}^2(\ell) \tag{26}$$

where

$$\mathrm{s}^2(\ell) = \mathrm{T}^{-1} \sum_{\mathrm{t}=1}^{\mathrm{T}} \hat{\varepsilon}_\mathrm{t}^2 + 2\mathrm{T}^{-1} + \sum_{\nu=1}^{\mathrm{T}} \varkappa(\nu / \ell(\mathrm{T})) + \sum_{\mathrm{t}=\nu+1}^{\mathrm{T}} \hat{\varepsilon}_\mathrm{t} \hat{\varepsilon}_{\mathrm{t-s}} \tag{27}$$

If $\varkappa(\nu / \ell(\mathrm{T}))$ is taken to be the Bartlett kernel, $\mathrm{s}^2(\ell)$ is identical to the denominator of the KPSS statistic. The Andrews supLM test is then simply given by $\sup_{\pi \in \Pi} LM_T(\pi)$ where Π is bounded away from 0 and 1.

The relationship between tests of stationarity and tests of structural change can be seen by a comparison of (13) and (26). The KPSS test for stationarity can be viewed as an average of the $LM_T(\pi)$ terms weighted by $\pi(1-\pi)$. Hightower and Parke (2001) show that these two tests (as well as the Andrews and Ploberger (1994) avgLM and expLM tests) have nearly identical power against many common alternatives including structural change, unit roots and long memory. In this context, the supLM test can provide insight into candidate breakpoints that may be driving rejections of short memory, if indeed these rejections are driven by a structural shift in the sample.

Bai and Perron (2001) develop this idea further, showing that multiple breakpoints may be estimated by using a sequential procedure. In Section IV, we use their procedure to assess whether instability in several potential macroeconomic state variables may lie behind the long memory evidence that we present next.

Section III. Data and Preliminary Results

A. Return and Yield Data

In the subsequent empirical analysis, we analyze gross and excess holding period returns on U.S. Treasury securities in addition to yields and term premia, all observed on a weekly basis. The basic return data set contains weekly holding period returns on one-, three-, six-, and 12-month Treasury Bills and three-, five-, and 10-year Treasury Bonds for the July 1962 through May 1996 period.[11] The yield data consist of weekly observations on three-, six-, and 12-month Treasury Bills and three-, five-, and 10-year Treasury bonds. We describe each in turn.

Weekly holding period returns were calculated by taking the log difference of this Wednesday's bid price and last Wednesday's ask price and adding in the percentage return associated with accrued interest. The bid and ask prices used to calculate the weekly return were for the same security. However, the security used in the computations was frequently changed so as to maintain a fairly constant maturity return series. In no case, though, were prices on different securities used to make return computations. The basic Treasury bill and bond price data is from the Wall Street Journal.[12]

We computed six excess return measures: the weekly holding period return on 10-year bonds (and five-year, three-year, 12-month, six-month, and three-month) less the weekly holding period return on one-month bills. This gives the extra return earned by holding a longer-maturity Treasury debt instrument vs. short-maturity Treasury debt.

The yield data is from the Federal Reserve Bank of St. Louis. The term premia studied here are calculated as the difference between the yield on 12-month (bill), three-year, five-year, and 10-year bond yields and the yield on three-month bills.

Figure 2 provides a graphical representation of the autocorrelations for gross holding period returns for Treasury Bills based on the full sample. Autocorrelations for shorter-maturity gross bill returns display the basic slow decay indicative of long memory. By contrast, the autocorrelations of Treasury Bond gross holding period returns presented in Figure 3 give no indication of long memory.

In Figure 4 (5), we provide a graphical representation of the autocorrelations of excess holding period returns for Treasury Bills (Bonds). By contrast with the gross holding period autocorrelations for bills, the decay in autocorrelations is very rapid. Based on the diagram alone, it would be hard to imagine these series would produce any evidence of long memory.

We present graphically Treasury Bill and Bond yield autocorrelations in Figures 6 and 7. Both show that yields are very persistent, a fact well known from the literature on bond pricing. The autocorrelations of term premia, presented in Figure 8, also display considerable persistence.

---

11. The original weekly return data was collected by Gautam Kaul and very graciously provided to us. We are grateful to Paisan Limratanamongkol who updated all the data series for us.

12. One motivation for using individual security data is to avoid aggregation of multiple security returns where possible. Granger (1980) provides an analysis of conditions under which aggregation can produce long memory.

The extent of this persistence is inversely related to the maturity of the bond for which the term premium is being calculated.

B. Long Memory Test Results

A fractionally-differenced model may produce the kind of dependence indicated by our autocorrelation graphs. We use the KPSS statistic (described in Section III, A.2) to test for the existence of fractional differencing. The results of the KPSS tests are reported in the panels of Table 1. For all nine return series except the weekly returns on five- and 10-year Treasury bonds, the null hypothesis of stationarity is strongly rejected in favor of an I(d) process. Together with the autocorrelation evidence, the KPSS test statistics clearly imply that weekly gross fixed income holding period returns display long memory which can be described using a fractionally-differenced ARIMA(p, d, q) model. The evidence of persistence is even stronger in the yields and term-premia, where almost every time series rejects stationarity for all choices of autocorrelation truncation parameter, ℓ.

Applied to weekly excess holding period returns, the KPSS test statistics show very little evidence of long-term dependence. This stands in sharp contrast to our just-noted findings about long memory in gross holding period returns. It suggests that empirical asset pricing work that focuses on excess returns can safely ignore the implications of neglected long memory.

Section IV. Diagnosing Long Memory

In this section, we examine whether the long memory evidence presented in this study is an artifact of structural breaks in the data. Since information on structural breaks in the data is not directly observable, to infer the timing of potential breaks we followed three approaches. In the first, we take a change in the Federal Reserve System operating procedures as a break point. Since changes in central bank behavior are known to affect financial markets, this choice seems to be a credible candidate for a structural break point.

In the second approach, we use the Bai-Perron tests to estimate the break points for two fundamental variables, namely money supply and inflation, that are expected to affect bond returns and yields. This method implicitly assumes that severe money supply and inflation shocks will produce immediate structural breaks in the bond return and yield series that we analyze here. We take the estimated structural break points for these fundamental variables as break points for the bond return and yield series. Subsamples constructed from these break points are analyzed for evidence of long memory. The advantage of this approach is that we are not looking at the series with the apparent long memory to find a structural break (and then using the structural break to argue that the long memory is an artifact of the structural break).

Instead of pinpointing structural change in a fundamental variable as the cause of a structural break/long memory in the bond return and yield series, the last method assumes that structural instability is precipitated by a latent variable. However, we assume that we can identify the shifts in this latent variable by analyzing the return and yield series directly. Accordingly, we apply Hamilton's (1994) Markov switching model methods to time the structural breaks in our bond return and yield series. In the rest of this section, we present findings based on applying the long memory tests discussed in Section II to subsamples chosen using these different methods.

A. Instability and Fed Operating Procedure

Borrowing from the Lobato-Savin approach, we initially explore the fragility of long memory evidence by comparing inferences across sub-samples and our full sample. More specifically, because of the shift in Federal Reserve System operating procedures in early October 1979, we split our U.S. Treasury bill and bond data at September 1979. As a first check on our earlier findings, we recalculated the KPSS test statistics using these sample breakpoints and checked the stability of the long-range dependence test statistics across the early and later samples. In this way, we hope to provide some initial evidence on the impact of potential structural instability on our inferences about long memory.

The various panels of Table 2 contain the sub-sample estimates of the KPSS test statistics for return series for which we reported results in Table 1. Analysis of the U.S. Treasury sub-sample results reveals that there is virtually no difference between the full sample and sub-sample inferences for any of the return, excess return, yield, or term premia series. Thus, it appears unlikely that the shift in Fed operating procedures is the culprit behind the long memory documented in Table 1.

B. Shifts in Fundamental Variables

B.1. Shifts in Money Supply (MS) Process

We also investigated the possibility that instability in Treasury debt markets may be related to instability in the underlying money supply or demand processes. The transmission mechanism from money supply or money demand shocks to bond markets has been a long-standing, central research theme in monetary economics. We chose M2 as the monetary aggregate in our empirical work.

The advantage of this approach is that we are not cooking the answer by searching for break points in the return series until we make the long memory evidence disappear. The disadvantage is that if the KPSS tests still find long memory, we can only conclude that potential structural instability in this series is probably not the source of long memory. Given the long history of research connecting money supply and demand processes with interest rates, this seems like a reasonable approach.

To assess the potential contributions of money supply or demand shifts to long memory in debt returns and yields, we sequentially estimate break points in the M2 series using the Bai-Perron procedure. We then use these break points to split our gross and excess return, yield, and term premia series and then retest the resulting subsamples for long memory using the KPSS test. The idea behind this is simple enough: if instability/persistence in the return series reflects instability in the monetary aggregate series, the evidence for long memory will be significantly weakened once the samples are chosen to isolate underlying structural breaks.

We report the estimated break points and the results of the KPSS tests in Table 3. The specific dates of the sample breaks are as follows: 1: July 4, 1962 – January 29, 1969, 2: February 5, 1969 – August 5, 1970, 3: August 12, 1970 – February 18, 1987, and 4: February 25, 1987 – May 29, 1996. Based on samples constructed using the four structural breaks that we found with the Bai-Perron procedure, we conclude that for the series analyzed there is still long memory in three of the four subsamples. Only in the second subsample did the evidence for long memory disappear, but the sample is so short (only 78 observations) that we believe it would be unwise to place much weight on this reversal.

B.2. Shifts in Inflation Process

Recall from Section III.B that we find persistence in yields, term premia, and gross (but not excess) returns. In the term premia analysis, the evidence favoring long memory appears to strengthen as the maturity difference increases. One partial explanation for this collection of results may come from an underlying instability in the inflation process that, in turn, produces differential persistence. There is some basic intuition that may prove useful. Hassler and Wolters (1995) find considerable evidence that inflation rates display long memory. Using the Fisher relation, we split the holding period return at time t on a riskless debt security with maturity of j months (R_{jt}^{N}) into three components, an expected real return over that horizon ($E(R_{jt}^{R})$), an inflation premium for that horizon ($E(\pi_t)$), and a contemporaneous random market-wide shock terms (υ_t) with mean zero:

$$R_{jt}^{N} = E(R_{jt}^{R}) + E(\pi_t) + \upsilon_t \tag{21}$$

The excess return is a linear combination of two holding period returns for securities with j and k months until maturity (and k =1 in practice):

$$R_{jt}^{N} - R_{1t}^{N} = E(R_{jt}^{R}) + E(\pi_t) + \upsilon_t - E(R_{1t}^{R}) - E(\pi_t) - \upsilon_t \tag{22}$$

$$R_{jt}^{N} - R_{1t}^{N} = E(R_{jt}^{R}) - E(R_{1t}^{R})\,. \tag{23}$$

Here, we rely on the assumptions that the inflation premium is independent of security maturity and the shock term is dominated by a market-wide component common to both maturities. This simple algebraic manipulation leads to our conjecture that if inflation is the source of the long memory in gross holding period returns (not expected returns), excess holding period returns will not display long memory.

In yields, by contrast, the expected inflation varies with maturity because the yield reflects the expected inflation over the life of the bond. We hypothesize that the nominal yield at time t on a riskless debt security with j periods until maturity (Y_{jt}^{N}) depends on four factors: the real yield for that maturity (Y_{jt}^{R}), the expected inflation rate over the life of the bond ($E(\pi_{jt})$), a maturity-specific disturbance term that captures shocks idiosyncratic to a particular maturity segment of the market (ν_{jt}), and a market-wide shock term (υ_t). Algebraically, we decompose the nominal yield as follows

$$Y_{jt}^{N} = Y_{jt}^{R} + E(\pi_{jt}) + \nu_{jt} + \upsilon_t\,. \tag{24}$$

The term premium is then given by

$$TP_{j-1,t} = Y_{jt}^{N} - Y_{1t}^{N} = Y_{jt}^{R} + E(\pi_{jt}) + \nu_{jt} + \upsilon_t - Y_{1t}^{R} - E(\pi_{1t}) - \nu_{1t} - \upsilon_t \tag{25}$$

$$TP_{j-1,t} = Y_{jt}^{R} - Y_{1t}^{R} + E(\pi_{jt}) - E(\pi_{1t}) + \nu_{jt} - \nu_{1t}\,. \tag{26}$$

From (26), it is clear that structural breaks in the expected inflation series or maturity-specific shocks, either one-time or in a continuing sequence as in Parke (1999), might produce evidence of long memory in the term premium series. Of course, by (24), it is clear that structural breaks in the expected inflation series or maturity-specific shocks might be able to produce long memory in the yield series, too.

The initial step in this portion of our empirical work is to find break points in the inflation series. The Bai-Perron procedure selects a three-break representation, with breaks in May 1967, January 1973, and July 1982. The shifts are matched against the underlying series in Figure 9. This selection of break points agrees closely with results in Bai and Perron (2001) who model U.S. real interest rates, proxied by the quarterly 3-month Treasury bill rate deflated by the CPI.

Despite this, running the KPSS tests on the subsamples based on inflation breaks seems to have very little impact on the evidence for long memory (see Table 4).

C. Modeling Shifts in Latent Variable

Finally, we investigated the possibility that structural instability might be due to a latent variable. In particular, we assume that a markov-switching model where which shifts between states follow an unobservable variable provides a straightforward approach to modeling structural change. Since markov-switching models have been analyzed extensively elsewhere (see Hamilton (1994), Ch. 22), we offer only a brief review of this modeling approach.

An easy starting point is to consider a segmented trends model where a latent variable s_t takes a value of 1 or 2 thereby indicating whether a time series is increasing in value or decreasing. The average change in the time series in the first state is given by μ_1, and the average change in state 2 is μ_2. The variance in each state is indicated by σ_1^2 and σ_2^2. The model is completed by assuming that a Markov chain describes the evolution of the latent (state) variable:

$$p(s_t = 1 \mid s_{t-1} = 1) = p_{11} \qquad p(s_t = 2 \mid s_{t-1} = 2) = p_{22}$$

$$p(s_t = 1 \mid s_{t-1} = 2) = 1 - p_{22} \qquad p(s_t = 1 \mid s_{t-1} = 1) = 1 - p_{11} \tag{1}$$

where the p_{ij} j are transition probabilities between states i and j.[13] Given the Markov chain assumption, s_{t-1} is sufficient to describe all past history of a time series and the latent variable.

The joint distribution of observed data (the value of the time series at time t is indicated by y_t) with sample size T may be written as $p(y_1, ..., y_t, s_1, ..., s_t \mid \theta)$ where s_t is the latent variable and $\theta = \{\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, p_{11}, p_{22}\}$. Given values for the parameter vector, the probability of being in regime s_t at date t given the available information at that time is the filter inference:

$$p(s_t \mid y_1, ..., y_t; \theta). \tag{2}$$

Alternatively, computing this probability with the full sample produces a "smoothed" inference about the regime:

$$p(s_t \mid y_1, ..., y_T; \theta). \tag{3}$$

When t=T, (2) and (3) produce the same value.

If the long memory evidence presented earlier is being produced by structural instability related to a latent variable, the smoothed state probabilities given by (3) are very interesting because they signal the timing of shifts in the time series from one state to another. Accordingly, if this type of structural instability is behind the long memory evidence, we expect the evidence for long memory to weaken considerably (if not disappear) when we limit long memory tests to samples drawn from only one state.

For the return and yield series where we find strong evidence of long memory (reported in earlier tables), we found clear evidence of switching between two regimes over the full sample. In addition, the state persistence is considerable so that there are long stretches of time when the return or yield series is in a single state. This makes it relatively less difficult to identify the dates of breaks in these time series. For the series where we don't find robust evidence of long memory, state persistence parameters are uniformly smaller, and the state probabilities indicate much more frequent switching between states. Dating the breaks is correspondingly more difficult, of course.

We report results of long memory tests for a subset of our time series using samples that are split using the state probabilities estimated by the markov-switching model. Figures 10 – 15 plot the smoothed state probabilities for six of our series. There is significant state persistence in the yield and term premia as well as three-month bill return series, but there is markedly less state persistence in the 10-year bond return and three-year excess return series.

If evidence of long memory is taken to be equivalent to evidence of structural instability, the smoothed state probabilities evidence just cited is potentially quite interesting. For the series where we find evidence of long memory, there seems to be clear evidence of regime switching and considerable state persistence. On the other hand, for the series where we find weak or nonexistent evidence of long memory, the states switch back and forth frequently making identification of a specific regime more difficult. We note in passing that many of the subsamples chosen from the inflation break dates overlap regimes from the Markov-switching model.

Perhaps the most interesting test centers on whether the long memory tests produce the same evidence on samples split using the regime shifts identified by the markov-switching model. In Table 5, we report results from these tests for an illustrative subsample of our series. On the whole, the test statistics appear to indicate nothing new: long memory (or the lack of long memory) where it is indicated in Table 1. Many of the test statistics, however, are considerably smaller than

---

13. In this application, we assume that the transition probabilities are constant, but Diebold, Lee, and Weinbach (1994) studied a markov-switching model in which the transition probabilities depend on a set of exogenous

before, suggesting that structural instability may have some role in explaining the long memory results reported in Table 1.

Section V. Summary and Further Discussion

As we noted in the introduction, a number of researchers have studied the relevance of long memory for macroeconomic and financial time series. There is reasonably widespread agreement that financial asset volatility displays long memory, but the evidence is mixed at best for other time series. One aim of this paper has been to extend the analysis of long memory models to U.S. bond markets. The other purpose of our research has been to explore potential foundations for the long memory evidence.

Our long memory tests show that Treasury security gross holding period returns, yields, and term premia all display long memory, but excess returns do not have this characteristic. We also document evidence that the long memory in returns decreases as the maturity of the debt instrument increases.

As we noted early in the paper, there is an important ambiguity in the interpretation of the KPSS test for long memory. This test shares an important connection with structural instability tests, and this renders the interpretation of large KPSS test statistics problematic. Accordingly, we have gone to some lengths to establish the sensitivity of these findings to structural instability.

We explored the impact of three different approaches to accounting for structural instability. The first method identified shifts in Federal Reserve operating procedure as the critical structural break. The second method used break dates in two fundamental variables, namely money supply (M2) and inflation, to date the critical structural breaks in return and yield series analyzed in this paper. These break dates in fundamental variables are identified by the Bai-Perron procedure. Our third approach relied on a latent variable approach and used Hamilton's markov-switching model to pick out potential breaks in the individual return, yield and term premia series. The first two methods have the advantage that they rely on more structural approaches to identifying sources of instability and do not use the return, yield and premia series themselves to find structural breaks. The third method makes fewer assumptions about the economic sources of instability, and dates any structural breaks using the time series themselves. This should give the sharpest identification of structural breaks.

---

variables.

The evidence from this analysis of structural instability is interesting, but it does not seem to provide an authoritative explanation of the long memory test results. The Federal Reserve operating procedure analysis identifies long memory as a feature of the post-October 1979 period. Indeed, even Treasury bond returns appear to display long memory in this subperiod. The M2-based analysis shows that long memory exists in all of the subsamples with a reasonably large set of observations. The inflation-based analysis also provides little new evidence of a structural instability explanation of our long memory results.

We also evaluated the sensitivity of long memory tests to splitting samples based on markov-switching model estimates. This portion of our work is grounded in the analytical results reported by Diebold and Inoue (2002). Here, we found that while test statistics were generally far smaller than in the tests reported in Table 1, our assessment was largely unchanged. That is, we still found long memory in gross return, yield and term premia series, but not in excess returns. We found long memory in shorter maturity return series, but not in returns of long maturity debt instruments.

We believe there is at least one alternative path yet to be explored. In particular, this involves assessing the impact of long memory on term structure, bond pricing, and fixed income derivative models. Specifically, it would be useful to know whether long memory is related to any pricing biases and whether there are circumstances in which the impact of ignoring long memory is the most (and the least) noticeable.

Table 1

Results of KPSS Test for Long Memory

Returns	Autocovariance Lag Length 0	4	8	12
1-month	22.298	5.334	3.069	2.178
3-month	12.067	4.420	2.806	2.098
6-month	5.109	2.632	1.932	1.595
12-month	1.991	1.235	1.034	0.951
3-year	0.601	0.405	0.356	0.347
5-year	0.415	0.310	0.272	0.260
10-year	0.174	0.141	0.129	0.128

Yields	Autocovariance Lag Length 0	4	8	12
3-month	30.114	6.078	3.412	2.388
6-month	30.950	6.240	3.501	2.449
12-month	33.886	6.829	3.828	2.676
3-year	45.847	9.226	5.163	3.602
5-year	52.849	10.628	5.941	4.140
10-year	63.024	12.661	7.068	4.917

Excess Returns	0	4	8	12
3m-1m	0.708	0.562	0.515	0.493
6m-1m	0.209	0.154	0.140	0.138
12m-1m	0.231	0.164	0.150	0.150
3y-1m	0.422	0.293	0.261	0.257
5y-1m	0.243	0.186	0.165	0.159
10y-1m	0.097	0.079	0.073	0.072

Term Premium	0	4	8	12
12m-3m	1.807	0.391	0.230	0.169
3y-3m	27.981	5.799	3.333	2.383
5y-3m	32.850	6.753	3.856	2.742
10y-3m	39.207	8.017	4.560	3.232

Notes: The KPSS test statistics reported here are calculated using (13) in the text. The number of lags in the autocovariance adjustment affects the nature of the consistent covariance matrix estimate ((14) in the text). KPSS test statistics that indicate a rejection of the null hypothesis at the five per cent significance level are shaded.

Table 2

Results of KPSS Test for Long Memory With Sample Split at October 1979

Panel A: First Half Sample

Returns	Autocovariance Lag Length 0	4	8	12	Yields	Autocovariance Lag Length 0	4	8	12
1-month	32.327	7.460	4.316	3.061	3-month	39.557	8.062	4.557	3.206
3-month	13.020	5.399	3.561	2.641	6-month	42.799	8.705	4.916	3.459
6-month	5.806	3.231	2.527	2.099	12-month	47.825	9.722	5.491	3.865
12-month	1.393	0.918	0.823	0.742	3-year	64.515	13.073	7.358	5.158
3-year	0.539	0.369	0.348	0.340	5-year	71.096	14.374	8.070	5.643
5-year	0.319	0.232	0.221	0.215	10-year	78.760	15.887	8.897	6.206
10-year	0.122	0.085	0.086	0.084					
Excess Returns					**Term Premium**				
3m-1m	0.278	0.260	0.259	0.241	12m-3m	3.735	0.848	0.507	0.371
6m-1m	0.151	0.116	0.116	0.118	3y-3m	11.386	2.374	1.360	0.963
12m-1m	0.111	0.081	0.079	0.076	5y-3m	10.947	2.260	1.287	0.908
3y-1m	0.106	0.073	0.069	0.068	10y-3m	11.240	2.304	1.307	0.921
5y-1m	0.105	0.076	0.073	0.071					
10y-1m	0.065	0.045	0.045	0.044					

Panel B: Second Half Sample

Returns	Autocovariance Lag Length 0	4	8	12	Yields	Autocovariance Lag Length 0	4	8	12
1-month	42.398	10.537	6.107	4.366	3-month	58.795	11.892	6.697	4.703
3-month	22.157	8.362	5.380	4.093	6-month	60.778	12.277	6.906	4.844
6-month	8.567	4.532	3.362	2.831	12-month	62.365	12.587	7.073	4.955
12-month	2.918	1.835	1.551	1.472	3-year	64.880	13.076	7.333	5.128
3-year	1.396	0.935	0.811	0.792	5-year	65.780	13.254	7.428	5.190
5-year	1.114	0.841	0.725	0.690	10-year	67.247	13.542	7.581	5.290
10-year	0.455	0.383	0.344	0.343					
Excess Returns					**Term Premium**				
3m-1m	0.974	0.761	0.703	0.694	12m-3m	9.829	2.097	1.236	0.910
6m-1m	0.215	0.158	0.142	0.140	3y-3m	4.211	0.882	0.514	0.374
12m-1m	0.141	0.100	0.092	0.093	5y-3m	6.178	1.283	0.743	0.537
3y-1m	0.391	0.272	0.241	0.239	10y-3m	9.512	1.965	1.133	0.816
5y-1m	0.412	0.321	0.281	0.270					
10y-1m	0.180	0.154	0.139	0.139					

Notes: The KPSS test statistics reported here are calculated using (13) in the text. The number of lags in the autocovariance adjustment affects the nature of the consistent covariance matrix estimate ((14) in the text). KPSS test statistics that indicate a rejection of the null hypothesis at the five per cent significance level are shaded.

Table 3

KPSS Tests from Samples Split Using M2 Break Points

Panel A: Returns and Yields

Returns	**Autocovariance Lag Length**				**Yields**	**Autocovariance Lag Length**			
1st Subsample	**0**	**4**	**8**	**12**	**1st Subsample**	**0**	**4**	**8**	**12**
1-month	20.548	5.164	3.041	2.187	**3-month**	28.020	5.744	3.270	2.322
3-month	14.315	4.760	2.955	2.162	**6-month**	28.784	5.895	3.356	2.383
6-month	4.953	2.481	1.795	1.453	**12-month**	29.133	5.959	3.387	2.402
12-month	1.461	0.919	0.741	0.639	**3-year**	30.164	6.158	3.493	2.470
3-year	0.106	0.070	0.060	0.053	**5-year**	30.849	6.298	3.571	2.522
5-year	0.102	0.070	0.059	0.054	**10-year**	30.464	6.222	3.525	2.488
10-year	0.060	0.045	0.049	0.050					
2nd Subsample	**0**	**4**	**8**	**12**	**2nd Subsample**	**0**	**4**	**8**	**12**
1-month	0.950	0.438	0.329	0.263	**3-month**	2.277	0.503	0.313	0.249
3-month	0.363	0.294	0.231	0.234	**6-month**	2.042	0.448	0.279	0.221
6-month	0.081	0.071	0.070	0.086	**12-month**	2.048	0.460	0.294	0.239
12-month	0.143	0.105	0.107	0.140	**3-year**	4.762	1.046	0.646	0.501
3-year	0.395	0.281	0.284	0.330	**5-year**	5.625	1.227	0.751	0.572
5-year	0.311	0.220	0.224	0.267	**10-year**	6.243	1.355	0.830	0.635
10-year	0.045	0.044	0.054	0.090					
3rd Subsample	**0**	**4**	**8**	**12**	**3rd Subsample**	**0**	**4**	**8**	**12**
1-month	19.450	4.950	2.898	2.085	**3-month**	28.505	5.780	3.263	2.297
3-month	8.951	3.720	2.496	1.937	**6-month**	29.171	5.906	3.331	2.343
6-month	3.478	1.946	1.525	1.333	**12-month**	31.868	6.447	3.634	2.554
12-month	1.614	1.019	0.887	0.847	**3-year**	40.750	8.223	4.619	3.234
3-year	0.334	0.226	0.208	0.215	**5-year**	43.956	8.861	4.968	3.473
5-year	0.340	0.251	0.225	0.220	**10-year**	49.027	9.868	5.520	3.849
10-year	0.312	0.246	0.230	0.231					
4th Subsample	**0**	**4**	**8**	**12**	**4th Subsample**	**0**	**4**	**8**	**12**
1-month	14.518	3.612	2.063	1.449	**3-month**	23.107	4.640	2.588	1.800
3-month	11.113	4.173	2.474	1.763	**6-month**	23.729	4.768	2.662	1.853
6-month	4.487	2.640	1.785	1.358	**12-month**	24.558	4.938	2.761	1.925
12-month	1.159	0.915	0.726	0.636	**3-year**	28.252	5.701	3.202	2.245
3-year	0.304	0.248	0.211	0.200	**5-year**	30.472	6.164	3.472	2.441
5-year	0.150	0.136	0.124	0.123	**10-year**	33.894	6.864	3.869	2.721
10-year	0.116	0.106	0.094	0.093					

Table 3 (cont.)

KPSS Tests from Samples Split Using M2 Break Points

Panel B: Excess Returns and Term Premia

Excess Returns	**Autocovariance Lag Length**			
1st Subsample	**0**	**4**	**8**	**12**
3m-1m	1.327	1.041	0.915	0.801
6m-1m	0.306	0.222	0.195	0.186
12m-1m	0.146	0.107	0.094	0.087
3y-1m	0.158	0.108	0.092	0.082
5y-1m	0.264	0.182	0.155	0.141
10y-1m	0.148	0.112	0.120	0.123
2nd Subsample	**0**	**4**	**8**	**12**
3m-1m	0.127	0.135	0.124	0.149
6m-1m	0.040	0.037	0.040	0.055
12m-1m	0.107	0.080	0.084	0.115
3y-1m	0.371	0.265	0.267	0.313
5y-1m	0.300	0.213	0.217	0.259
10y-1m	0.047	0.046	0.056	0.093
3rd Subsample	**0**	**4**	**8**	**12**
3m-1m	0.637	0.503	0.476	0.468
6m-1m	0.239	0.171	0.156	0.155
12m-1m	0.339	0.234	0.214	0.214
3y-1m	0.149	0.102	0.094	0.096
5y-1m	0.187	0.139	0.124	0.121
10y-1m	0.226	0.178	0.165	0.165
4th Subsample	**0**	**4**	**8**	**12**
3m-1m	0.568	0.506	0.402	0.341
6m-1m	0.138	0.138	0.126	0.119
12m-1m	0.134	0.125	0.113	0.112
3y-1m	0.556	0.448	0.379	0.356
5y-1m	0.176	0.162	0.148	0.148
10y-1m	0.156	0.144	0.128	0.127

Term Premia	**Autocovariance Lag Length**			
1st Subsample	**0**	**4**	**8**	**12**
12m-3m	2.051	0.451	0.268	0.198
3y-3m	1.673	0.367	0.220	0.163
5y-3m	4.218	0.896	0.526	0.385
10y-3m	9.895	2.066	1.200	0.870
2nd Subsample	**0**	**4**	**8**	**12**
12m-3m	0.647	0.186	0.125	0.105
3y-3m	4.595	1.162	0.730	0.567
5y-3m	4.649	1.114	0.693	0.537
10y-3m	4.094	0.913	0.554	0.424
3rd Subsample	**0**	**4**	**8**	**12**
12m-3m	4.372	0.949	0.563	0.414
3y-3m	8.509	1.775	1.027	0.739
5y-3m	8.300	1.720	0.991	0.710
10y-3m	8.359	1.726	0.993	0.711
4th Subsample	**0**	**4**	**8**	**12**
12m-3m	4.256	0.885	0.509	0.364
3y-3m	3.334	0.685	0.392	0.280
5y-3m	4.748	0.966	0.546	0.386
10y-3m	5.732	1.159	0.651	0.456

Notes: The KPSS test statistics reported here are calculated using (13) in the text. The number of lags in the autocovariance adjustment affects the nature of the consistent covariance matrix estimate ((14) in the text). KPSS test statistics that indicate a rejection of the null hypothesis at the five per cent significance level are shaded. The specific dates of the sample breaks are as follows: 1: 7/4/62 – 1/29/69, 2: 2/5/69 – 8/5/70, 3: 8/12/70 – 2/18/87, and 4: 2/25/87 – 5/29/96.

Table 4

KPSS Tests from Samples Split Using Inflation Break Points

Panel A: Returns and Yields

Returns	Autocovariance Lag Length			
1st Subsample	**0**	**4**	**8**	**12**
1-month	15.426	4.162	2.443	1.748
3-month	13.132	4.052	2.407	1.723
6-month	5.438	2.437	1.638	1.256
12-month	2.590	1.268	0.962	0.790
3-year	0.324	0.169	0.139	0.120
5-year	0.243	0.149	0.122	0.107
10-year	0.088	0.057	0.067	0.070
2nd Subsample	**0**	**4**	**8**	**12**
1-month	6.404	1.495	0.872	0.624
3-month	2.312	0.971	0.640	0.493
6-month	0.574	0.376	0.312	0.285
12-month	0.305	0.210	0.182	0.166
3-year	0.314	0.224	0.204	0.191
5-year	0.407	0.277	0.257	0.251
10-year	0.203	0.156	0.158	0.152
3rd Subsample	**0**	**4**	**8**	**12**
1-month	19.786	5.210	3.088	2.243
3-month	7.598	3.383	2.354	1.873
6-month	2.420	1.375	1.126	1.030
12-month	0.585	0.356	0.325	0.333
3-year	0.058	0.039	0.038	0.043
5-year	0.047	0.034	0.032	0.033
10-year	0.037	0.027	0.027	0.030
4th Subsample	**0**	**4**	**8**	**12**
1-month	29.155	7.338	4.252	3.022
3-month	20.395	7.141	4.376	3.239
6-month	10.000	5.126	3.392	2.614
12-month	4.271	3.029	2.240	1.842
3-year	2.629	1.779	1.354	1.148
5-year	1.802	1.432	1.159	1.019
10-year	1.017	0.906	0.758	0.685

Yields	Autocovariance Lag Length			
1st Subsample	**0**	**4**	**8**	**12**
3-month	21.670	4.387	2.471	1.739
6-month	20.751	4.204	2.370	1.671
12-month	20.428	4.141	2.336	1.646
3-year	20.338	4.127	2.328	1.638
5-year	20.631	4.190	2.364	1.662
10-year	20.191	4.139	2.350	1.655
2nd Subsample	**0**	**4**	**8**	**12**
3-month	7.667	1.568	0.893	0.634
6-month	7.299	1.497	0.855	0.610
12-month	6.279	1.294	0.743	0.533
3-year	4.472	0.925	0.533	0.384
5-year	5.296	1.096	0.632	0.453
10-year	7.990	1.653	0.951	0.681
3rd Subsample	**0**	**4**	**8**	**12**
3-month	29.948	6.109	3.466	2.452
6-month	30.747	6.264	3.553	2.513
12-month	32.215	6.564	3.723	2.632
3-year	35.371	7.199	4.069	2.866
5-year	36.121	7.348	4.147	2.916
10-year	38.147	7.748	4.363	3.059
4th Subsample	**0**	**4**	**8**	**12**
3-month	42.767	8.698	4.894	3.424
6-month	44.000	8.948	5.036	3.527
12-month	44.987	9.140	5.144	3.606
3-year	49.161	9.994	5.634	3.956
5-year	50.241	10.212	5.756	4.044
10-year	52.057	10.575	5.958	4.183

Table 4 (cont.)

KPSS Tests from Samples Split Using Inflation Break Points

Panel B: Excess Returns and Term Premia

Excess Returns	**Autocovariance Lag Length**			
1st Subsample	**0**	**4**	**8**	**12**
3m-1m	2.045	1.416	1.100	0.900
6m-1m	0.736	0.489	0.401	0.353
12m-1m	0.708	0.398	0.330	0.290
3y-1m	0.295	0.157	0.130	0.112
5y-1m	0.295	0.182	0.150	0.132
10y-1m	0.149	0.096	0.112	0.119
2nd Subsample	**0**	**4**	**8**	**12**
3m-1m	0.266	0.232	0.208	0.203
6m-1m	0.169	0.144	0.143	0.156
12m-1m	0.111	0.084	0.078	0.075
3y-1m	0.378	0.273	0.250	0.235
5y-1m	0.470	0.321	0.299	0.292
10y-1m	0.239	0.184	0.186	0.179
3rd Subsample	**0**	**4**	**8**	**12**
3m-1m	0.268	0.229	0.239	0.252
6m-1m	0.051	0.037	0.035	0.037
12m-1m	0.083	0.055	0.053	0.057
3y-1m	0.147	0.100	0.098	0.110
5y-1m	0.106	0.078	0.073	0.075
10y-1m	0.133	0.098	0.096	0.106
4th Subsample	**0**	**4**	**8**	**12**
3m-1m	2.099	1.425	1.046	0.889
6m-1m	1.298	0.983	0.749	0.640
12m-1m	0.970	0.824	0.664	0.584
3y-1m	1.514	1.069	0.831	0.717
5y-1m	1.148	0.943	0.779	0.697
10y-1m	0.689	0.625	0.529	0.483

Term Premia	**Autocovariance Lag Length**			
1st Subsample	**0**	**4**	**8**	**12**
12m-3m	0.741	0.184	0.123	0.098
3y-3m	7.316	1.704	1.089	0.840
5y-3m	14.606	3.155	1.883	1.386
10y-3m	18.337	3.819	2.209	1.589
2nd Subsample	**0**	**4**	**8**	**12**
12m-3m	5.968	1.338	0.789	0.577
3y-3m	17.013	3.577	2.051	1.454
5y-3m	18.126	3.746	2.131	1.503
10y-3m	17.982	3.679	2.085	1.469
3rd Subsample	**0**	**4**	**8**	**12**
12m-3m	10.529	2.296	1.358	0.993
3y-3m	3.933	0.834	0.489	0.355
5y-3m	5.030	1.057	0.616	0.447
10y-3m	5.893	1.230	0.715	0.517
4th Subsample	**0**	**4**	**8**	**12**
12m-3m	3.440	0.733	0.436	0.325
3y-3m	14.702	3.035	1.753	1.268
5y-3m	9.582	1.964	1.123	0.803
10y-3m	6.426	1.312	0.746	0.529

Notes: The KPSS test statistics reported here are calculated using (13) in the text. The number of lags in the autocovariance adjustment affects the nature of the consistent covariance matrix estimate ((14) in the text). KPSS test statistics that indicate a rejection of the null hypothesis at the five per cent significance level are shaded. The specific dates of the sample breaks are as follows: 1: 7/4/62 – 5/3/67, 2: 5/10/67 – 1/3/72, 3: 1/10/72 – 6/16/82, and 4: 6/23/82 – 5/29/96.

Table 5

KPSS Tests from Samples Split Using Markov-Switching Model

Number of Lags in Autocovariance Adjustment

3-Mth. Returns

start	end	0	4	8	12
1	**847**	9.884	4.178	2.822	2.117
848	**1198**	0.702	0.391	0.303	0.271
1199	**1769**	11.265	4.307	2.570	1.838

3-Mth Yield

start	end	0	4	8	12
1	**852**	33.254	6.774	3.832	2.700
853	**1262**	16.651	3.443	1.988	1.435
1263	**1378**	7.614	1.722	1.038	0.777
1379	**1511**	10.545	2.200	1.275	0.925
1512	**1770**	10.888	2.207	1.244	0.875

10-Yr. Returns

start	end	0	4	8	12
1	**280**	0.216	0.144	0.157	0.156
281	**1769**	0.223	0.182	0.166	0.164

10-Yr. Yield

start	end	0	4	8	12
1	**855**	74.474	15.026	8.418	5.874
856	**1258**	6.729	1.386	0.793	0.566
1259	**1322**	5.240	1.145	0.704	0.540
1322	**1505**	5.578	1.198	0.721	0.546
1506	**1770**	5.335	1.097	0.629	0.453

3-Yr. Excess Returns

start	end	0	4	8	12
1	**924**	0.327	0.214	0.201	0.198
925	**1191**	0.125	0.090	0.080	0.077
1092	**1769**	0.535	0.398	0.324	0.294

3-Yr. Term Premium

start	end	0	4	8	12
1	**469**	8.444	1.851	1.100	0.807
470	**567**	3.432	0.748	0.453	0.340
568	**679**	7.711	1.719	1.048	0.777
680	**866**	11.102	2.396	1.409	1.036
867	**1025**	3.596	0.787	0.485	0.376
1026	**1396**	7.418	1.576	0.940	0.706
1397	**1513**	8.069	1.753	1.055	0.793
1514	**1733**	2.488	0.525	0.309	0.229
1734	**1770**	2.337	0.539	0.353	0.300

Notes: The KPSS test statistics reported here are calculated using (13) in the text. The number of lags in the autocovariance adjustment affects the nature of the consistent covariance matrix estimate ((14) in the text). KPSS test statistics that indicate a rejection of the null hypothesis at the five per cent significance level are shaded. Sample break points are as follows:

1) 3-month Gross Returns [1: 7/4/62 – 9/20/78, 2: 9/27/78 – 6/12/85, and 3: 6/19/85 – 5/29/96],
2) 10-year Gross Returns [1: 7/4/62 – 11/8/67 and 2: 11/15/67 – 5/29/96],
3) 3-year Excess Returns [1: 7/4/62 – 3/12/80, 2: 3/19/80 – 4/24/85, and 3: 5/1/85 – 5/29/96],
4) 3-month Yield [1: 7/4/62 – 10/25/78, 2: 11/1/78 – 9/10/86, 3: 9/17/86 – 11/30/88, 4: 12/7/91 – 6/19/91, and 5: 6/26/91 – 5/29/96],
5) 10-year Bond Yield [1: 7/4/62 – 11/15/78, 2: 11/22/78 – 8/13/86, 3: 8/19/86 – 11/4/87, 4: 11/11/87 – 5/8/91, and 5: 5/15/91 – 5/29/96],
6) 3-year Term Premium [1: 7/4/62 – 6/23/71, 2: 6/30/71 – 5/9/73, 3: 5/16/73 – 7/2/75, 4: 7/9/75 – 1/31/79, 5: 2/7/79 – 2/17/82, 6: 2/24/82 – 4/5/89, 7: 4/12/89 – 7/3/91, 8: 7/10/91 – 9/20/95, and 9: 9/27/95 – 5/29/96]

Figure 1: Autocorrelogram for AR(1) Series (Rho = .5) and Fractionally-Differenced Series (d = .40)

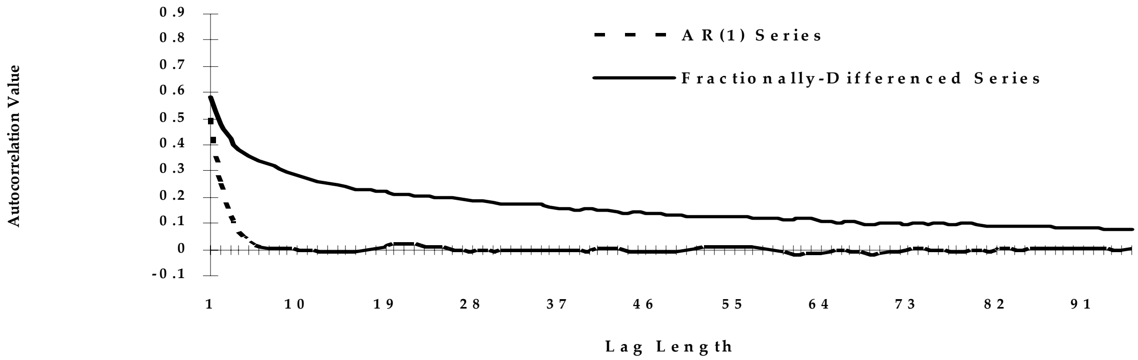


Figure 2: Autocorrelations of Weekly Gross Treasury Bill Holding Period Returns

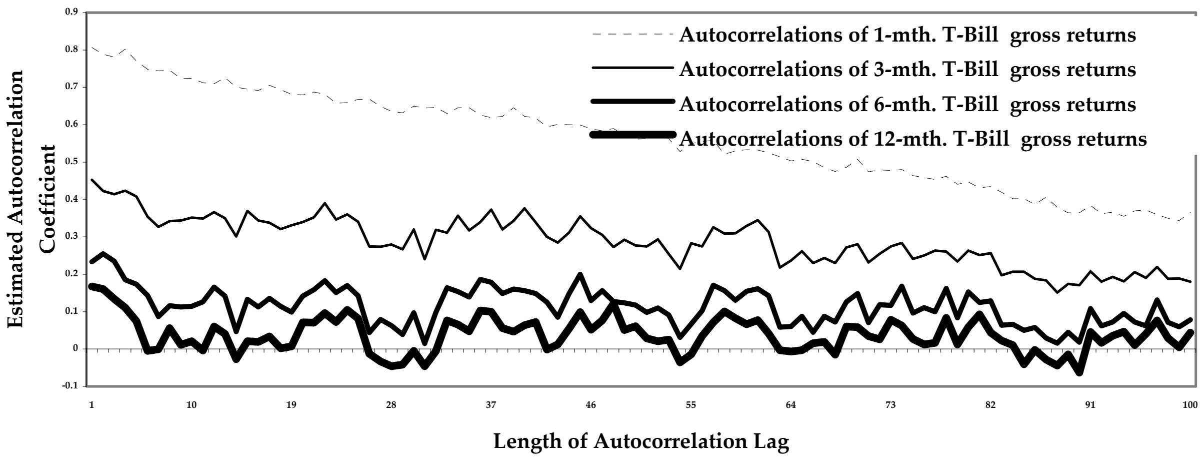


Figure 3: Autocorrelations of Weekly Gross Holding Period Returns on Treasury Bonds

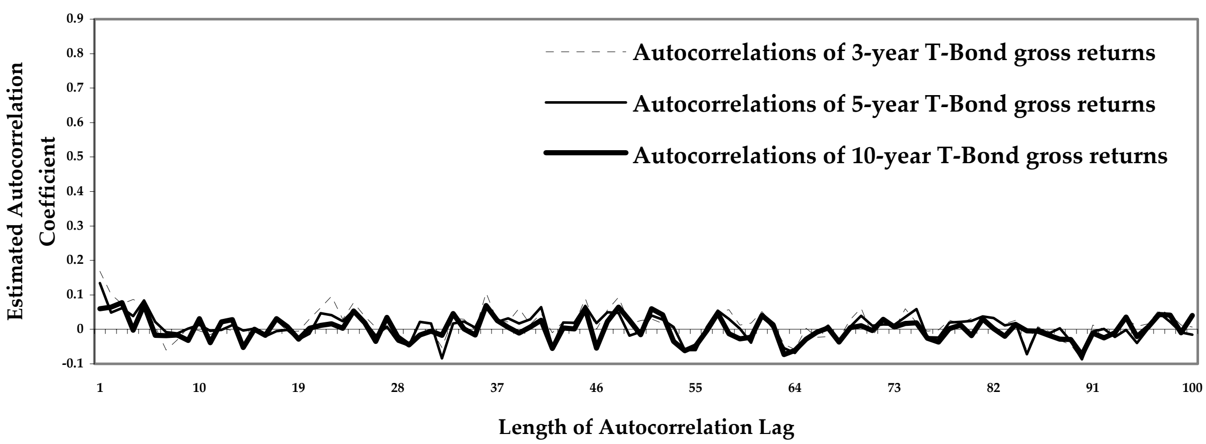


Figure 4: Autocorrelations of Weekly Excess Treasury Bill Holding Period Returns

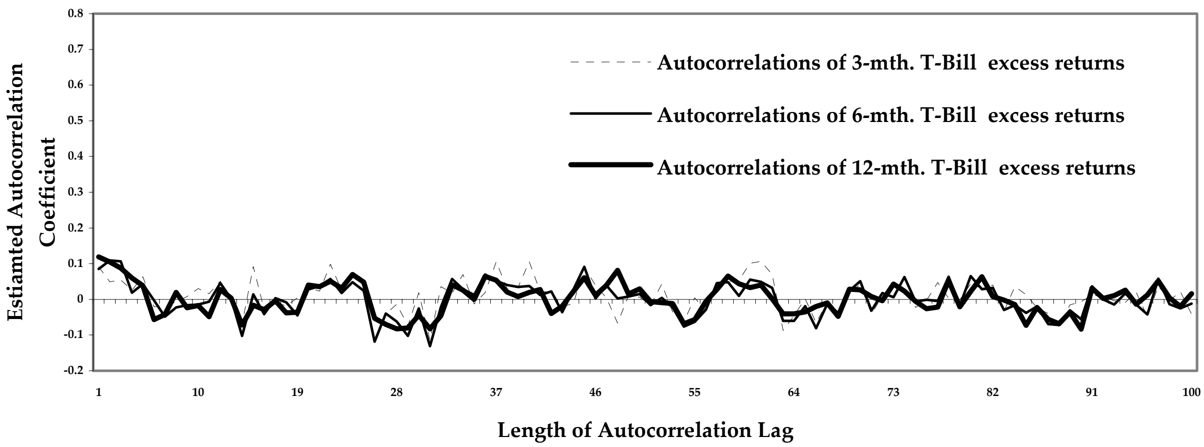


Figure 5: Autocorrelations of Weekly Excess Holding Period Returns on Treasury Bonds

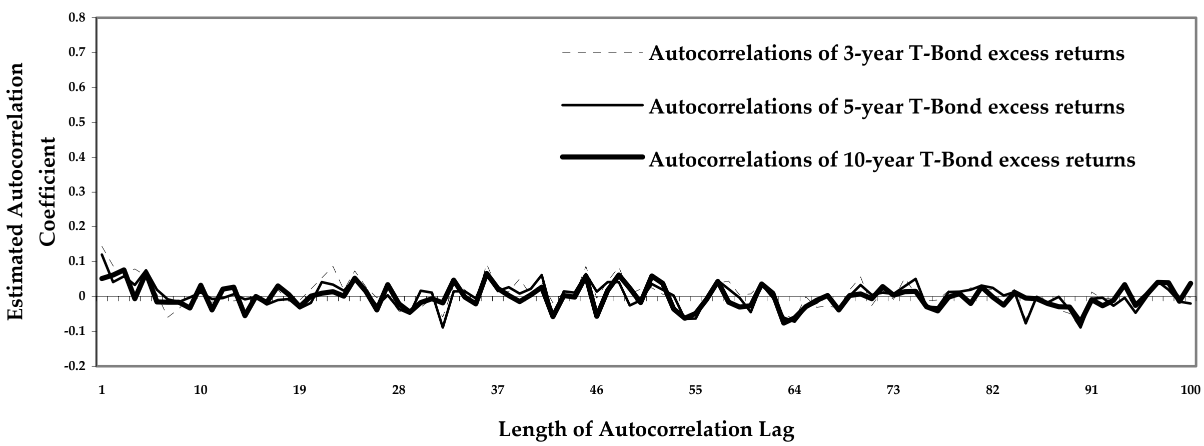


Figure 6: Autocorrelations of Weekly Treasury Bill Yields

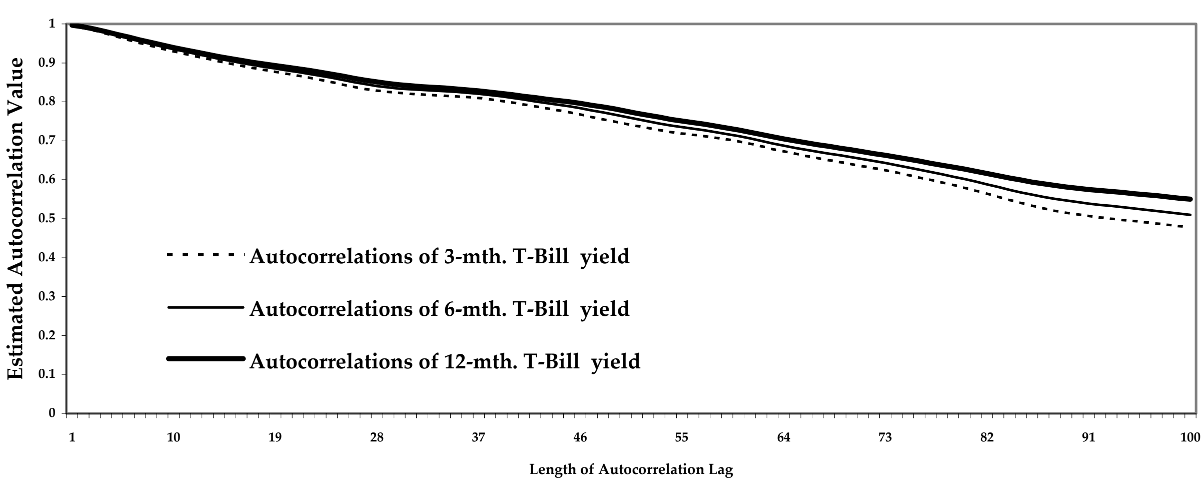


Figure 7: Autocorrelations of Weekly Treasury Bond Yields

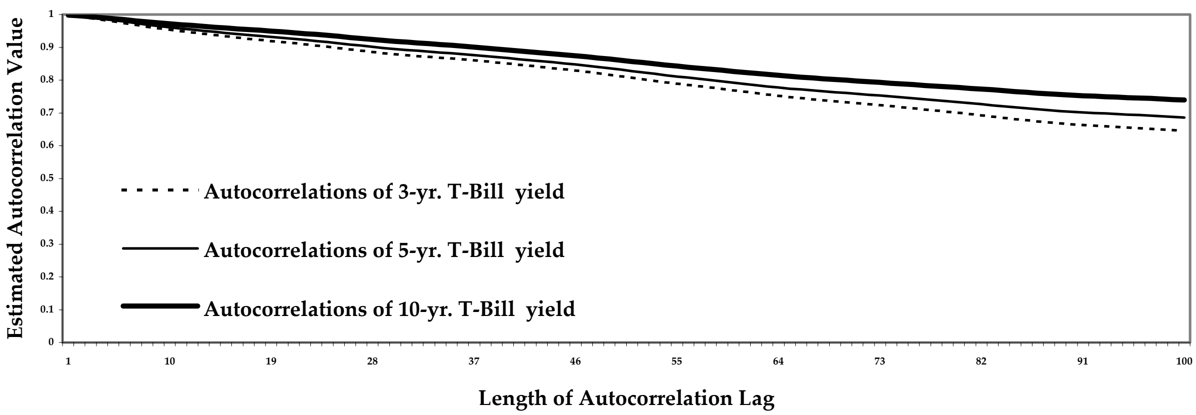


Figure 8: Autocorrelations of Weekly Treasury Bill and Bond Term Premia

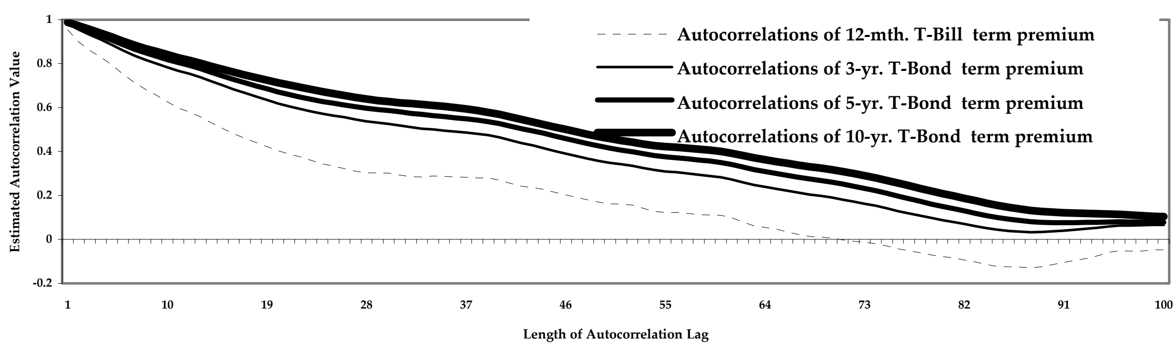


Figure 9 - Inflation Series with Estimated Break Points

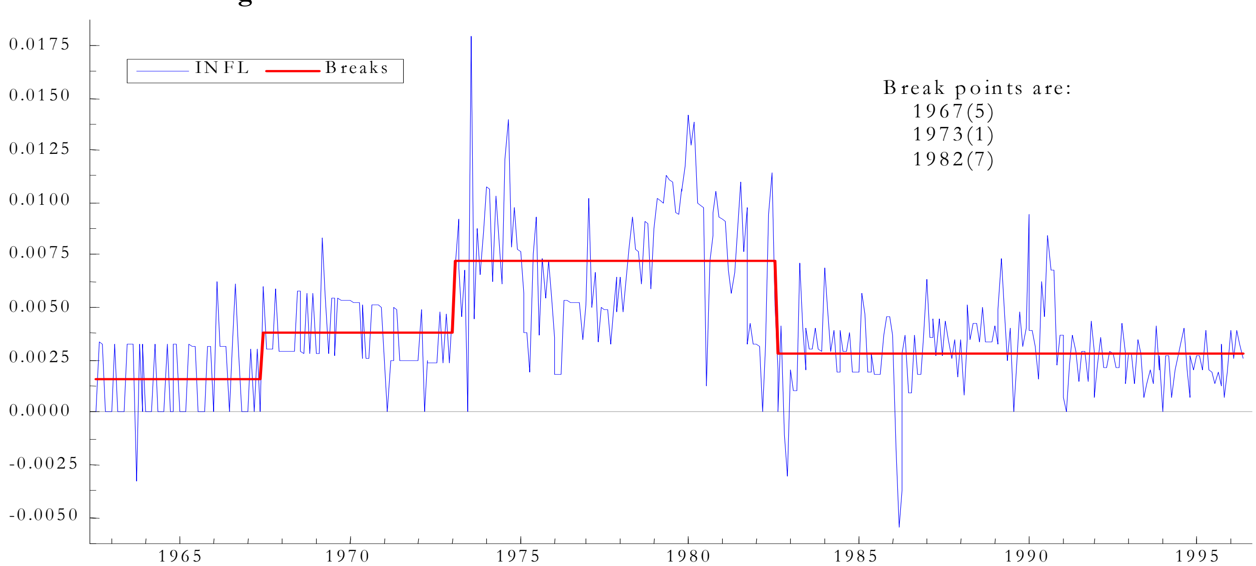


Figure 10 – Three-Month Bill Return Data and Smoothed State Probabilities

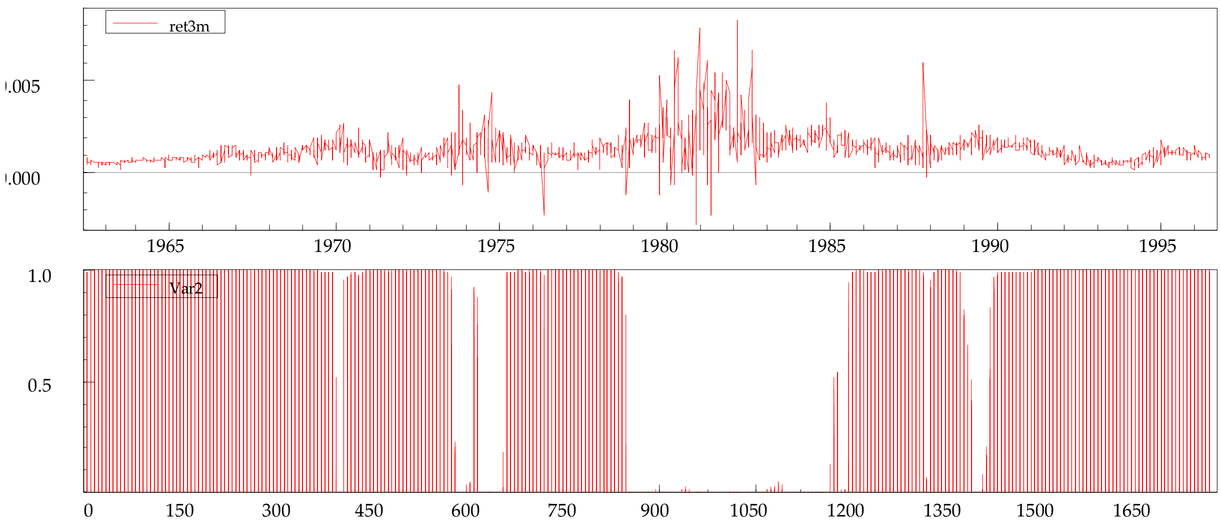


Figure 11 – Ten-Year Bond Return Data and Smoothed State Probabilities

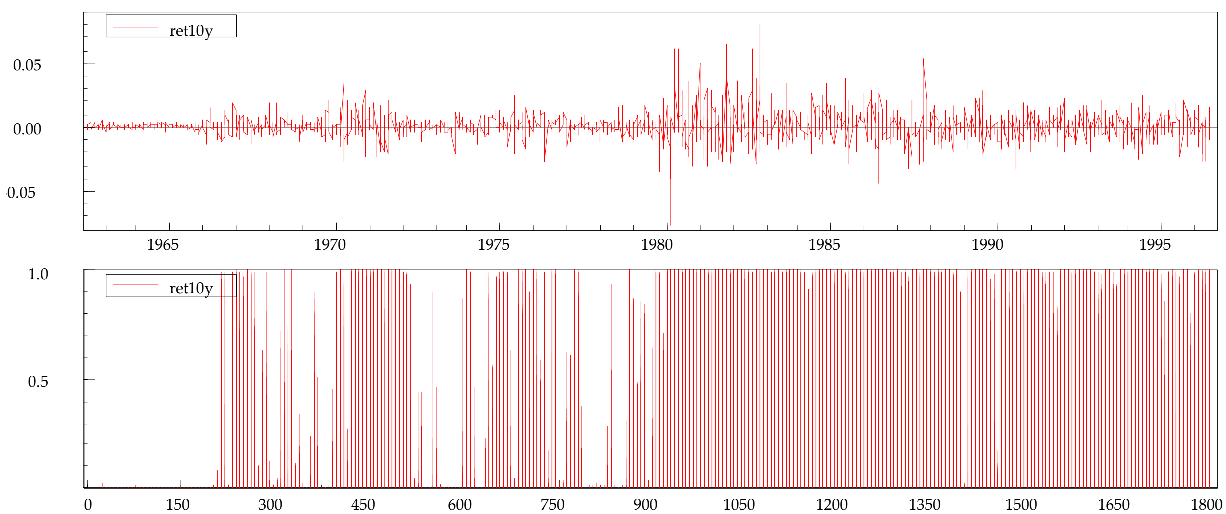


Figure 12 – Three-Month Bill Yield Data and Smoothed State Probabilities

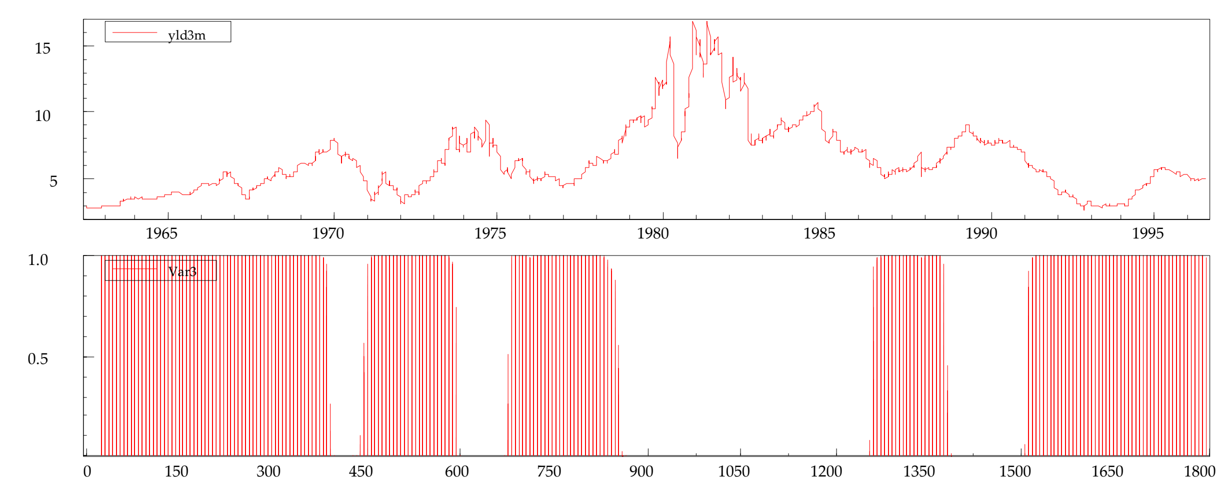


Figure 13 – Ten-Year Bond Yield Data and Smoothed State Probabilities

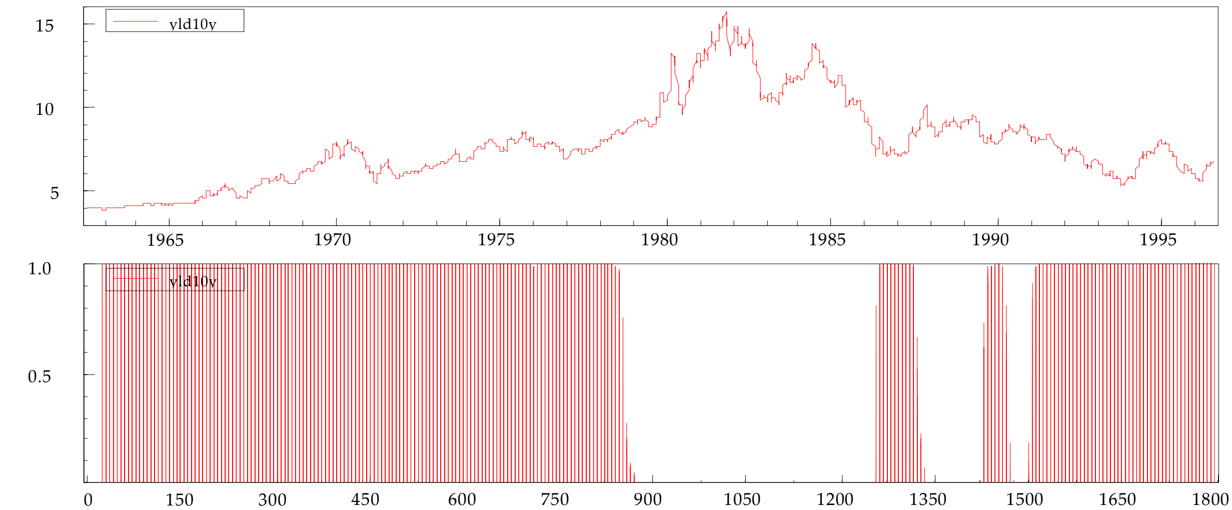


Figure 14 – Three-Month Bill Excess Return Data and Smoothed State Probabilities

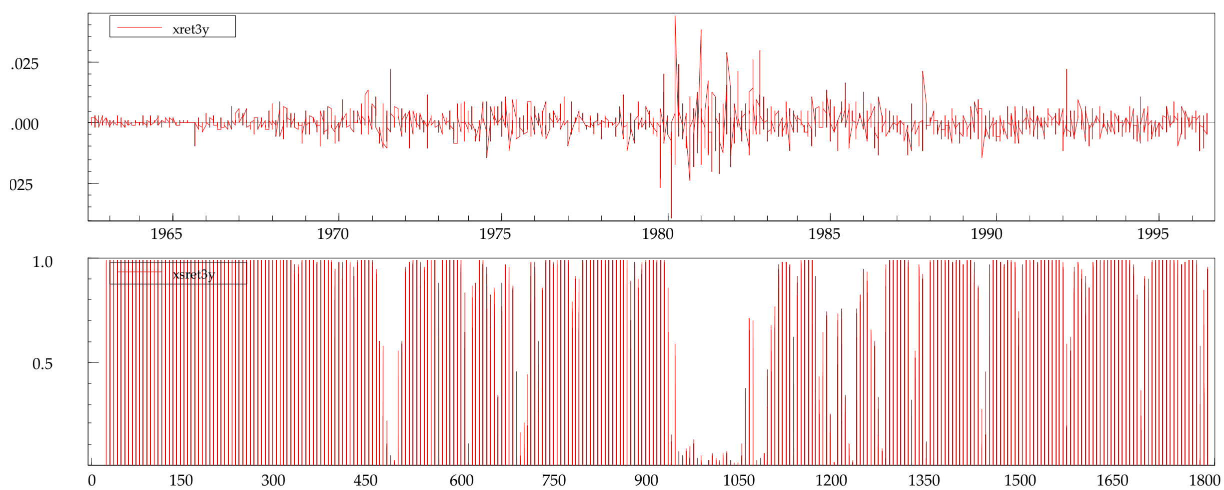
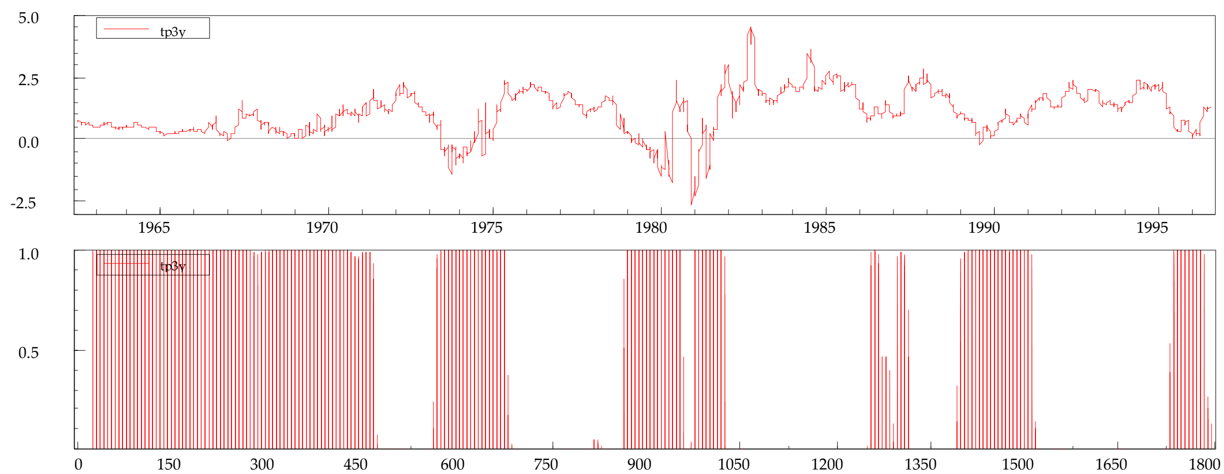


Figure 15 – Three-Year Term Premium Data and Smoothed State Probabilities



REFERENCES

Anderson, T. G., Bollerslev, T., Diebold, F. X., and Labys, P., (2001) "The Distribution of Exchange Rate Volatility," *Journal of the American Statistical Association*, 96, 42-55.

Andersson, M. K. (1998), "On the Effects of Imposing or Ignoring Long Memory when Forecasting," Working Paper Series in Economics and Finance, No. 225, Stockholm School of Economics.

Andrews, D. W. (1993), "Tests for Parameter Instability and Structural change with Unknown change Point," *Econometrica*, 61, 821-856.

Andrews, D. W. and Ploberger, W., (1994), "Optimal Tests When a Nuisance Parameter is Present Only Under the Alternative," *Econometrica*, 62, 1383-1414.

Ang, A. and Bekaert, G., (2002a), "Short Rate Nonlinearities and Regime Switches," *Journal of Economic Dynamics and Control*, 26, 1243 – 1274.

Ang, A. and Bekaert, G., (2002b), "Regime Switches in Interest Rates," *Journal of Business and Economic Statistics*, 20, 163-182

Backus, D. K., and Zin, S. E. (1993), "Long-Memory Inflation Uncertainty: Evidence from the Term Structure of Interest Rates," *Journal of Money, Credit, and Banking*, 25, 681-700.

Bai, J. and Perron, P., (2001), "Computation and Analysis of Multiple Structural Change Models, working paper, Boston University.

Baillie, R. T. (1996), "Long Memory Processes and Fractional Integration in Econometrics," *Journal of Econometrics*, 73, 5-59.

Baillie, R. T., and Bollerslev, T. (1994), "The Long Memory of the Forward Premium," *Journal of International Money and Finance*, 13, 565-571.

Baillie, R. T., Bolleslev, T., and Mikkelsen, H. (1996), "Fractionally Integrated Generalized Autoregressive Conditional Heteroscedasticity," *Journal of Econometrics*, 74, 3-30.

Bollerslev, T., Cai, J., and Song, F., (2000), "Intraday Periodicity, Long-Memory Volatility, and Macroeconomic Announcements Effects in the U.S. Treasury Bond Market," *Journal of Empirical Finance*, 7, 37-55.

Bollerslev, T., and Mikkelsen, H.-O. (1996), "Modeling and Pricing Long-Memory in Stock Market Volatility," *Journal of Econometrics*, 74, 3-30.

Booth, G., Kaen, F., and Kaveos, P. (1982), "Persistent Dependence in Gold Prices," *Journal of Financial Research*, 5, 85-93.

Campbell, J., Lo, A., and MacKinlay, C., (1997), *The Econometrics of Financial Markets*, (Princeton, N.J.: Princeton University Press).

Cheung, Y.-W. (1993), "Long Memory in Foreign-Exchange Rates," *Journal of Business and Economic Statistics*, 11, 93-101.

Comte, F. and Renault, E. (1996), "Long Memory Continuous Time Models," *Journal of Econometrics*, 73, 101-149.

Conrad, J. and Kaul, G., (1989) "Mean Reversion in Short-Horizon Expected Returns," *Review of Financial Studies*, 2, 225-240.

Diebold, F. X. and Inoue, A., (2002), "Long Memory and Regime Switching," *Journal of Econometrics*, 105, 131-159.

Geweke, J. F. and Porter-Hudak, S. (1983), The Estimation and Application of Long Memory Time Series Models," *Journal of Time Series Analysis*, 4, 221-238.

Granger, C. W. J. (1980), Long Memory Relationships and the Aggregation of Dynamic Models," *Journal of Econometrics*, 14, 227-238.

Granger, C. W. J., and Joyeux, R. (1980), "An Introduction to Long-Memory Time Series Models and Fractional Differencing," *Journal of Time Series Analysis*, 1, 15-29.

Granger, C. W. J. (1999), "Aspects of Research Strategies for Time Series Analysis," unpublished presentation to New Developments in Time Series Economics conference, New Haven, CT.

Granger, C. W. J., and Hyung, Namwon (1999), "Occasional Structural Breaks and Long Memory," unpublished working paper 99-14, University of California-San Diego, Department of Economics (available online at http://weber.ucsd.edu/Depts/Econ/Wpapers/dp99.html).

Gray, S., (1996), "Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process," *Journal of Financial Economics*, 42, 27-62.

Hamilton, J., (1994), *Time Series Analysis*, (Princeton, N.J.: Princeton University Press).

Hansen, Bruce E. (1992), "Testing for Parameter Instability in Linear Models," *Journal of Policy Modeling*, 14, 517-533.

Hassler, U. and Wolters, J. (1995), "Long Memory in Inflation Rates: International Evidence," *Journal of Business and Economic Statistics*, 13, 37-45.

Helms, B., F. Kaen, and Rosenman R. (1984), "Memory in Commodity Futures Contracts," *The Journal of Futures Markets*, 4, 559-567.

Hidalgo, J. and Robinson, P. M. (1996), "Testing for Structural Change in a Long-Memory Environment," *Journal of Econometrics*, 70, 159-174.

Hightower, K. N., and Parke, W. R. (2001), "Stability," unpublished working paper, University of North Carolina – Chapel Hill, Department of Economics.

Hosking, J. (1981), "Fractional Differencing," *Biometrika*, 68, 165-176.

Kim, M. and Nelson, C., (1998), "Testing for Mean Reversion in Heteroskedastic Data (II): Autoregression Tests Based on Gibbs-Sample-Augmented Randomization," *Journal of Empirical Finance*, 5, 385-396.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., and Shin, Y. (1992), "Testing the Null Hypothesis of Stationarity Against the Alternatives of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root?," *Journal of Econometrics*, 54, 159-178.

Lee, D. and Schmidt, P. (1996), "On the Power of the KPSS Test of Stationarity Against Fractionally-Integrated Alternatives," *Journal of Econometrics*, 73, 285-302.

Liu, M., (2000), "Modeling Long Memory in Stock Market Volatility," *Journal of Econometrics*, 99, 139-171.

Lo, A. (1991), "Long-Term Memory in Stock Market Prices," *Econometrica*, 59, 1279-1313.

Lo, A. and MacKinlay, C., (1988), "Stock Market Prices Do Not Follow Random Walks, Evidence from a Simple Specification Test," *Review of Financial Studies*, 1, 41-66.

Lobato, I. N., and Savin, N. E. (1998), "Real and Spurious Long-Memory Properties of Stock-Market Data," *Journal of Business and Economic Statistics*, 16, 261-268.

Lobato, I. N., and Savin, N. E. (1998), "Reply," *Journal of Business and Economic Statistics*, 16, 280-283.

Newey, W. and West, K. (1987), "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703-708.

Parke, W. R. (1999), "What is Fractional Integration?," *Review of Economics and Statistics*, 81, 632-638.

Phillips, P. C. B. (1987), "Time Series Regression with a Unit Root," *Econometrica*, 55, 277-302.

Robinson, P. M. (1994a), "Time Series with Strong Dependence," in Sims, C. A., ed., *Advances in Econometrics: Sixth World Congress, Vol. 1* (Cambridge: Cambridge University Press), 47-96.

Taqqu, M., Willinger, W., and Sherman, R. (1997), "Proof of a Fundamental Result in Self-Similar Traffic Modeling," Computer Communication Review, 27, 5-23.