Stationarity and the Term Structure of Interest Rates: A Characterisation of Stationary and Unit Root Yield Curves

Clive G. Bowsher and Roland Meeks

Research Department
Working Paper 0811

Federal Reserve Bank of Dallas

Stationarity and the Term Structure of Interest Rates: A Characterisation of Stationary and Unit Root Yield Curves

Clive G. Bowsher[*]and Roland Meeks[†]

October 24, 2008

Abstract

The nature of yield curve dynamics and the determinants of the integration order of yields are investigated using a benchmark economy in which the logarithmic expectations theory holds and the regularity condition of a limiting yield and limiting term premium is satisfied. By considering a zero-coupon yield curve with a complete term structure of maturities, a linear vector autoregressive process is constructed that provides an arbitrarily accurate moving average representation of the complete yield curve as its cross-sectional dimension (n) goes to infinity. We use this to prove the following novel results. First, any I(2) component vanishes owing to the almost sure (a.s.) convergence of the innovations to yields, $\nu_t(n)$, as $n \to \infty$. Second, the yield curve is stationary if and only if $n\nu_t(n)$ converges a.s., or equivalently the innovations to log discount bond prices converge a.s.; otherwise yields are I(1). A necessary condition for either stationarity or the absence of arbitrage is that the limiting yield is constant over time. Since the time-varying component of term premia is small in various fixed-income markets, these results provide insight into the critical determinants of the stationarity properties of the term structure.

Keywords: Term structure of interest rates, yield curve, integration, cointegration, expectations hypothesis, long rate, ET-VAR.

1 Introduction

Whilst econometric models of the term structure typically treat the yield curve as a non-stationary, integrated process, the models of continuous time finance typically imply stationary yield curves. Methods widely used to evaluate the well-known expectations theory of the term structure – such as those based on the present value model tests of Campbell and Shiller (1987, 1991) – are often only valid when yields are integrated of order one or I(1). More recently, there has been a growing interest in allowing even for the possibility that nominal variables such as money supply and interest rate levels or spreads are I(2) (see, e.g., Johansen, Juselius, Frydman, and Goldberg 2008). However there are not currently available parsimonious, economically interpretable conditions that determine the order of integration of the yield curve and there is hence a poor theoretical understanding of the determinants of the stationarity properties of the term structure. Furthermore there is little

---

[*]Corresponding author: C.Bowsher@statslab.cam.ac.uk; *tel:* +44 (0)1223 766859, *fax:* +44 (0)1223 337956; Statistical Laboratory, Centre for Mathematical Sciences, University of Cambridge (Wilberforce Road, Cambridge, U.K.) and Nuffield College, University of Oxford.

[†]Roland.Meeks@dal.frb.org; *tel:* +1 214 922 6804; Research Dept., Federal Reserve Bank of Dallas, 220 North Pearl St., Dallas TX75201, US. The views expressed in this paper are not necessarily those of the Federal Reserve Bank of Dallas or the Federal Reserve System.

theoretical guidance, apart from the implied long-run behaviour of yields, for the choice between the various alternatives in empirical work.

We address this problem in the context of the logarithmic Expectations Theory (ET) since it both constitutes an important empirical and theoretical benchmark, and also provides a tractable framework for analysis of the problem at hand. By considering a term structure of maturities that is complete or 'has no gaps', we reveal the critical role played by the convergence properties of the innovations to zero-coupon yields and to log discount bond prices at the long maturity end of the term structure. Under a weak regularity condition, these convergence properties are the sole determinants of the integration order of the yield curve.[1]

The main contributions of the paper are as follows. A *linear* vector autoregressive process (the ET-VAR) is constructed that provides an arbitrarily accurate moving average representation of complete, ET-consistent yield curves as the cross-sectional dimension (n) goes to infinity. We thus prove that the ET is incompatible with non-linear vector autoregressive specifications of the dynamics of the (complete) term structure. The moving average representation is used to prove the following novel results. First, although models with I(2) yield curves and stationary curvatures satisfying the ET may be constructed, the I(2) component vanishes in the presence of a limiting yield and term premium, owing to the *a.s.* convergence of the innovations to yields, $\nu_t(n)$, as maturity $n \to \infty$. Second, the yield curve is either I(1) or stationary depending on the rate of convergence of $\nu_t(n)$: the yield curve is stationary if and only if $n\nu_t(n)$ converges *a.s.*, that is if and only if the innovations to log discount bond prices converge. A necessary condition for either stationarity or the absence of arbitrage is that $\nu_t(n)$ converges to zero, which implies the time-invariance of the limiting yield. Since for many fixed-income markets, such time-invariance seems implausibly strong from an empirical perspective, our results give support to the substantial econometric literature that evaluates the ET using procedures that are only valid when yields are indeed I(1). However, the lack of mean reversion of an I(1) process indicates that the ET should generally be applied only to finite intervals of time.

Introduced over a hundred years ago and brought to prominence in the writings of Fisher (1930), Keynes (1930) and Hicks (1953), the expectations theory has been one of the most intensively studied models in financial economics. Nevertheless, the contributions just discussed significantly extend understanding of the implications of the ET for the dynamics of the yield curve. The theory continues to provide both an important empirical benchmark and baseline theory, and to find widespread application. For example, the ET is a direct building block in the modern economic forecasting and financial models used in practice by most central banks (Roush 2007); and recent

---

[1]We do not require the almost sure existence of a limiting yield or limiting term premium – see Condition 1.

2

theoretical, macro-finance models that jointly derive the dynamics of the term structure and the macroeconomy within a DSGE setting generate bond yields that satisfy the ET by adopting a homoscedastic approximation to the economy's pricing kernel (see Wu 2006).

Research during the last 15 years has provided substantial empirical evidence in support of the expectations theory in the case of developed economies other than the United States (Hardouvelis 1994, Dahlquist and Jonsson 1995, and Gerlach and Smets 1997), shorter term maturities and alternative securities less affected by the factors driving the specialness of US Treasuries (Longstaff 2000b). It seems that a particular focus on US Treasury bill and bond yields has perhaps led in the past to an overemphasis on the empirical deviations from the ET found using such data. For seminal evaluations of the ET for US Treasuries see, *inter alia*, Fama and Bliss (1987), Campbell and Shiller (1991), and Cochrane and Piazzesi (2005). In careful econometric work, Bekaert and Hodrick (2001) employ small sample tests which find no evidence against the ET for the British Pound and which generally find evidence more in favour of the theory than under asymptotic inference. Bekaert, Hodrick, and Marshall (2001) demonstrate that the introduction of only small time variation of term premia in their regime-switching model is enough mostly to match the regression-based evidence regarding the ET for US Treasury data. In summary, there are many fixed income markets for which the ET appears to hold, at least to a close approximation (the various non-US markets referenced above and the US market for repo rates – see Longstaff 2000b) . For other markets, the time variation in term premia will often be small enough for our analysis to provide important insight into the determinants of the stationarity or 'stability' properties of the term structure.

We work with the logarithmic version of the ET for the following reasons: first, unlike the non-logarithmic version, statements of the theory in terms of multi-period holding returns, one-period holding returns, or forward rates are equivalent (see Eq.'s 2, 3 and 4); and second, the overwhelming majority of empirical evaluations of the ET consider the logarithmic version. Both McCulloch (1993) and Bekaert and Hodrick (2001) establish that the logarithmic ET is consistent with the absence of arbitrage. Furthermore, Longstaff (2000a) generalises the Cox, Ingersoll, and Ross (1981) framework to markets where bonds are not redundant securities and shows that all traditional forms of the ET can be consistent with the absence of arbitrage if the market is not overspanned (or 'complete' in that sense).

The structure of the paper is as follows. Section 2 defines the logarithmic ET, introduces some of the main ideas of the paper using three concrete examples, and derives some implications of the existence of a limiting yield and limiting term premium under the ET. Section 3 is concerned with the construction of the ET-VAR; the proof of the *a.s.* convergence to zero of the distance between

ET-consistent yield curve processes and the ET-VAR; and the derivation of the integration and cointegration properties of the ET-VAR using its moving average (MA) representation. Section 4 contains our main theorems on the connections between the order of integration of the yield curve and the convergence properties of the sequence of innovations to yields, $\{\nu_t(n)\}$, as maturity $n \to \infty$. Section 5 concludes the paper with a discussion of the implications of our results.

The following notation is used throughout. If α is an $n \times r$ matrix of full rank, we define $\alpha_\perp$ to be some $n \times (n-r)$ matrix of full rank such that $\alpha'\alpha_\perp = 0$. We denote by $\gamma[i]$, $i = 1, ..., n$, the ith row of any $n \times r$ matrix γ, and denote by $\gamma[i][j]$ the jth element of that row. The usual notation $||\cdot||$ is used for the Euclidean norm of a vector, and $||\cdot||_\infty$ for the uniform norm of a vector (i.e. the maximum of the absolute values of its elements). The abbreviation $a.s.$ is used when a property holds almost surely, that is with probability one.

2 Yield Curve Dynamics

A zero-coupon or 'discount' bond with face value \$1 and *maturity* τ is a security that makes only a certain payment of \$1 τ periods from today. Its *yield* (to maturity), $y_t(\tau)$, is defined as the per period continuously compounded return obtained by holding the bond from time t to $t+\tau$, so that

$$y_t(\tau) = -\tau^{-1} p_t(\tau), \tag{1}$$

where $p_t(\tau)$ is the log price of the discount bond at t. The *yield curve* consists of the yields on discount bonds of different maturities. Discrete time is indexed by $t \in \{1, 2, ...\}$, and bond maturity τ and t are taken to be measured in the same physical units. We focus in this paper on *complete* yield curves of cross-sectional dimension n, which contain yields for all maturities $\tau \in \{1, 2, ..., n\}$. Formally, an n-complete yield curve is the vector $y_t(1:n) := (y_t(1), y_t(2), ..., y_t(n))'$. The notation $s_t(\tau_2, \tau_1) := y_t(\tau_2) - y_t(\tau_1)$ is used for the *spread* between two yields, whilst the $(n-1) \times 1$ vector of spreads between the yields and the short rate is denoted $s_{nt} := (s_t(2, 1), ..., s_t(n, 1))'$.

2.1 The logarithmic Expectations Theory

The logarithmic expectations theory states that a longer term, τ-period yield differs only by a time-invariant constant from the conditionally expected, per period log return obtained by successively rolling over 1-period discount bonds for τ periods. A formal definition is as follows.

Definition 1 *The discrete time process for yields $\{y_t(\cdot)\}$ satisfies the **logarithmic Expectations Theory (ET)** if and only if*

$$y_t(\tau) = \left\{ \tau^{-1} \sum_{r=0}^{\tau-1} \mathsf{E}[y_{t+r}(1)|\mathcal{F}_t] \right\} + \rho(\tau), \quad \tau = 2, 3, ..., \ \forall t, \tag{2}$$

4

where the real-valued, time-invariant constants $\rho(\tau)$ are known as term premia and $\{\mathcal{F}_t\}$ denotes the filtration of publicly available information, which includes the natural filtration of all yields, $(y_t(\tau); \tau = 1, 2, ...)$.

Note that the above definition is equivalent to the following forms of the logarithmic ET:

$$y_t(1) = \mathsf{E}[r_{t+1}(\tau)|\mathcal{F}_t] + (\tau - 1)\rho(\tau - 1) - (\tau)\rho(\tau), \quad \tau = 2, 3, ..., \forall t, \tag{3}$$

where $r_{t+1}(\tau)$ is the 1-period log holding return obtained by purchasing the τ-maturity bond at time t and selling it at $(t + 1)$; and

$$f_t(\tau) = \mathsf{E}[y_{t+\tau}(1)|\mathcal{F}_t] + (\tau + 1)\rho(\tau + 1) - (\tau)\rho(\tau), \quad \tau = 1, 2, ..., \forall t, \tag{4}$$

where $f_t(\tau)$ is the τ-period ahead *forward rate,* i.e. the guaranteed, continuously compounded, one-period interest rate on a \$1 investment to be made at $(t + \tau)$. Such equivalences do not hold for the non-logarithmic version of the expectations theory as a result of Jensen's inequality terms (see Campbell, Lo, and MacKinlay 1997, p.414).

2.2 Illustrative examples

In order to introduce some of the main ideas of the paper, we provide below three concrete examples which illustrate how the convergence behaviour of the sequence of innovations to yields, $\{\nu_t(\tau)\}_{\tau=1,2,...}$, determinines the order of integration of the yield curve. By the *innovation* to a yield we mean the difference between the yield and its conditional expectation with respect to the public information set $\mathcal{F}_t$, that is the 'shock' $\nu_{t+1}(\tau) := y_{t+1}(\tau) - \mathsf{E}[y_{t+1}(\tau)|\mathcal{F}_t]$.

Example 1 *I(2) yield curve.* *Denote by $F(u, T)$ the instantaneous forward rate at continuous time u on a riskless loan with investment date T. Let the forward rate process follow the Heath-Jarrow-Morton (1992) type specification*

$$F(u, T) - F(0, T) = \int_0^u \sigma(T - s)dB(s), \quad u \in [0, T], \tag{5}$$

where $B(u)$ is a standard Brownian motion, and let the spot volatility $\sigma(T - u) = [T - u]$. This example satifies the ET (Definition 1) with zero term premia (at any time series frequency), since $\mathsf{E}[F(T, T)|\mathcal{F}_u] = F(u, T)$. It is possible to show that

$$F(u, T) - F(0, T) = [T - u]B(u) + \int_0^u B(s)ds,$$

see Øksendal (2000, Theorem 4.1.5); thus with a flat initial forward rate curve $F(0, T) = F(0) \ \forall T$, the yield curve at time u is given by

$$y(u, \tau_c) = \tau_c^{-1} \int_u^{u+\tau_c} F(u, T)dT = F(0) + \frac{\tau_c}{2}B(u) + \int_0^u B(s)ds, \tag{6}$$

5

where τ_c denotes maturity measured in continuous time. It is clear that each yield $y(u, \tau_c)$ has a non-zero I(2) component, $\int_0^u B(s)ds$, and a non-zero I(1) component, $0.5B(u)$, and that the yield curve is hence integrated of order 2. Note that $y(u, \tau_c)$ is linear and hence diverges as a function of τ_c.

What drives the I(2) behaviour of the yield curve in the above example? Suppose, without loss of generality, that we generate a discrete time process by sampling the continuous time process using a sampling interval equal to one. Then the innovation to the yield in discrete time is given by $\nu_{t+1}(\tau) = y(t+1, \tau) - \mathsf{E}[y(t+1, \tau)|\mathcal{F}_t] = 0.5\tau[B(t+1) - B(t)] - B(t) + \int_t^{t+1} B(s)ds$. It follows that the sequence of innovations to yields $\nu_t(\tau)$ diverges linearly in τ, as $\tau \to \infty$. It is this divergence that results in the non-zero I(2) component of the yield curve.[2] We will show in Theorem 9 that (under Condition 1 to be stated later) the *a.s. convergence* of $\nu_t(\tau)$ as $\tau \to \infty$ gurantees that the yield curve is at most integrated of order one.

In the second example below, $\nu_t(\tau)$ converges *a.s.* to a limiting random variable which is not equal to zero *a.s.*, and therefore the yield curve is I(1). We will show in Theorem 10 that the *a.s.* convergence of $\nu_t(\tau)$ to *zero* as $\tau \to \infty$ is a necessary (but not sufficient) condition for the yield curve to be stationary.

Example 2 *I(1) Vasicek Model.* *Consider the discrete time version of the Vasicek (1977) model given by Campbell, Lo, and MacKinlay (1997, p.429), with the autoregressive parameter ϕ equal to 1. Then $\Delta y_{t+1}(\tau) = \xi_{t+1}$ for $\tau = 1, ..., n$, where ξ_{t+1} is the scalar innovation to the single factor of the model (with variance σ^2), and $\mathsf{E}[y_{t+r}(1)|\mathcal{F}_t] = y_t(1) \ \forall r$. It follows that such a process for the yield curve satisfies the ET if and only if $s_t(\tau, 1) := y_t(\tau) - y_t(1)$ is a time-invariant constant (see Eq. 2). The term premium $\rho(\tau)$ is then given by the constant spread $s_t(\tau, 1)$.*

Indeed, it is possible to show that in the Vasicek model of Campbell, Lo, and MacKinlay (1997):

$$s_t(\tau, 1) = -\frac{\sigma^2}{2}\left[\beta(\tau - 1) + \tau^{-1}\sum_{j=1}^{\tau-1}(\tau - j)^2\right] \quad for \ \tau = 2, 3, ..., \tag{7}$$

where we have used the same notation for their (price of risk) parameter β. It follows that $s_t(\tau + 1, \tau) - s_t(\tau, \tau - 1) = -\sigma^2/3$, and hence that $y_t(\tau) \to -\infty$ as $\tau \to \infty$.

Clearly $\nu_t(\tau) \to \xi_t$ *a.s.* in Example 2. It turns out that a necessary and sufficient condition for the yield curve to be stationary is that the innovations to the log price of the discount bonds converge *a.s.* to a real-valued random variable (r.v.), or equivalently that $\tau\nu_t(\tau)$ converges to such a r.v.

---

[2] We note for completeness that since the divergence of $\nu_t(\tau)$ is linear in τ, it follows from Eq. (57) that the limit of $\alpha_{n\perp}^{ET\prime}\nu_{n,t}$ is non-zero and hence that the I(2) component is non-zero.

(which is allowed to depend on t). This condition fails in the Vasicek model above, since there $|\tau \xi_t| \to \infty$, but is satisfied in the third and final example below. Note that if $\tau \nu_t(\tau)$ converges $a.s.$ then it must be the case that $\nu_t(\tau) \to 0$ $a.s.$ as $\tau \to \infty$.

Example 3 *I(0) yield curve.* *Let the forward rate process follow the Heath-Jarrow-Morton type specification in Eq. (5) with $\sigma(T-u) = \exp(\lambda[T-u])$ and $\lambda < 0$. Again, this example satifies the ET (Definition 1) with zero term premia (at any time series frequency), since $\mathsf{E}[F(T,T)|\mathcal{F}_u] = F(u,T)$. Note that by Itô's Lemma, $X(u) := [\sigma(T-u)]^{-1}[F(u,T)-F(0,T)]$ is an Ornstein-Uhlenbeck process that does not depend on T. It is therefore possible to show (again using a sampling interval equal to 1 for convenience) that*

$$\tau \Delta y_{t+1}(\tau) = \xi_{t+1}[1 + e^{\lambda} + ... + e^{(\tau-1)\lambda}] + [e^{\tau\lambda} - 1]X(t),$$

where ξ_{t+1} is the Gaussian innovation to $X(t)$. Hence $\lim_{\tau\to\infty} \nu_t(\tau) = 0$ and $\lim_{\tau\to\infty} \tau\nu_t(\tau) = -\xi_{t+1}e^{(1-\lambda)}$.

Note that all 3 examples above satisfy the condition

$$\lim_{n\to\infty} [s_t(n+1,n) - s_t(n,n-1)] - [s^\rho(n+1,n) - s^\rho(n,n-1)] = 0 \quad a.s., \forall t, \tag{8}$$

where the spread between term premia $s^\rho(n+1,n) := \rho(n+1) - \rho(n)$. Example 1 satisfies the condition because term premia are zero and its yield curve is linear in maturity (see Eq. 6), and Example 2 satisfies it trivially because $s_t(n+1,n) = s^\rho(n+1,n)$ $\forall n$. These are specialised examples chosen for their tractability. The condition in Eq. (8) will be required for the validity of the ET-VAR asymptotic representation method discussed in Section 3.3. A formal statement of the condition and discussion of its role here as a weak regularity condition is given later in Section 3.2. It suffices to note at this stage that a particular case in which the condition clearly holds is when the zero-coupon yields converge $a.s.$ to a finite limiting yield denoted by $y_{t,L} := \lim_{\tau\to\infty} y_t(\tau)$, $\forall t$, and the term premia converge to $\rho_L := \lim_{\tau\to\infty} \rho(\tau)$ (as in Example 3 above). This case is considered in the following section in order both to build intuition for our main results and to provide a link with the well-known Dybvig, Ingersoll, and Ross (1996) theorem on the monotonicity of limiting rates. We note in passing, however, that Eq. (8) clearly includes cases where the yield curve is $a.s.$ unbounded as a function of maturity (as in Examples 1 and 2).

2.3 Long rates and the ET

Many well-known term structure models possess both limiting forward rates and yields. If a limiting forward rate, $f_{t,L} := \lim_{\tau\to\infty} f_t(\tau)$, exists $a.s.$ then $y_{t,L} = f_{t,L}$, although a limiting forward rate

is not necessary for the existence of a limiting yield. In the absence of seasonal effects, a limiting forward rate is an intuitively appealing condition on economic grounds. If a limiting yield exists and the conditionally expected average over time of short rates, $\mathsf{E}[\tau^{-1}\sum_{r=0}^{\tau-1} y_{t+r}(1)|\mathcal{F}_t]$, converges *a.s.* as the time horizon $\tau \to \infty$, then $[y_t(\tau) - \rho(\tau)]$ must converge *a.s.* under the ET (see Eq. 2), and hence there must also exist a limiting term premium.

What is the dynamic behaviour of a limiting zero-coupon yield under the ET? In Theorem 1 below, we show that if a term structure model satisfies the ET the limiting yield must be a martingale. Furthermore, the innovations to yields $\nu_t(\tau)$ converge *a.s.* as $\tau \to \infty$ to the change in the limiting rate, $y_{t,L} - y_{t-1,L}$. This provides intuition for the result stated later in Theorem 9 that the *a.s.* convergence of $\nu_t(\tau)$ implies that the yield curve is at most integrated of order one: in the presence of a limiting yield, long-maturity rates behave like a martingale, which will be I(1) (when the associated martingale difference sequence is stationary). Furthermore, if the limiting yield $y_{t,L}$ is time-invariant, then it follows immediately that $\nu_t(\tau) \to 0$ *a.s.* (In this case $y_{t,L}$ is of course still a martingale but has innovations equal to zero *a.s.*)

Theorem 1 *Suppose that the logarithmic ET holds (Definition 1) and that for each t there exists a.s. a finite limiting zero-coupon yield denoted by* $y_{t,L} := \lim_{\tau \to \infty} y_t(\tau)$. *Suppose also that the term premia converge to* $\rho_L := \lim_{\tau \to \infty} \rho(\tau)$ *and that* $|y_t(\tau)| < Y_t$ *for all t and* τ, *where the* Y_t *are integrable random variables.*

Then the limiting yield process is an $\mathcal{F}_t$-*martingale and hence* $\mathsf{E}[y_{t+1,L}|\mathcal{F}_t] = y_{t,L}$. *Furthermore, denoting the limit of the innovations to yields by* $\nu_{t,L} := \lim_{\tau \to \infty} \nu_t(\tau)$,

$$\nu_{t+1,L} = y_{t+1,L} - y_{t,L} \quad a.s., \tag{9}$$

that is the innovation to the limiting yield is equal to the limit of the innovations to yields, which gurantees the a.s. existence of the latter here.

Combining the Dybvig, Ingersoll, and Ross (1996) theorem with the martingale property of the long zero-coupon yield then shows that if there is also no arbitrage, $y_{t,L}$ *must be time invariant, that is* $y_{t+1,L} = y_{t,L}$ *a.s. for all t.*

The well-known result of Dybvig, Ingersoll, and Ross (1996, Theorem 2) states that when there is an absence of arbitrage, the probability of the limiting yield decreasing over time is zero. We are therefore able to prove that if a model satisfies both the logarithmic ET and an absence of arbitrage (and Y_t, some r.v. that bounds the yield curve, has finite expectation) then its limiting yield must be constant over time and the innovations to yields converge to zero with growing maturity. This is an interesting connection with the work of Dybvig, Ingersoll, and Ross (1996, Theorem 2) given their

view that "we should think of the ordinary situation as one in which the long [...] rate is constant." Theorem 10 will show a related result, namely that under the logarithmic ET the limiting yield *must* be constant if yields are stationary, irrespective of whether arbitrage opportunities exist or not.

Of course, limiting yields are not empirically observable since real-world bond markets possess maximum bond maturities. Nevertheless, just as zero-coupon yields themselves are commonly inferred from coupon bond data, the time series behaviour of the limiting yields implied by commonly used parametric yield curve estimation procedures may be examined to assess how reasonable the assumption of a time-invariant limiting yield might be for a given bond market. Such assessments are less reliable when identification of the limiting yield involves a large degree of extrapolation beyond the observed maturities. We take the view that for some, perhaps most fixed-income markets, the assumption of a time-invariant limiting yield seems implausibly strong from an empirical perspective (see, e.g., Cairns 2004 for an examination of the data for UK gilts).

3   Moving Average Representation

We establish below that the ET fully determines the conditional expectation of the $(n-1)$-complete yield curve, $y_{t+1}(1:n-1)$, given any information set that includes the current n-complete yield curve, $y_t(1:n)$. Whilst the proof of this result is straightforward, its use of complete yield curves and its statement in multivariate form paves the way for the derivation of a *linear* VAR representation of the yield curve that is arbitrarily accurate as the cross-sectional dimension $n \to \infty$. We term this VAR the ET-VAR. The corresponding MA representation is then derived in order to investigate the integration and cointegration properties of yield curves under the ET.

3.1   Conditional expectations under the ET

First note that for a given maturity τ,

$$\mathsf{E}[\Delta y_{t+1}(\tau)|\mathcal{F}_t] = \frac{\tau+1}{\tau}\left\{s_t(\tau+1,1) - \rho(\tau+1)\right\} - \left\{s_t(\tau,1) - \rho(\tau)\right\}, \quad \tau = 1,2,\dots \quad (10)$$

The conditional expectation $\mathsf{E}[\Delta y_{t+1}(1:n-1)|\mathcal{F}_t]$ is an affine function of the current spread vector s_{nt}.

Theorem 2 (*Conditional Expectation of Yields*)   *Let $n \geq 2$ and suppose that the discrete time process for yields $\{y_t(\cdot)\}$ satisfies the logarithmic ET (Definition 1). Then,*

$$\Delta y_{t+1}(1:n-1) = \bar{\alpha}_{n-1}^{ET}[\beta_n' y_t(1:n) - \rho_n] + \nu_{n-1,t+1} \quad \forall t, \quad (11)$$

where $\mathsf{E}[\nu_{n-1,t+1}|\mathcal{F}_t] = 0$ and $\rho_n = [\rho(2), ..., \rho(n)]'$. The $(n-1) \times (n-1)$ matrix $\bar{\alpha}_{n-1}^{ET}$ is, for $n > 2$, given by

$$
\bar{\alpha}_{n-1}^{ET} = \begin{pmatrix}
2 & 0 & 0 & 0 & \dots & 0 & 0 \\
-1 & 3/2 & 0 & 0 & \dots & 0 & 0 \\
0 & -1 & 4/3 & 0 & \dots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \dots & -1 & \frac{n}{n-1}
\end{pmatrix}, \tag{12}
$$

$\bar{\alpha}_1^{ET} = 2$, and the τth row of the $(n-1) \times n$ matrix β_n' is $(-1, \mathbf{0}_{1 \times \tau - 1}, 1, \mathbf{0}_{1 \times n - \tau - 1})$. Thus $\beta_n' y_t(1 : n) = s_{nt}$, the vector of spreads.

It follows from the definition of excess returns on a τ period bond realised at $(t + 1)$, denoted here by $rx_{t+1}(\tau)$, that these are constant under the ET and given by:

$$
\mathsf{E}[rx_{t+1}(\tau)|\mathcal{F}_t] = \tau \rho(\tau) - (\tau - 1)\rho(\tau - 1), \quad \tau = 2, 3, ... \tag{13}
$$

In the case of a complete yield curve, the conditional mean $\mathsf{E}[\Delta y_{t+1}(1 : n - 1)|\mathcal{F}_t]$ is thus a known *linear* function of the difference between the spread and term premia vectors.

We note that for $h > 1$, the ET also fully determines the h-step ahead conditional mean of the $(n-1)$-complete yield curve, given any information set that includes the current $(n-1+h)$-complete yield curve, $y_t(1 : n - 1 + h)$.

Theorem 3 (h-Step Ahead Conditional Expectation of Yields) *Suppose that the logarithmic ET (Definition 1) is satisfied. Then, for $\tau = 1, 2, ...,$*

$$
\mathsf{E}[y_{t+h}(\tau)|\mathcal{F}_t] = y_t(\tau) + \frac{\tau + h}{\tau}\{s_t(\tau + h, h) - s^\rho(\tau + h, h)\} - \{s_t(\tau, h) - s^\rho(\tau, h)\}, \quad h = 1, 2, ..., \tag{14}
$$

where we define the difference or spread between term premia $s^\rho(\tau_2, \tau_1) := \rho(\tau_2) - \rho(\tau_1)$. The conditional mean $\mathsf{E}[y_{t+h}(\tau)|\mathcal{F}_t]$ is thus a linear function of only 2 yields, namely $y_t(h)$ and $y_t(\tau+h)$.[3]

Under the ET, two properties of all h-step ahead conditional means $\mathsf{E}[\Delta_h y_{t+h}(1 : n - 1)|\mathcal{F}_t]$ are noteworthy ($h \geq 1$): they do not involve lagged variables dated prior to t; and the only regressor is the spread vector $s_{n-1+h,t}$. In particular, under the ET neither macroeconomic variables nor financial variables other than current spreads are regressors given the information set $\mathcal{F}_t$ (which includes yields of all maturities). Equation (14) may also be used to establish the equivalence of the ET and the 1-step ahead conditional mean given by Eq. (11) – see the proof of the corollary below.

---

[3]Campbell and Shiller (1991) state in their Eq. (2) a result closely related to Eq. (14) here, but do so without proof.

Corollary 4 *The logarithmic ET (Definition 1) holds if and only if*

$$\Delta y_{t+1}(1:n-1) = \bar{\alpha}_{n-1}^{ET}(\beta_n' y_t(1:n) - \rho_n) + \nu_{n-1,t+1} \quad \forall t, \ n \geq 2, \tag{15}$$

where $\mathsf{E}[\nu_{n-1,t+1}|\mathcal{F}_t] = 0$, $\rho_n = (\rho(2), ..., \rho(n))'$ *is a vector of real-valued constants, and the matrices* $\bar{\alpha}_{n-1}^{ET}$ *and* β_n *are defined for* $n \geq 2$ *as in Theorem 2.*

3.2 Construction of the ET-VAR

Suppose that we have a time series of observed yields $\{y_t(1:n)\}$ that satisfies the ET and we seek a VAR representation for $\{y_t(1:n)\}$ that holds asymptotically when n is large. Theorem 2 provides a sufficiently detailed description of the dynamics of a complete yield curve under the ET that its combination with information only about the asymptotic behaviour of yields (and term premia) at long maturities allows the derivation of such a VAR representation.

Let $\{\mathcal{G}_{nt}\}$ be the natural filtration of the n-complete yield curves $\{y_t(1:n)\}$. We know from Theorem 2 that the conditional mean w.r.t. $\mathcal{G}_{nt}$ of the first $(n-1)$ yields is given by $\mathsf{E}[\Delta y_{t+1}(1:n-1)|\mathcal{G}_{nt}] = \bar{\alpha}_{n-1}^{ET}[\beta_n' y_t(1:n) - \rho_n]$. However, it follows from Eq. (10) that the conditional mean of the longest maturity yield w.r.t. $\mathcal{F}_t$ is not $\mathcal{G}_{nt}$-measurable since the spread $s_t(n+1,1)$ is not observed given the information in $\mathcal{G}_{nt}$. Therefore, in general, $\mathsf{E}[\Delta y_{t+1}(n)|\mathcal{G}_{nt}] \neq \mathsf{E}[\Delta y_{t+1}(n)|\mathcal{F}_t]$.

However, suppose that the following condition, already discussed as Eq. (8), holds in the limit as the cross-sectional dimension of the yield curve $n \to \infty$.

Condition 1 *(For asymptotic validity of ET-VAR representation).* *The condition is given by:*

$$\lim_{n \to \infty} \{[s_t(n+1,n) - s_t(n,n-1)] - [s^\rho(n+1,n) - s^\rho(n,n-1)]\} = 0 \quad a.s., \ \forall t, \tag{16}$$

As will be seen in Theorems 5 and 6 below, this condition implies the asymptotic validity of the ET-VAR representation of a discrete time process for yields that satisfies the ET. The condition states that the function $[y_t(\tau) - \rho(\tau)]$ is asymptotically linear as $\tau \to \infty$ for all t, which is clearly the case when both limiting yields and a limiting term premium exist. One could maintain that the existence of such limits is already a weak condition for analysis of the problem at hand. Nevertheless it follows from Eq. (2) that, under the ET, this function is given by the conditionally expected average over time of short rates up to time $(\tau - 1)$, i.e. $[y_t(\tau) - \rho(\tau)] = \mathsf{E}[\tau^{-1} \sum_{r=0}^{\tau-1} y_{t+r}(1)|\mathcal{F}_t]$. Condition 1 thus includes all well-behaved cases where this conditionally expected time average converges $a.s.$ to a (possibly stochastic) limit as the time horizon $\tau \to \infty$, since then $[y_t(\tau) - \rho(\tau)]$ converges $a.s.$ to the same limit. The limit may vary over time and can be interpreted as the long-run, expected average short rate at time t. However, Condition 1 also allows the conditionally expected average

short rate to diverge as the time horizon $\tau \to \infty$ (provided the growth is asymptotically linear in τ). This is the case in Example 1 where as a result the yield curve diverges as a function of maturity.

Condition 1 thus includes the cases of principle interest and constitutes a weak regularity condition that allows us to characterise sharply the conditions determining the integration order of yields, without the need to specify other features of the process or to resort to particular, parametric term structure models.

We now give a formal definition of an ET-VAR process, the motivation for which is explained in Eq. (20) below.

Definition 2 (ET-VAR) *The ET-VAR approximation of the process $\{y_t(1:n)\}$ is the linear VAR(1) process $\{z_{tn}\}$ given by*

$$\Delta z_{t+1,n} = \alpha_n^{ET}[\beta_n' z_{tn} - \rho_n] + \nu_{n,t+1}, \quad n \geq 2, \tag{17}$$

where the initial condition $z_{0,n} = y_0(1:n)$ a.s. holds, and $\nu_{n,t+1} = y_{t+1}(1:n) - \mathsf{E}[y_{t+1}(1:n)|\mathcal{F}_t]$ is the true innovation to the yield curve. The matrix α_n^{ET} is given by

$$\alpha_n^{ET} = \begin{pmatrix} & \bar{\alpha}_{n-1}^{ET} \\ 0_{1\times(n-3)} & \frac{-(n+1)}{n} & \frac{n+2}{n} \end{pmatrix}, \quad n > 2, \tag{18}$$

and $\alpha_2^{ET} = (2,2)'$, with $\bar{\alpha}_{n-1}^{ET}$ as defined in Theorem 2. The real-valued constants $\rho_n = [\rho(2), ..., \rho(n)]'$ are chosen to satisfy Eq. (15), i.e. ρ_n is the true vector of term premia for the process $\{y_t(1:n)\}$. The characteristic polynomial of the ET-VAR is given by $A_n^{ET}(z) := I_n - (I_n + \alpha_n^{ET}\beta_n')z$.

Since the aim is an asymptotic, autoregressive representation of $\{y_t(1:n)\}$, the ET-VAR process $\{z_{tn}\}$ shares its initialisation, and is defined using the true term premia and true innovations of $\{y_t(1:n)\}$. Associated with the ET-VAR is the point predictor of $\Delta y_{t+1}(1:n)$ given by Eq. (17), which we now define formally below.

Definition 3 (ET-VAR Predictor) *The ET-VAR predictor $\boldsymbol{\mu}_n^{ET}$ is an R^n-valued function of $y_t(1:n)$ that is understood as a 1-step ahead predictor of $\Delta y_{t+1}(1:n)$ and is given by*

$$\boldsymbol{\mu}_n^{ET}[\Delta y_{t+1}(1:n)] = \alpha_n^{ET}[\beta_n' y_t(1:n) - \rho_n]. \tag{19}$$

In a slight abuse of notation we denote by $\boldsymbol{\mu}_n^{ET}[\Delta y_{t+1}(\tau)]$ the τth element of $\boldsymbol{\mu}_n^{ET}[\Delta y_{t+1}(1:n)]$.

It is important to note that the first $(n-1)$ elements of $\boldsymbol{\mu}_n^{ET}[\Delta y_{t+1}(1:n)]$ are equal to $\mathsf{E}[\Delta y_{t+1}(1:n-1)|\mathcal{F}_t]$ a.s., and that its point prediction of the nth yield is given by $\frac{n+2}{n}\{s_t(n,1) - \rho(n)\}$

$-\frac{(n+1)}{n}\{s_t(n-1,1) - \rho(n-1)\}$. It follows straightforwardly that

$$||\mathsf{E}[\Delta y_{t+1}(1 \quad : \quad n)|\mathcal{F}_t] - \boldsymbol{\mu}_n^{ET}[\Delta y_{t+1}(1:n)]|| =$$
$$\left|\frac{n+1}{n}\left[\{s_t(n+1,n) - s_t(n,n-1)\} - \{s^\rho(n+1,n) - s^\rho(n,n-1)\}\right]\right|, \quad (20)$$

from which it is clear that the ET-VAR predictor will be close to the true conditional mean when n is large and Condition 1 holds. We state this property formally in the theorem below, which in turn will be central to establishing the asymptotic validity of the ET-VAR representation.

Theorem 5 *Suppose that the discrete time process for yields $\{y_t(\cdot)\}$ satisfies the logarithmic Expectations Theory (see Definition 1), and that Condition 1 is satisfied. Then,*

$$\lim_{n\to\infty} ||\mathsf{E}[\Delta y_{t+1}(1 \quad : \quad n)|\mathcal{F}_t] - \boldsymbol{\mu}_n^{ET}[\Delta y_{t+1}(1:n)]||$$
$$= \quad \lim_{n\to\infty} |\mathsf{E}[\Delta y_{t+1}(n)|\mathcal{F}_t] - \boldsymbol{\mu}_n^{ET}[\Delta y_{t+1}(n)]| \quad (21)$$
$$= \quad 0 \ a.s. \ \forall t = 0, 1, ... \quad (22)$$

Similarly, $\lim_{n\to\infty} ||\mathsf{E}[\Delta y_{t+1}(1:n)|\mathcal{F}_t] - \boldsymbol{\mu}_n^{ET}[\Delta y_{t+1}(1:n)]||_\infty = 0 \ a.s. \ \forall t = 0, 1, ...,$ where the Euclidean norm has been replaced with the uniform norm.

3.3 Asymptotic properties of the ET-VAR

Our aim then is to establish that under Condition 1,

$$\lim_{n\to\infty} ||\mathbf{y}_n^T - \mathbf{z}_n^T||_\infty = 0 \ a.s. \quad \forall T \in \{1, 2, ...\}, \quad (23)$$

We use $\mathbf{x}_n^T$ to denote the column vector formed by vertically stacking the n-dimensional vector elements of a time series $\{x_{tn}\}_{t=0,...,T-1}$. Intuitively Eq. (23) states that, with probability one, the sample path of the ET-consistent yield curve $y_t(1:n)$ and the sample path of its ET-VAR approximation z_{tn} can be made arbitrarily close by setting n sufficiently large. Since we are able to derive the integration and cointegration properties of the ET-VAR process (see Section 3.4), we will be able to conclude that $y_t(1:n)$ must share these properties of z_{tn} in the limit as $n \to \infty$. Note that Eq. (23) is equivalent to the statement that given $\epsilon > 0$, $\exists N(\epsilon)$ such that for all $n > N(\epsilon)$,

$$|y_t(\tau) - z_{tn}(\tau)| < \epsilon \ \forall \tau \in \{1, ..., n\}, \ \forall t = 0, 1, ..., (T-1). \quad (24)$$

This is exactly the property that should be established, namely that at *all* points in time before T and for *all* the maturities of the yield curve, the distance $|y_t(\tau) - z_{tn}(\tau)|$ is uniformly bounded by ϵ.

Let us denote by $w_t(1:n) := y_t(1:n) - z_{tn}$ the error that results from approximating $y_t(1:n)$ by the corresponding 'observation' of the ET-VAR at time t. Then the next time period's error is given by

$$\Delta w_{t+1}(1:n) = \mathsf{E}[\Delta y_{t+1}(1:n)|\mathcal{F}_t] - \boldsymbol{\mu}_n^{ET}[\Delta y_{t+1}(1:n)] - \alpha_n^{ET}\beta_n' w_t(1:n) \quad a.s. \qquad (25)$$

When $t = 1$, $w_t(1:n) = 0$ a.s. due to the initialisation $z_{0n}(1:n) = y_0(1:n)$ a.s., and $||w_{t+1}(1:n)||_\infty$ is equal to the distance between the true conditional mean and the ET-VAR predictor studied in Theorem 5. When $t > 1$, in general $w_t(1:n) \neq 0$ and the contribution of the term $\alpha_n^{ET}\beta_n' w_t(1:n)$ to $\Delta w_{t+1}(1:n)$ must be taken into account. Thus the proof of Theorem 6 below proceeds by induction on t. The triangle inequality is applied to $||\Delta w_{t+1}(1:n)||_\infty$, with Theorem 5 applying to the term $||\mathsf{E}[\Delta y_{t+1}(1:n)|\mathcal{F}_t] - \boldsymbol{\mu}_n^{ET}[\Delta y_{t+1}(1:n)]||_\infty$, whilst the boundedness of the matrix norm $||\alpha_n^{ET}\beta_n'||_\infty$ ensures that $||\alpha_n^{ET}\beta_n' w_t(1:n)||_\infty$ converges to zero. Since the Euclidean norm is perhaps more familiar, we state and prove the result for this norm also.

Theorem 6 *Suppose that the discrete time process for yields $\{y_t(\cdot)\}$ satisfies the logarithmic Expectations Theory (Definition 1), and Condition 1 is satisfied. Let $\{z_{tn}\}$ be the ET-VAR approximation of the process $\{y_t(1:n)\}$ given by Definition 2. Then*

$$\lim_{n\to\infty} ||y_t(1:n) - z_{tn}|| = 0 \ a.s., \quad t = 0, 1, 2, ... \qquad (26)$$

Arrange the n time series as $n \times T$ matrices, and define $\mathbf{y}_n^T := vec\{y_t(1:n); t = 0, 1, 2, ..., (T-1)\}$, with $\mathbf{z}_n^T$ defined analogously. Then,

$$\lim_{n\to\infty} ||\mathbf{y}_n^T - \mathbf{z}_n^T|| = 0 \ a.s. \quad \forall T \in \{1, 2, ...\}. \qquad (27)$$

The same properties hold using the uniform norm. That is, $\lim_{n\to\infty} ||y_t(1:n) - z_{tn}||_\infty = 0$ a.s., $t = 0, 1, 2, ...,$ and $\lim_{n\to\infty} ||\mathbf{y}_n^T - \mathbf{z}_n^T||_\infty = 0$ a.s.

Notice in Eq. (26) that the convergence holds even though the cross-sectional dimension n of the yield curve $y_t(1:n)$ is allowed to equal that of the approximating ET-VAR, and hence allowed to grow asymptotically.

An immediate and important implication of Theorem 6 is that under its conditions, the process for the complete yield curve $y_t(1:n)$ can be described arbitrarily well by an ET-VAR, which is a *linear*, first order vector autoregression. Thus, under Condition 1, non-linear dynamics of (complete) yield curves in which the conditional expectation is a non-linear function of current and past yields are ruled out. The moving average representation of the ET-VAR (see Eq. 28 below) makes clear that z_{tn} is a linear function of the current and past innovations $\nu_{n,t-i}$ ($i = 0, 1, ...$).

3.4 Integration and cointegration

This section derives the integration and cointegration properties of the ET-VAR process z_{tn}, beginning with its moving average representation. We will describe a vector process X_t as integrated of order d, I(d), $d = 0, 1, 2, \ldots$ if it is stationary after differencing d times, i.e. if $\Delta^d(x_t - \mathsf{E}[x_t])$ is stationary, but $\Delta^{(d-1)}(x_t - \mathsf{E}[x_t])$ is not stationary The notation $\chi_n := \bar{\beta}_n \bar{\alpha}_n^{ET\prime}$ is used throughout this section.

Theorem 7 below establishes that the yield curve of an ET-VAR is I(2) and that its spread vector $\beta'_n z_{tn}$ is I(1). We then go on in Theorem 8 to show that the spread vector is itself cointegrated, with cointegrating rank $(n - 2)$ and stationary cointegrating relations given by the *curvatures* of the yield curve. These are the integration and cointegration properties of Example 1, since any discrete time process obtained by sampling the continuous time yield curve there obeys an ET-VAR *exactly* (because the yield curve is linear). However we will see very shortly, in Theorem 9 below, the importance of situations in which the I(2) component of the ET-VAR equals zero in the limit as $n \to \infty$, and hence its associated spread vector is stationary.

Theorem 7 (MA Representation) *Let $\{z_{tn}\}$ be generated according to an ET-VAR (see Eq. 17) with $n \geq 2$ and $\mathsf{E}[\nu_{nt}\nu'_{nt}] = \Omega_n < \infty$.[4] Since the matrices $\alpha_n^{ET}\beta'_n$ and $\alpha_{n\perp}^{ET\prime}\beta_{n\perp}$ have reduced ranks given by $(n-1)$ and zero respectively, and $\det(\alpha_{n\perp}^{ET\prime}\chi_n\beta_{n\perp}) \neq 0$, the process $\{z_{tn}\}$ is I(2) and has the representation*

$$
\begin{aligned}
z_{tn} &= M_{2n}\sum_{u=1}^{t}\sum_{r=1}^{u}\nu_{n,r} + M_{1n}\sum_{r=1}^{t}(\nu_{n,r} - \alpha_n^{ET}\rho_n) \\
&\quad + \sum_{r=1}^{\infty}M_{0n,r}(\nu_{n,t-r} - \alpha_n^{ET}\rho_n) + M_{3n} + M_{4n}t,
\end{aligned}
\tag{28}
$$

where

$$
\begin{aligned}
M_{2n} &= \beta_{n\perp}(\alpha_{n\perp}^{ET\prime}\chi_n\beta_{n\perp})^{-1}\alpha_{n\perp}^{ET\prime} \neq 0, \\
M_{1n} &= \{\chi_n - M_{2n}\chi_n[\chi_n + I_n]\}\,M_{2n} + M_{2n}\chi_n \neq 0,
\end{aligned}
\tag{29}
$$

and the coefficients M_{3n} and M_{4n} depend on the initial conditions with M_{4n} satisfying $\beta'_n M_{4n} = 0$. Note that $\beta'_n M_{2n} = 0$, $\beta'_n M_{1n} = (\alpha_n^{ET\prime}\alpha_n^{ET})^{-1}\alpha_n^{ET\prime}M_{2n} \neq 0$, and that $M_{2n}\alpha_n^{ET}\rho_n = \beta'_n M_{1n}\alpha_n^{ET}\rho_n = 0$.

It follows immediately that $\Delta^2 z_{tn}$ is stationary with mean zero and that the spread vector $\beta'_n z_{tn}$ is I(1).[5]

---

[4]If the term stationary is taken to mean covariance (or 'weakly') stationary then all that is needed here is the constant variance Ω_n, since ν_{nt} is a Martingale Difference Sequence by definition. We impose this henceforth for simplicity.

[5]We note in passing that if, for all t, the distribution of $\nu_{n,t}$ possesses a density with respect to n-dimensional

The ET-VAR (17) implies that the spread vector $s_{nt} = \beta_n' z_{tn}$ follows the VAR process

$$\Delta s_{n,t+1} = \beta_n' \alpha_n^{ET}(s_{nt} - \rho_n) + \beta_n' \nu_{n,t+1}. \tag{30}$$

Theorem 7 has already established that the spread vector is non-stationary and I(1). It follows that the matrix $\beta_n' \alpha_n^{ET}$ must have reduced rank, and it is shown in Lemma 11 of the Appendix that its rank is equal to $(n-2)$, which in turn is equal to the cointegrating rank of the process for the spread vector.

Theorem 8 *(Cointegrated I(1) Spreads)* *Let $\{z_{tn}\}$ be generated according to an ET-VAR (see Eq. 17) with $n > 2$. Then the spread vector s_{nt} is a cointegrated I(1) process,[6] with cointegrating rank $(n-2)$ and stationary cointegrating relations given by the curvatures c_{nt} of the yield curve, where*

$$c_{nt} := (s_t(3,2) - s_t(2,1), ..., s_t(n, n-1) - s_t(n-1, n-2))'. \tag{31}$$

4   Main Theorems

We now use the MA representation of the ET-VAR derived above in order to investigate the determinants of the integration properties of yields under the logarithmic ET. It turns out that with limiting yields and a limiting term premium (or more generally, under the regularity condition 1), these properties depend only on the convergence behaviour of the innovations to yields $\nu_t(\tau)$ as the maturity $\tau \to \infty$, in a manner made specific below.

Example 1 demonstrates the possibility under the ET of I(2) yield curves and the associated stationary curvatures (see Theorem 8). The theorem below establishes that the I(2) component vanishes whenever the innovations to yields $\nu_t(\tau)$ converge *a.s.* to a limiting, real-valued random variable $\nu_{t,L}$ for all t. Note that the r.v. $\nu_{t,L}$ may vary over time, and that with limiting yields and a limiting term premium, $\nu_{t,L} = y_{t,L} - y_{t-1,L}$ *a.s.* (see Eq. 9).

Theorem 9 *Suppose that the logarithmic ET holds (Definition 1), that Condition 1 is satisfied and that the innovations to yields $\nu_t(\tau)$ converge a.s. to the limiting, real-valued random variable $\nu_{t,L}$ as $\tau \to \infty$. By Theorem 6, an ET-VAR z_{tn} may be constructed such that:*

$$\lim_{n \to \infty} ||\mathbf{y}_n^T - \mathbf{z}_n^T||_\infty = 0 \; a.s. \quad \forall T \in \{1, 2, ...\}. \tag{32}$$

Lebesgue measure λ_n, $\mathsf{P}[\nu_{n,t} \in col(\alpha_n^{ET})] = 0 \; \forall t$, since $\lambda_n[col(\alpha_n^{ET})] = 0$. It then follows that the I(2) component is non-zero *a.s.* for all t, since $M_2 \nu_{n,r} = 0$ iff $\nu_{n,r} \in col(\alpha_n^{ET})$. Compare with Example 2 in which $\alpha_{n\perp}^{ET'} \nu_{n,t} = \alpha_{n\perp}^{ET'} \beta_{n\perp} \xi_t = 0 \; \forall n$ since $\alpha_{n\perp}^{ET'} \beta_{n\perp} = 0$.

[6] We note in passing that if $\mathsf{P}[\nu_{n,t} \in col(\alpha_n^{ET})] = 1 \; \forall t$, then $\mathsf{P}[\beta_n' M_{1n} \nu_{n,t} = 0] = 1 \; \forall t$ because $\beta_n' M_1 \nu_{n,t} = (\alpha_n^{ET'} \alpha_n^{ET})^{-1} \alpha_n^{ET'} M_2 \nu_{n,t} = 0$ whenever $\nu_{n,t} \in col(\alpha_n^{ET})$, and s_{nt} is stationary since both its I(2) and I(1) components are equal to zero *a.s.* (as in Example 2, where $\alpha_{n\perp}^{ET'} \nu_{n,t} = \alpha_{n\perp}^{ET'} \beta_{n\perp} \xi_t = 0 \; \forall n$).

Furthermore the I(2) components of the MA representation (Eq. 28) of z_{tn}, $\sum_{u=1}^{t}\sum_{r=1}^{u}(M_{2n}\nu_{n,r})[\tau]$, are identical across maturity τ and converge a.s. to zero as $n \to \infty$. Therefore,

$$\lim_{n\to\infty}\left\|\sum_{u=1}^{t}\sum_{r=1}^{u}M_{2n}\nu_{n,r}\right\|_{\infty} = 0 \quad a.s., \quad \forall t, \tag{33}$$

where $||\cdot||_{\infty}$ denotes the uniform norm.

Suppose that $\mathsf{E}[\nu_{\tau t}\nu_{\tau t}'] = \Omega_{\tau}$ is time-invariant. For any $\tau < \infty$, the vector time series of yields $\mathbf{y}_n^T$ can therefore, by choosing n sufficiently large, be represented arbitrarily accurately by a MA representation that is asymptotically at most integrated of order 1.

Note that in effect the I(2) component vanishes from *all* yields in the economy. Claims of the sort just made that, "the vector time series of yields $\mathbf{y}_n^T$ can [...] be represented arbitrarily accurately by a MA representation that is asymptotically at most integrated of order 1" merit a little more explanation. As n increases, both $||\mathbf{y}^T(1:n) - \mathbf{z}_n^T(1:n)||_{\infty}$ and the I(2) component of the sample path of the ET-VAR process become ever closer to zero. The distance of both objects from zero is made arbitrarily small by choosing n to be sufficiently large. Since in the limit the ET-VAR is at most integrated of order one and 'indistinguishable' from $\mathbf{y}_n^T$, it is appropriate to regard the latter as I(d), $d \le 1$ for large n.[7]

Theorem 9 implies that in a benchmark economy with a limiting yield and limiting term premium (which of course satisfies Condition 1), the I(2) component vanishes from the yield curve. In such an economy the yield curve is either I(1) or stationary, and this in turn depends on the rate of convergence of $\nu_t(\tau)$.

Theorem 10 *Suppose that the logarithmic ET holds (Definition 1) and that Condition 1 is satisfied. As in Theorem 9, an ET-VAR z_{tn} may be constructed such that:*

$$\lim_{n\to\infty}||\mathbf{y}_n^T - \mathbf{z}_n^T||_{\infty} = 0 \; a.s. \quad \forall T \in \{1, 2, ...\}.$$

A necessary and sufficient condition is sought for the following to hold:

$$\lim_{n\to\infty}\left\|\sum_{r=1}^{t}M_{1n}\nu_{n,r}\right\|_{\infty} = 0 \quad a.s., \forall t. \tag{34}$$

Such a necessary and sufficient condition is that $-\tau\nu_{t+1}(\tau)$, *the innovation to the log price of the τ-maturity bond, converges a.s. to some real-valued random variable $\nu_{t+1,p}$ as $\tau \to \infty$, that is*

$$p_{t+1}(\tau) - \mathsf{E}[p_{t+1}(\tau)|\mathcal{F}_t] = -\tau\nu_{t+1}(\tau) \to \nu_{t+1,p} \quad a.s., \forall t. \tag{35}$$

---

[7]Of course the I(2) component may remain non-zero $\forall n$ despite becoming arbitrarily small, but this has no consequence in this context.

Suppose also that $\mathsf{E}[\nu_{\tau t}\nu'_{\tau t}] = \Omega_\tau$ *is time-invariant. Then the implication of Eq. (35) is that the vector time series of yields* $\mathbf{y}_n^T$ *can be represented arbitrarily accurately by a MA representation that is asymptotically stationary.*

Unlike the I(2) component, the I(1) component of the ET-VAR varies across maturity. However, the use of the uniform norm in Eq. (34) ensures that, analogously to Eq. (24), the I(1) components $\left|\sum_{r=1}^t (M_{1n}\nu_{n,r})[\tau]\right|$ can simultaneously be made arbitrarily small across all maturities (and $\forall t < T$).

The I(1) component vanishes from all yields in the economy and yields are stationary if and only if $\tau\nu_t(\tau)$ converges $a.s.$ to a real-valued, possibly time-varying random variable as $\tau \to \infty$. This implies (but is not implied by) the convergence of $\nu_t(\tau)$ to zero $a.s.$ Recall from Theorem 1 that in a benchmark economy with a limiting yield and limiting term premium, $\nu_t(\tau) \to 0$ $a.s.$ is also a necessary condition for the absence of arbitrage, and that this condition implies the time-invariance of the limiting yield.

We have shown that the convergence properties of the innovations to zero-coupon yields and log discount bond prices at the long maturity end of the term structure determine the order of integration of yields. One way to think about these results is that the more 'regular' is the behaviour across maturities and time of the innovations to the long end of the term structure, the more 'stable' is the time series evolution of the yield curve. If the 'surprise in' or 'shock to' yields is effectively constant across maturities (at $\nu_{t,L}$) for long yields at every time t, then I(2) yields are ruled out, even when $\nu_{t,L}$ is both stochastic and time-varying. However, the stationarity of yields requires that there effectively be no surprise in long yields for any time t ($\nu_{t,L} = 0$ $\forall t$). Of course, log discount bond prices diverge to $-\infty$ with increasing maturity since the discount function must converge to zero. But when the surprise in long, log discount bond prices is effectively constant across maturities (at $\nu_{t,p}$) at each time t, then yields are stationary.

5 Discussion and Possible Extensions

Theorems 9 and 10 imply three impossibility results in a benchmark economy that possesses a limiting yield and term premium, and that satisfies the logarithmic expectations theory. First, a non-linear autoregressive specification of the dynamics of complete, high-dimensional yield curves is inconsistent with the ET. Second, a stationary yield curve is impossible when the limiting yield varies over time. Third, an I(2) yield curve is impossible, irrespective of how the limiting yield behaves over time.

An appreciation of these impossibilities can be brought to bear on empirical work concerned with modelling the yield curve and evaluating the expectations theory. For some, perhaps most fixed

income markets, the time invariance of limiting yields does not square well with an examination of the data. One should then, as a result of the second and third impossibilities, regard yields as I(1) under the null of the ET. Our results thus give reassuring, theoretical support to a substantial econometric literature that evaluates the ET using procedures that are only valid when yields are indeed I(1).

It is widely recognised however that an I(1) process is a poor long-run description of the behaviour of yields since such a process is unbounded with probability one as time $t \to \infty$. Unless the setting is one where a time-invariant limiting yield seems plausible empirically, this highlights a limitation of the ET – stationary, mean-reverting yields arise only under a very restrictive condition. A response in the literature to the empirically observed, (near-)integrated behaviour of yields and the desirability of (eventual) mean-reversion has been the use of non-linear autoregressive models (see Lanne and Saikkonen 2002, and Nicolau 2002). However, as a result of the first impossibility above, such models are very likely ruled out by the ET of Definition 1. Taken together these points suggest that, just as I(1) processes for the yield curve are best regarded as local approximations, so too one should not expect the ET literally to hold for all time, but rather over finite time intervals or regimes. Simple models might allow variation of term premia between but not within such regimes, or allow deviation from the ET within certain regimes but not others.

We conclude by briefly mentioning two possible extensions to this work. The aim has been to provide parsimonious, economically interpretable conditions determining the order of integration of the yield curve and the stationarity properties of the term structure without resorting to particular parametric models. It would be interesting to extend our results on the importance of the convergence behaviour of innovations to yields and log bond prices to settings where stationary, moderate time-variation in term premia is allowed. Finally, our analysis has been in discrete time owing to the relative sparsity of results on integration and cointegration in continuous time (although see Comte 1999). However, the extension to continuous time and to processes satisfying the local expectations theory would enable an analysis of the problem for particular classes such as continuous time, affine term structure models. Conditions established under the risk neutral measure might then imply economically interpretable conditions under the data generating measure.

19

Any remaining errors are our own. All computations were performed using the Ox language of Doornik (2001).

References

Bekaert, G. and R. Hodrick (2001). Expectations hypothesis tests. *Journal of Finance 56*, 1357–1394.

Bekaert, G., R. Hodrick, and D. Marshall (2001). Peso problem explanations for term structure anomalies. *Journal of Monetary Economics 48*, 241–270.

Billingsley, P. (1995). *Probability and Measure.* New York: Wiley.

Cairns, A. J. (2004). *Interest Rate Models: An Introduction.* Princeton: Princeton University Press.

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

Campbell, J. Y. and R. J. Shiller (1987). Cointegration and tests of present value models. *Journal of Political Economy 95*, 1062–1088.

Campbell, J. Y. and R. J. Shiller (1991). Yield spreads and interest rate movements: A bird's eye view. *Review of Economic Studies 58*, 495–514.

Cochrane, J. H. and M. Piazzesi (2005). Bond risk premia. *American Economic Review 95*, 138–160.

Comte, F. (1999). Discrete and continuous time cointegration. *Journal of Econometrics 88*, 207–226.

Cox, J., J. Ingersoll, and S. Ross (1981). A reexamination of traditional hypotheses about the term structure of interest rates. *Journal of Finance 36*, 321–346.

Dahlquist, M. and G. Jonsson (1995). The information in Swedish short-maturity forward rates. *European Economic Review 39*, 1115–1131.

Davidson, J. (2002). *Stochastic Limit Theory.* Oxford: Oxford University Press.

Doornik, J. A. (2001). *Ox 3.0 - An Object-Oriented Matrix Programming Language.* London: Timberlake Consultants Ltd.

Dybvig, P. H., J. E. Ingersoll, and S. A. Ross (1996). Long forward and zero-coupon rates can never fall. *Journal of Business 69*, 1–25.

Fama, E. F. and R. R. Bliss (1987). The information in long-maturity forward rates. *American Economic Review 77*, 680–692.

Fisher, I. (1930). *The Theory of Interest*. New York: Macmillan Press.

Gerlach, S. and F. Smets (1997). The term structure of Euro-rates: some evidence in support of the expectations hypothesis. *Journal of International Money and Finance 16*, 305–321.

Hardouvelis, G. A. (1994). The term structure spread and future changes in long and short rates in the G7 countries. Is there a puzzle? *Journal of Monetary Economics 33*, 255–283.

Heath, D., R. Jarrow, and A. Morton (1992). Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica 60*, 77–105.

Hicks, J. (1953). *Value and Capital*. London: Oxford University Press.

Hubalek, F., I. Klein, and J. Teichmann (2002). A general proof of the Dybvig-Ingersoll-Ross theorem: Long forward rates can never fall. *Mathematical Finance 12*, 447–451.

Johansen, S. (1996). *Likelihood-based inference in cointegrated vector autoregressive models*. Oxford: Oxford University Press.

Johansen, S. (2008). Representation of cointegrated autoregressive processes with application to fractional processes. Forthcoming, Econometric Reviews.

Johansen, S., K. Juselius, R. Frydman, and M. Goldberg (2008). Testing hypotheses in an I(2) model with applications to the persistent long swings in Dmk-Usd rate. Economics Discussion Paper 07-34, University of Copenhagen.

Keynes, J. M. (1930). *A Treatise on Money*. London: Macmillan Press.

Lanne, M. and P. Saikkonen (2002). Threshold autoregressions for strongly autocorrelated time series. *Journal of Business and Economic Statistics 20*, 282–289.

Longstaff, F. (2000a). Arbitrage and the expectations hypothesis. *Journal of Finance 55*, 989–994.

Longstaff, F. (2000b). The term structure of very short-term rates: New evidence for the expectations hypothesis. *Journal of Financial Economics 58*, 397–415.

McCulloch, J. H. (1993). A reexamination of traditional hypotheses about the term structure of interest rates: a comment. *Journal of Finance 48*, 779–789.

Nicolau, J. (2002). Stationary processes that look like random walks: the bounded random walk process in discrete and continuous time. *Econometric Theory 18*, 99–118.

Øksendal, B. (2000). *Stochastic Differential Equations: An Introduction with Applications.* Berlin: Springer.

Roush, J. (2007). The expectations theory works for monetary policy shocks. *Journal of Monetary Economics 54*, 1631–1643.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics 5*, 177–188.

Wu, T. (2006). Macro factors and the affine term structure of interest rates. *Journal of Money, Credit and Banking 38*, 1847–1875.

APPENDIX

Proof. (Theorem 1) It follows from Eq. (13) that $\lim_{\tau \to \infty} \mathsf{E}[\tau^{-1} rx_{t+1}(\tau)|\mathcal{F}_t] = \rho_L - \rho_L = 0$ $\forall t$. By definition, $rx_{t+1}(\tau) = \tau y_t(\tau) - (\tau - 1)y_{t+1}(\tau - 1) - y_t(1)$ and hence

$$
\begin{aligned}
\lim_{\tau \to \infty} \mathsf{E}[\tau^{-1} rx_{t+1}(\tau)|\mathcal{F}_t] &= y_{t,L} - \lim_{\tau \to \infty} \mathsf{E}[y_{t+1}(\tau - 1)|\mathcal{F}_t] \quad (36)\\
&= y_{t,L} - \mathsf{E}[y_{t+1,L}|\mathcal{F}_t] \quad a.s.,
\end{aligned}
$$

where the second equality follows from the integrability of Y_{t+1} (see, e.g., Theorem 34.2(v) of Billingsley 1995). Since we have shown that the l.h.s. of Eq. (36) must be zero under the ET, it follows that $\mathsf{E}[y_{t+1,L}|\mathcal{F}_t] = y_{t,L}$ $a.s.$ and that $\{y_{t,L}\}$ is an $\mathcal{F}_t$-martingale. Then $\lim_{\tau \to \infty} \mathsf{E}[\Delta y_{t+1}(\tau)|\mathcal{F}_t] = 0$ $a.s.$, and hence

$$
\nu_{t+1,L} = \lim_{\tau \to \infty} \{\Delta y_{t+1}(\tau) - \mathsf{E}[\Delta y_{t+1}(\tau)|\mathcal{F}_t]\} = y_{t+1,L} - y_{t,L} \quad a.s.,
$$

which is an $\mathcal{F}_t$-martingale difference sequence (MDS).

The Dybvig-Ingersoll-Ross theorem states that when there is no arbitrage, $\nu_{t,L} = y_{t,L} - y_{t-1,L} \geq 0$ $a.s.$ (see Hubalek, Klein, and Teichmann 2002, Theorem 3.1 for a general proof). Since $\nu_{t,L}$ is an $\mathcal{F}_t$-MDS, $\mathsf{E}[\nu_{t,L}] = 0$ and hence $y_{t,L} = y_{t-1,L}$ $a.s.$ ∎

Proof. (Theorem 2) Eq. (2) implies that

$$
\mathsf{E}[y_{t+1}(\tau - 1)|\mathcal{F}_t] = \left\{ \frac{1}{\tau - 1} \sum_{i=1}^{\tau-1} \mathsf{E}[y_{t+i}(1)|\mathcal{F}_t] \right\} + \rho(\tau - 1), \quad \tau = 2, 3, ..., \text{ and} \quad (37)
$$

$$
\tau y_t(\tau) - y_t(1) = \left\{ \sum_{i=1}^{\tau-1} \mathsf{E}[y_{t+i}(1)|\mathcal{F}_t] \right\} + \tau \rho(\tau), \quad \tau = 2, 3, ... \quad (38)
$$

Combining (37) and (38) gives

$$
\mathsf{E}[\Delta y_{t+1}(\tau)|\mathcal{F}_t] = \frac{\tau + 1}{\tau} \{s_t(\tau + 1, 1) - \rho(\tau + 1)\} - \{s_t(\tau, 1) - \rho(\tau)\}, \quad \tau = 1, 2, ...,
$$

22

which, after taking conditional expectations w.r.t. $\mathcal{F}_t$, is (11) stated equation-by-equation. ∎

Proof. (Theorem 3) Eq. (14) clearly holds for $h = 1$. The proof is by induction on h. Suppose that Eq. (14) holds for some $h \geq 1$. Then

$$\mathsf{E}[\Delta_{h+1}y_{t+h+1}(\tau)|\mathcal{F}_t] = \mathsf{E}[\Delta_1 y_{t+1}(\tau)|\mathcal{F}_t] + \mathsf{E}[\mathsf{E}[\Delta_h y_{(t+1)+h}(\tau)|\mathcal{F}_{t+1}]|\mathcal{F}_t],$$

where the first term on the right follows from Eq. (11) and the inner conditional expectation of the second term on the right follows from the induction hypothesis. Then, noting that, for $\tau = 1, 2, ...$,

$$
\begin{aligned}
\mathsf{E}[s_{t+1}(\tau, 1)|\mathcal{F}_t] &= \mathsf{E}[y_{t+1}(\tau) - y_{t+1}(1))|\mathcal{F}_t] \\
&= \frac{\tau + 1}{\tau}\left\{s_t(\tau + 1, 1) - \rho(\tau + 1)\right\} + \rho(\tau) - 2\left\{s_t(2, 1) - \rho(2)\right\},
\end{aligned}
$$

we obtain

$$
\begin{aligned}
\mathsf{E}[\Delta_{h+1}y_{t+h+1}(\tau)|\mathcal{F}_t] &= \mathsf{E}[\Delta y_{t+1}(\tau)|\mathcal{F}_t] + \frac{\tau + h}{\tau}\left\{\mathsf{E}[s_{t+1}(\tau + h, 1)|\mathcal{F}_t] - s^{\rho}(\tau + h, h)\right\} \\
&\quad - \mathsf{E}[h\tau^{-1}s_{t+1}(h, 1) + s_{t+1}(\tau, 1)|\mathcal{F}_t] + s^{\rho}(\tau, h) \\
&= \frac{\tau + h + 1}{\tau}\left\{s_t(\tau + h + 1, h + 1) - s^{\rho}(\tau + h + 1, h + 1)\right\} \\
&\quad - \left\{s_t(\tau, h + 1) - s^{\rho}(\tau, h + 1)\right\},
\end{aligned}
$$

as required to complete the proof by induction. ∎

Proof. (Corollary 4) The necessity of the condition for the ET to hold has been established by Theorem 2. Its sufficiency may be established as follows. First note from the proof of Theorem 3 that Eq. (14) is a direct implication of Eq. (15). Hence Eq. (15) implies that, for all $\tau \geq 2$,

$$
\begin{aligned}
\tau^{-1}\sum_{r=0}^{\tau-1}\mathsf{E}[y_{t+r}(1)|\mathcal{F}_t] &= y_t(1) + \tau^{-1}\sum_{r=1}^{\tau-1}\left[\frac{r+1}{1}\left\{s_t(1 + r, r) - s^{\rho}(1 + r, r)\right\} - \left\{s_t(1, r) - s^{\rho}(1, r)\right\}\right] \\
&= y_t(\tau) - \rho(\tau),
\end{aligned}
$$

which, on comparison with Definition 1, completes the proof. ∎

Proof. (Theorem 5) Recall the fundamental property of the ET-VAR predictor, namely that

$$\mathsf{E}[\Delta y_{t+1}(1 : n - 1)|\mathcal{F}_t] = \boldsymbol{\mu}_n^{ET}[\Delta y_{t+1}(1 : n - 1)] \ a.s.,$$

which yields the first equality in (22). Since by Definition 3 $\boldsymbol{\mu}_n^{ET}[\Delta y_{t+1}(1 : n)] = \alpha_n^{ET}[\beta_n' y_t(1 : n) - \rho_n]$, it follows directly that

$$\boldsymbol{\mu}_n^{ET}[\Delta y_{t+1}(n)] = \frac{n+2}{n}\left\{s_t(n, 1) - \rho(n)\right\} - \frac{(n+1)}{n}\left\{s_t(n-1, 1) - \rho(n-1)\right\}. \tag{39}$$

Applying Theorem 2 to obtain $\mathsf{E}[\Delta y_{t+1}(n)|\mathcal{F}_t]$ gives

$$\mathsf{E}[\Delta y_{t+1}(n)|\mathcal{F}_t] = \frac{n+1}{n}\{s_t(n+1,1) - \rho(n+1)\} - \{s_t(n,1) - \rho(n)\}. \tag{40}$$

Combining (39) and (40) gives

$$\begin{aligned}
\lim_{n\to\infty} \mathsf{E}[\Delta y_{t+1}(n)|\mathcal{F}_t] - \boldsymbol{\mu}_n^{ET}[\Delta y_{t+1}(n)] &= \lim_{n\to\infty}\left[\frac{n+1}{n}\{s_t(n+1,n) - s_t(n,n-1)\} - \right. \\
&\qquad\qquad \left. \frac{n+1}{n}\{s^\rho(n+1,n) - s^\rho(n,n-1)\}\right] \\
&= 0 \ a.s.,
\end{aligned}$$

by Eq. (16). For convergence of the uniform norm, it suffices to note that $\sup_{i=1,\dots,n}|\mathsf{E}[\Delta y_{t+1}(i)|\mathcal{F}_t] - \boldsymbol{\mu}_n^{ET}[\Delta y_{t+1}(i)]| = |\Delta y_{t+1}(n)|\mathcal{F}_t] - \boldsymbol{\mu}_n^{ET}[\Delta y_{t+1}(n)|. \quad \blacksquare$

Proof. (Theorem 6) Define the approximation error $w_t(1:n) := y_t(1:n) - z_{tn}$. The proof is by induction on t. Eq. (26) holds for $t = 0$ since $z_{0n} = y_0(1:n)$ a.s. Eq. (26) also holds for $t = 1$ by Theorem 5 since $w_0(1:n) = 0$ a.s., and hence

$$\begin{aligned}
\lim_{n\to\infty}||w_1(1:n)|| &= \lim_{n\to\infty}||\Delta y_1(1:n) - \Delta z_{1n}|| \\
&= \lim_{n\to\infty}||\mathsf{E}[\Delta y_1(1:n)|\mathcal{F}_0] - \alpha_n^{ET}(\beta_n' y_0(1:n) - \rho_n)|| \\
&= \lim_{n\to\infty}||\mathsf{E}[\Delta y_1(1:n)|\mathcal{F}_0] - \boldsymbol{\mu}_n^{ET}[\Delta y_1(1:n)]|| \\
&= 0 \ a.s. \quad [\text{by Eq. (22)}].
\end{aligned}$$

Suppose Eq. (26) holds for some $t \in \{1, 2, \dots\}$. It is required to show that $\lim_{n\to\infty}||\mathbf{w}_{t+1}(1:n)|| = 0$ a.s., i.e. Eq. (26) holds for $t+1$. Let $||.||_2$ denote the spectral norm of a square matrix and note that

$$\begin{aligned}
0 \leq ||w_{t+1}(1:n)|| &= ||w_t(1:n) + \Delta y_{t+1}(1:n) - \Delta z_{t+1,n}|| = \tag{41} \\
&= ||w_t(1:n) + \mathsf{E}[\Delta y_{t+1}(1:n)|\mathcal{F}_t] - \alpha_n^{ET}\{\beta_n'[y_t(1:n) - w_t(1:n)] - \rho_n\}|| \ a.s. \\
&\leq ||w_t(1:n)|| + ||\mathsf{E}[\Delta y_{t+1}(1:n)|\mathcal{F}_t] - \boldsymbol{\mu}_n^{ET}[\Delta y_{t+1}(1:n)|\mathcal{F}_t]|| + ||\alpha_n^{ET}\beta_n'||_2 \cdot ||w_t(1:n)|| \ a.s.,
\end{aligned}$$

by the triangle inequality and since $||\alpha_n^{ET}\beta_n' w_t(1:n)|| \leq ||\alpha_n^{ET}\beta_n'||_2 \cdot ||w_t(1:n)||$ for the spectral norm. The final line of Eq. (41) satisfies

$$\lim_{n\to\infty}\left\{||w_t(1:n)|| + ||\mathsf{E}[\Delta y_{t+1}(1:n)|\mathcal{F}_t] - \boldsymbol{\mu}_n^{ET}[\Delta y_{t+1}(1:n)|\mathcal{F}_t]|| + ||\alpha_n^{ET}\beta_n'||_2 \cdot ||w_t(1:n)||\right\}$$
$$= 0 \ a.s., \tag{42}$$

by the induction hypothesis, Theorem 5, and since $\lim_{n\to\infty}||\alpha_n^{ET}\beta_n'||_2 \cdot ||w_t(1:n)|| = 0$ a.s. when $||\alpha_n^{ET}\beta_n'||_2$ is bounded above for all n.[8]

---

[8]Whilst analytic results for general n are unavailable, computation of $||\alpha_n^{ET}\beta_n'||_2$ confirms that $||\alpha_n^{ET}\beta_n'||_2 \leq ||\alpha_2^{ET}\beta_2'||_2 = 4 = ||\alpha_n^{ET}\beta_n'||_\infty$ for $n = 2, 3, \dots, 1000$. Furthermore, it appears that $||\alpha_n^{ET}\beta_n'||_2$ converges to a limit approximately equal to 2.912 as $n \to \infty$.

Since the inequalities in Eq. (41) hold $\forall n$, it follows immediately from Eq. (42) that $\lim_{n \to \infty}$ $||w_{t+1}(1:n)|| = 0$ $a.s.$ as required. This completes the proof of (26) for all $t \in \{0, 1, 2, ...\}$. Eq. (27) follows straightforwardly by noting that

$$\lim_{n \to \infty} ||\mathbf{y}_n^T - \mathbf{z}_n^T||^2 = \sum_{t=0}^{T-1} \lim_{n \to \infty} ||\mathbf{w}_t(1:n)||^2 = 0 \quad a.s.$$

The proof using the uniform norm proceeds along exactly the same lines, except that instead of the spectral norm we use the the natural matrix norm induced by the uniform norm, i.e., $||\alpha_n^{ET}\beta_n'||_\infty = \max_{i=1,...,n} \sum_{j=1}^n |(\alpha_n^{ET}\beta_n')[i][j]|$. ∎

Proof. (Theorem 7) The proof follows as an interesting special case of Theorem 10 of Johansen (2008) in which $\alpha'_\perp \dot{A}(1)\beta_\perp$ is not only of reduced rank, but that rank is equal to zero. Theorem 10 and the associated Theorem 5 of Johansen (2008) continue to hold in this case,[9] setting

$$\alpha_1 = 0_{n \times 1}, \quad \beta_1 = 0_{n \times 1}, \quad \alpha_2 = \alpha_{n\perp}^{ET}, \quad \beta_2 = \beta_{n\perp}. \tag{43}$$

Inspection of the final column of α_n^{ET} reveals that $\alpha_n^{ET}[n-1]$ cannot be written as a linear combination of the previous rows $\{\alpha_n^{ET}[i]\}_{i=1}^{n-2}$. Hence $rank(\alpha_n^{ET}) = n - 1 = rank(\alpha_n^{ET}\beta_n')$, since β_n' has full row rank. Denote the ith element of the $n \times 1$ matrix $\alpha_{n\perp}^{ET}$ as $\alpha_{n\perp}[i]$. Then, for $n \geq 2$, we can take

$$\alpha_{n\perp}[n] = -1, \; \alpha_{n\perp}[n-1] = \frac{(n-1)(n+2)}{n^2}, \; \alpha_{n\perp}[n-2] = \frac{2(2-n)}{n^2(n-1)}, \quad \text{and}$$

$$\alpha_{n\perp}[i] = i \times \alpha_{n\perp}(1) \text{ for } i = 2, 3, ..., (n-2), \text{ and } \beta_{n\perp} = \mathbf{1}_{n \times 1}. \tag{44}$$

Note that we can write $\beta_{n\perp} = \alpha_n^{ET}\zeta_n$ and $\alpha_{n\perp}^{ET} = \beta_n\psi_n$, where $\zeta_n' = (\frac{1}{2}, 1, \frac{3}{2}, ..., \frac{(n-1)}{2})'$ and $\psi_n' = (\alpha_{n\perp}[2], \alpha_{n\perp}[3], ..., \alpha_{n\perp}[n])'$. It is then immediate that $\alpha_{n\perp}^{ET'}\beta_{n\perp} = 0_{1 \times 1}$. Furthermore the so-called I(2) condition, $\det(\alpha_{n\perp}^{ET'}\chi_n\beta_{n\perp}) \neq 0$, is satisfied since

$$\left| \alpha_{n\perp}^{ET'}\chi_n\beta_{n\perp} \right| = |\psi_n'\beta_n'\bar{\beta}_n\bar{\alpha}_n^{ET'}\alpha_n^{ET}\zeta_n| = |\psi_n'\zeta_n| = (n+1)/3n \neq 0. \tag{45}$$

The expressions for M_{2n} and M_{1n} follow from Eq.'s (12) and (13) of Johansen (2008), where $\theta_n = \dot{A}_n^{ET}(1)\chi_n\dot{A}_n^{ET}(1) + \frac{1}{2}\ddot{A}_n^{ET}(1)$. For the ET-VAR, $\dot{A}_n^{ET}(1) = -(I_n + \alpha_n^{ET}\beta_n')$ and $\ddot{A}_n^{ET}(1) = 0$. The expression for M_{1n} simplifies considerably since $\theta_n M_{2n} = \chi_n M_{2n} + M_{2n}$ and hence

$$M_{1n} = \chi_n M_{2n} + M_{2n}\chi_n(I_n - \theta_n M_{2n})$$

$$= \{\chi_n - M_2\chi_n[\chi_n + I_n]\} M_2 + M_2\chi_n.$$

_____

[9] Whilst analytic results for general n are unavailable, computation of the roots of $|A_n^{ET}(z)| = 0$ for $n = 2, 3, ..., 1000$ confirms that either $z = 1$ or $|z| > 1$, as required by Theorem 10 of Johansen (2008). In order to allow for imprecision in the computation, we deem $z = 1$ if the distance of the computed eigenvalue from 1 in the complex plane is less than 10^{-6} and deem $|z| > 1$ if the modulus of the computed eigenvalue is less than $1 - 10^{-6}$.

We note that an alternative, more straightforward proof of the MA representation may be constructed along the lines of Johansen (1996), where using the notation $\tilde{\Gamma}$ found there, the invertible matrix $\tilde{\Gamma} = (\bar{\alpha}_n^{ET} \bar{\alpha}_{n\perp}^{ET})'(\bar{\beta}_n \bar{\beta}_{n\perp})$. The MA representation of the resultant VAR(1) can be derived using Theorem 4.2 of Johansen (1996) since the transformed process obtained is I(1). However, this approach does not yield a closed form expression for M_{1n}. ∎

Lemma 11 (Reduced rank of $\beta_n' \alpha_n^{ET}$) *Let $n > 2$. The matrix $\beta_n' \alpha_n^{ET}$, has reduced rank equal to $(n-2)$ since (1) each row $(\beta_n' \alpha_n^{ET})[i]$, $i = 2, ..., (n-2)$ cannot be written as a linear combination of its predecessor rows; and (2) there exists a unique vector $\boldsymbol{\phi} = (\phi_1, ..., \phi_{n-2})'$ satisfying*

$$(\beta_n' \alpha_n^{ET})[n-1] = \sum_{i=1}^{n-2} \phi_i (\beta_n' \alpha_n^{ET})[i], \tag{46}$$

so that the final row is a linear combination of its predecessors. The vector $\boldsymbol{\phi}$ is given for $n > 4$ by

$$\phi_1 = \frac{-4}{n^2(n-1)}, \quad \phi_i = \frac{i+1}{2}\phi_1 \quad for \ i = 2, 3, ..., (n-3), \quad \phi_{n-2} = \frac{(n-1)(n+2)}{n^2}. \tag{47}$$

Note that $\phi_1 = \phi_{n-2}$ in (47) for $n = 3$, $\phi_2 = \phi_{n-2}$ in (47) for $n = 4$, and $\phi_1 = -4/n^2(n-1)$ for $n = 4$.

Proof. We give the proof for the more difficult case $n > 4$. The matrix $\beta_n' \alpha_n^{ET}$ is given by

$$\beta_n' \alpha_n^{ET} = \begin{pmatrix} -3 & 3/2 & 0 & 0 & \dots & 0 & 0 \\ -2 & -1 & 4/3 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -2 & 0 & 0 & 0 & \dots & \frac{n-1}{n-2} & 0 \\ -2 & 0 & 0 & 0 & \dots & -1 & \frac{n}{n-1} \\ -2 & 0 & 0 & 0 & \dots & -\frac{(n+1)}{n} & \frac{(n+2)}{n} \end{pmatrix}. \tag{48}$$

Note that the jth column of this matrix has exactly 2 non-zero elements for $j = 2, ..., (n-3), (n-1)$. Inspection of the jth columns for $j = 2, ..., (n-3)$ yields $\phi_i = \frac{i+1}{i}\phi_{i-1}$, $i = 2, ..., (n-3)$; and $\phi_{n-2} = (n-1)(n+2)/n^2$ follows directly from inspection of the final column. The first column implies that ϕ_1 must then satisfy

$$3\phi_1 + 2\phi_{n-2} + 2\sum_{i=2}^{n-3} \frac{(i+1)}{2}\phi_1 = 2, \tag{49}$$

implying $\phi_1 = -4/n^2(n-1)$. The $(n-2)$th column is the only column that now remains unused. Hence Lemma 11 is true since it is straightforward to show that

$$\phi_{n-3}\frac{(n-1)}{(n-2)} - \phi_{n-2} = -\frac{(n+1)}{n}. \tag{50}$$

∎

Proof. (Theorem 8) Denote the first $(n-2)$ rows of $\beta'_n \alpha_n^{ET}$ by

$$B'_n = \begin{pmatrix} -3 & 3/2 & 0 & 0 & \ldots & 0 & 0 \\ -2 & -1 & 4/3 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -2 & 0 & 0 & 0 & \ldots & -1 & \frac{n}{n-1} \end{pmatrix}, \tag{51}$$

and perform the decomposition $\beta'_n \alpha_n^{ET} = A_n B'_n$, where $A'_n = (I_{n-2}, \phi)$ and ϕ is given by Lemma 11. Both A_n and B_n are $(n-1) \times (n-2)$ of full rank. We can take $A_{n\perp} = (\phi', -1)'$ and $B_{n\perp} = (-1, -2, \ldots, -(n-1))'$. Since

$$A'_{n\perp} B_{n\perp} = (n-1) - \sum_{i=1}^{n-2} i\phi_i, \tag{52}$$

it follows that $n^2 A'_{n\perp} B_{n\perp} = \frac{2}{3} n(n+1)$ and hence $|A'_{n\perp} B_{n\perp}| \neq 0$ for all $n > 2$. Theorem 8 is then an implication of Theorem 4.2 of Johansen (1996),[10] with $B'_n s_{nt}$ as the $(n-2) \times 1$ vector of stationary cointegrating relations. Finally, it is possible to show that for all $n > 2$,

$$B'_n s_{nt} = D_n c_{nt}, \tag{53}$$

where the $(n-2) \times (n-2)$ non-singular matrix D_n is given by

$$D_n = \begin{pmatrix} 3/2 & 0 & 0 & \ldots & 0 & 0 \\ 5/3 & 4/3 & 0 & \ldots & 0 & 0 \\ 7/4 & 6/4 & 5/4 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{2n-3}{n-1} & \frac{2n-4}{n-1} & \frac{2n-5}{n-1} & \ldots & \frac{n+1}{n-1} & \frac{n}{n-1} \end{pmatrix}. \tag{54}$$

It follows that $c_{nt} = D_n^{-1} B'_n s_{nt}$ is itself stationary. The $(n-2)$ rows of $D_n^{-1} B'_n$ are linearly independent and are cointegrating vectors. $\blacksquare$

Lemma 12 *Suppose the existence of the almost sure limit $\nu_{t,L} := \lim_{\tau \to \infty} \nu_t(\tau)$. Then*

$$\lim_{n \to \infty} \frac{1}{n^2} \sum_{\tau=1}^{n-2} \nu_t(\tau)\tau = \frac{\nu_{t,L}}{2} \quad a.s. \tag{55}$$

Proof. Since the proof is straightforward, only an outline is given here. Define $\tilde{\nu}_t(\tau) := \nu_t(\tau) - \nu_{t,L}$, and fix some $\delta > 0$. Then decompose the sum as follows:

$$\frac{1}{n^2} \sum_{\tau=1}^{n-2} |\tilde{\nu}_t(\tau)|\tau = \frac{1}{n^2} \sum_{\tau=1}^{m-1} |\tilde{\nu}_t(\tau)|\tau + \frac{1}{n^2} \sum_{\tau=m}^{n-2} |\tilde{\nu}_t(\tau)|\tau,$$

---

[10]Whilst analytic results for general n are unavailable, computation of the roots of the characteristic polynomial of the VAR (30), $|I_{n-1} - (I_{n-1} + \beta'_n \alpha_n^{ET})z| = 0$, confirms that either $z = 1$ or $|z| > 1$ for $n = 3, \ldots, 1000$. We deem $z = 1$ if the distance of the computed eigenvalue from 1 in the complex plane is less than 10^{-10} and deem $|z| > 1$ if the modulus of the computed eigenvalue is less than $1 - 10^{-10}$.

where $|\tilde{\nu}_t(\tau)| < \delta \; \forall \tau \geq m$. It follows that for sufficiently large n

$$\frac{1}{n^2}\sum_{\tau=1}^{n-2}|\tilde{\nu}_t(\tau)|\tau < \delta + \frac{1}{n^2}\sum_{\tau=1}^{n-2}\delta\tau < 2\delta,$$

hence $\lim_{n\to\infty}\frac{1}{n^2}\sum_{\tau=1}^{n-2}\tilde{\nu}_t(\tau)\tau = 0$ $a.s.$, which implies Eq. (55). ∎

Proof. (Theorem 9) Since $M_{2n} = \beta_{n\perp}(\alpha_{n\perp}^{ET\prime}\chi_n\beta_{n\perp})^{-1}\alpha_{n\perp}^{ET\prime}$ (see Eq. 29), $(\alpha_{n\perp}^{ET\prime}\chi_n\beta_{n\perp})^{-1} = -3n/(n+1)$ (see Eq. 45), and $\beta_{n\perp} = \mathbf{1}_n$, it follows that

$$\lim_{n\to\infty}\sup_{i\in\{1,..,n\}}|(M_{2n}\nu_{n,r})[i]| = \lim_{n\to\infty}\left|\frac{-3n}{(n+1)}\alpha_{n\perp}^{ET\prime}\nu_{n,r}\right|, \tag{56}$$

since $(M_{2n}\nu_{n,r})[i]$ does not depend on i. Now recalling the definition of $\alpha_{n\perp}^{ET}$ from Eq. (44), we find that

$$\alpha_{n\perp}^{ET\prime}\nu_{n,r} = \frac{[n-1][n+2]}{n^2}\nu_r(n-1) - \nu_r(n) - \frac{2}{n^2[n-1]}\sum_{\tau=1}^{n-2}\tau\nu_r(\tau). \tag{57}$$

Lemma 12 implies that the third term on the RHS of Eq. (57) converges to zero $a.s.$ and therefore that

$$\lim_{n\to\infty}(M_{2n}\nu_{n,r})[1] = -3[1\cdot\nu_{r,L} - \nu_{r,L}] - 0 = 0 \quad a.s. \tag{58}$$

Eq. (33) then follows directly for all finite t, since $||\sum_{u=1}^{t}\sum_{r=1}^{u}M_{2n}\nu_{n,r}||_\infty = |\sum_{u=1}^{t}\sum_{r=1}^{u}(M_{2n}\nu_{n,r})[1]|$. ∎

Proof. (Theorem 10) Recall that $M_{1n} = -\{\chi_n - M_{2n}\chi_n[\chi_n + I_n]\}\beta_{n\perp}k_n\alpha_{n\perp}^{ET\prime} + M_{2n}\chi_n$, where the scalar $k_n := 3n/(n+1)$ (see Eq. 29). We note that the almost sure convergence of $-\tau\nu_{t+1}(\tau)$ implies that $\nu_{t+1}(\tau) \to 0$ $a.s.$ We note also that the n-vector $\Psi_n := \{\chi_n - M_{2n}\chi_n[\chi_n + I_n]\}\beta_{n\perp}$ can be written as

$$\Psi_n = \tilde{\Psi}_n + r_n := \left[\frac{-(n+2)}{4}, \frac{-n}{4}, ..., \frac{(n-4)}{4}\right]' + r_n,$$

where the remainder term satisfies that $||r_n||_\infty \leq ||r_2||_\infty = 0.25$ $\forall n$ (which of course implies that $r_n[1]$ is $O(1)$ as a sequence in n).

(i) **Sufficiency.** It is enough to establish sufficiency to show that, under the condition in Eq. (35), both $\lim_{n\to\infty}||\Psi_n k_n\sum_{r=1}^{t}\alpha_{n\perp}^{ET\prime}\nu_{n,r}||_\infty = 0$ $a.s.$ and $\lim_{n\to\infty}||M_{2n}\chi_n\sum_{r=1}^{t}\nu_{n,r}||_\infty = 0$ $a.s.$

Consider the first of these limits, and note that

$$\arg\sup_{i\in\{1,..,n\}}\left|\{\tilde{\Psi}_n[i] + r_n[i]\}k_n\sum_{r=1}^{t}\alpha_{n\perp}^{ET\prime}\nu_{n,r}\right|$$

$$= \arg\sup_{i\in\{1,..,n\}}\left|\tilde{\Psi}_n[i] + r_n[i]\right| = 1 \quad \textit{for all } n,$$

by examination of the form of $\tilde{\Psi}_n$ and since the maximum is obtained either by minimising or maximising $(\tilde{\Psi}_n[i] + r_n[i])$. It is readily seen, since $||r_n||_\infty \leq 0.25$, that the maximiser is given

by $i = n$, the minimiser by $i = 1$, and that $\left|\tilde{\Psi}_n[1] + r_n[1]\right| > \left|\tilde{\Psi}_n[n] + r_n[n]\right|$ $\forall n$. Therefore the sequence $\{\|\Psi_n k_n \sum_{r=1}^{t} \alpha_{n\perp}^{ET\prime} \nu_{n,r}\|_\infty\}_n$ has its nth element equal to $|\Psi_n[1] k_n \sum_{r=1}^{t} \alpha_{n\perp}^{ET\prime} \nu_{n,r}|$. We will now show that $\lim_{n\to\infty} \sum_{r=1}^{t} \Psi_n[1] k_n \alpha_{n\perp}^{ET\prime} \nu_{n,r} = 0$ a.s. We have that:

$$
\begin{aligned}
\left(\Psi_n k_n \alpha_{n\perp}^{ET\prime} \nu_{n,r}\right)[1] &= \frac{(2 - n - 4)k_n}{4} \alpha_{n\perp}^{ET\prime} \nu_{n,r} + r_n[1] \cdot k_n \alpha_{n\perp}^{ET\prime} \nu_{n,r} \\
&= f(1,n)\{[n-1]\nu_r(n-1) - n\nu_r(n)\} + f(1,n)\left\{\frac{2(n-1)}{n}\nu_r(n-1)\right\} \\
&\quad - f(1,n)\left\{\frac{2n}{(n-1)}\frac{1}{n^2}\sum_{\tau=1}^{n-2}\tau\nu_r(\tau)\right\} + O(1)\alpha_{n\perp}^{ET\prime}\nu_{n,r}, \quad (59)
\end{aligned}
$$

where $f(i,n) := \frac{(2i-n-4)k_n}{4n}$ and use is made of Eq. (57). Notice that Theorem 9 establishes that $\lim_{n\to\infty} O(1)\alpha_{n\perp}^{ET\prime}\nu_{n,r} = 0$ a.s. when $\nu_r(\tau)$ converges a.s. to $\nu_{r,L}$ (see Eq. 56). Since $\nu_{t,L} = 0$ here, Lemma 12 implies that $\lim_{n\to\infty} f(1,n)\frac{2n}{(n-1)}\frac{1}{n^2}\sum_{\tau=1}^{n-2}\tau\nu_r(\tau) = 0$ a.s. since $f(1,n)$ is $O(1)$. It remains to consider the term $f(1,n)\{[n-1]\nu_r(n-1) - n\nu_r(n)\}$, which clearly converges a.s. to zero since $\tau\nu_r(\tau) \to -\nu_{r,p}$. Therefore, $\lim_{n\to\infty} \Psi_n[1] k_n \alpha_{n\perp}^{ET\prime}\nu_{n,r} = 0$ a.s., which completes this part of the proof.

Consider now the second of the limits, namely $\lim_{n\to\infty} \|M_{2n}\chi_n \sum_{r=1}^{t}\nu_{n,r}\|_\infty = 0$ a.s.

The $(n \times n)$ matrix $M_{2n}\chi_n = \beta_{n\perp}(\alpha_{n\perp}^{ET\prime}\chi_n\beta_{n\perp})^{-1}\alpha_{n\perp}^{ET\prime}\chi_n = -k_n\alpha_n^{ET}\zeta_n\psi_n'\beta_n'\bar{\beta}_n\bar{\alpha}_n^{ET\prime} = -k_n \alpha_n^{ET}\zeta_n\psi_n'\bar{\alpha}_n^{ET\prime}$. The matrix $M_{2n}\chi_n$ satisfies the following two properties: i) for all n, the matrix consists of identical rows and the elements of each row are positive; ii) $\lim_{n\to\infty}(M_{2n}\chi_n)[1][\tau] = 0$ for fixed, finite maturity τ. Property i) implies the further Property iii) the sum of each row equals unity. That is $\sum_{\tau=1}^{n}(M_{2n}\chi_n)[1][\tau] = 1$, since $tr[M_{2n}\chi_n] = -k_n tr[\psi_n'\bar{\alpha}_n^{ET\prime}\alpha_n^{ET}\zeta_n] = 1$. These 3 properties and the convergence of $\nu_r(\tau)$ together imply, by application of Toeplitz's lemma for triangular arrays (see Davidson 2002, p.34), that $\lim_{n\to\infty}(M_{2n}\chi_n\nu_{n,r})[1] = \nu_{r,L} = 0$ a.s. Since the rows of $(M_{2n}\chi_n\sum_{r=1}^{t}\nu_{n,r})$ are identical, $\lim_{n\to\infty}\|M_{2n}\chi_n\sum_{r=1}^{t}\nu_{n,r}\|_\infty = \lim_{n\to\infty}|(M_{2n}\chi_n\sum_{r=1}^{t}\nu_{n,r})[1]| = 0$ a.s.

(ii) **Necessity.** Suppose then that for some r, $\nu_r(\tau) \to \nu_{r,L}$ a.s., but $\tau\nu_r(\tau)$ is not convergent a.s. and hence $\nu_{r,L} \neq 0$ a.s. Then $\lim_{n\to\infty}(M_{2n}\chi_n\nu_{n,r})[1] = \nu_{r,L} \neq 0$ a.s. Noting that $\lim_{n\to\infty} f(1,n) = -3/4$, it is readily seen that for $i = 1$, the last 3 terms on the RHS of Eq. (59) also converge to finite limits a.s. However the first term, $f(1,n)\{[n-1]\nu_r(n-1) - n\nu_r(n)\}$, now fails to converge a.s. Therefore, $(M_{1n}\nu_{n,r})[1]$ does not converge a.s. If Eq. (34) holds then $\lim_{n\to\infty}\sum_{r=1}^{t}(M_{1n}\nu_{n,r})[1] = 0$ a.s. $\forall t$ and hence $\lim_{n\to\infty}(M_{1n}\nu_{n,r})[1] = 0$ a.s. $\forall r$, which contradicts the previous sentence.Therefore, the a.s. existence of the finite limit $\nu_{r,p}$ $\forall r$ is necessary for Eq. (34) to hold. $\blacksquare$