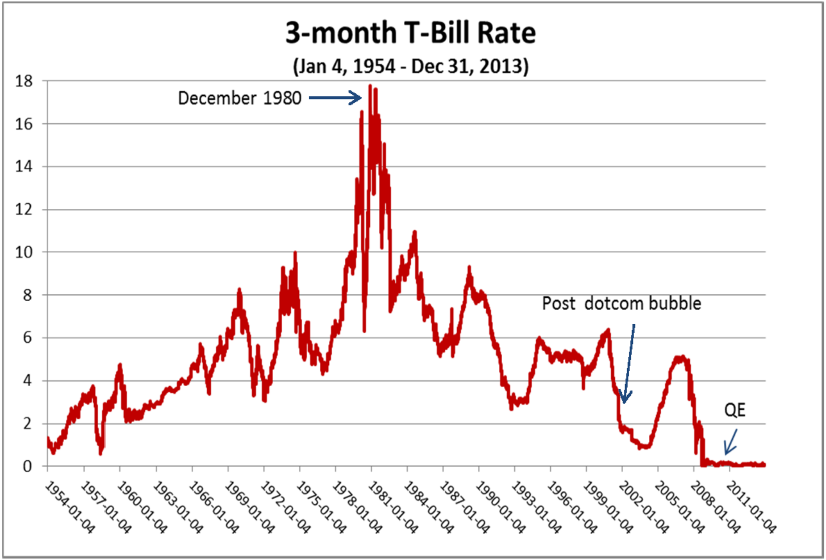
Local-Momentum Autoregression and the Modeling of Interest Rate Term Structure

Jin-Chuan Duan
Business School, Risk Management Institute, and Department of Economics
(www.rmi.nus.edu.sg/duanjc)

National University of Singapore

(April 2015)

(The 7th Annual NYU-Stern School Volatility Institute Conference)


3-month T-Bill Rate
(Jan 4, 1954 - Dec 31, 2013)
December 1980
Post dotcom bubble
QE
18
16
14
12
10
8
6
4
2
0
1954-01-04
1957-01-04
1960-01-04
1963-01-04
1966-01-04
1969-01-04
1972-01-04
1975-01-04
1978-01-04
1981-01-04
1984-01-04
1987-01-04
1990-01-04
1993-01-04
1996-01-04
1999-01-04
2002-01-04
2005-01-04
2008-01-04
2011-01-04

Standard mean-reversion

$$\begin{aligned} \Delta X_t &= \kappa_x(\mu - X_{t-1}) + \sigma_x \varepsilon_t \\ \varepsilon_t | \mathcal{F}_{t-1} &\sim D(0,1) \end{aligned}$$

Is the AR(1) model of this type suitable for modeling interest rates or other highly persistent systems with long periods of directional moves?

How about adding a latent stochastic central tendency factor?

$$
\begin{aligned}
\Delta X_t &= \kappa_x(\mu_t - X_{t-1}) + \sigma_x \varepsilon_t \\
\Delta \mu_t &= \kappa_\mu(\bar{\mu} - \mu_{t-1}) + \sigma_\mu \epsilon_t \\
\varepsilon_t | \mathcal{G}_{t-1} &\sim D(0, 1)
\end{aligned}
$$

Note: Adding a latent stochastic central tendency factor is an idea in Balduzzi, Das and Foresi (1998, *Review of Economics and Statistics*), because empirical evidence suggests that the short-term interest rate tends to move towards the long term interest rate.

Local-momentum with latent central tendency (LM-CTAR)

$$
\begin{aligned}
\Delta X_t &= \kappa_x\left(\mu_t - X_{t-1}\right) + \omega\left(\bar{X}_{(t-1)|n} - X_{t-1}\right) + \sigma_x \varepsilon_t \\
\Delta \mu_t &= \kappa_\mu\left(\bar{\mu} - \mu_{t-1}\right) + \sigma_\mu \epsilon_t \\
\bar{X}_{(t-1)|n} &= \sum_{i=t-n}^{t-1} b_{t-i} X_i \\
\varepsilon_t |\, \mathcal{G}_{t-1} &\sim D(0,1),\ \ \epsilon_t |\, \mathcal{G}_{t-1} \sim D(0,1)
\end{aligned}
$$

where $\sum_{i=1}^{n} b_i = 1$ with $b_i \geq 0$ for $i = 1, 2, \cdots, n$. $\bar{X}_{(t-1)|n}$ is meant to be some sort of moving weighted sample mean.

(Note: Exponentially decaying weights with n being set to ∞ lead to a 3-dimensional Markov system, i.e., $(X_t, \bar{X}_{t|\infty}, \mu_t)$, with one extra decaying parameter and an additional latent factor, $\bar{X}_{t|\infty}$.)

Local-momentum without latent central tendency (LM-AR)

$$
\begin{aligned}
\Delta X_t &= \kappa_x \left(\bar{\mu} - X_{t-1}\right) + \omega \left(\bar{X}_{(t-1)|n} - X_{t-1}\right) + \sigma_x \varepsilon_t \\
\bar{X}_{(t-1)|n} &= \sum_{i=t-n}^{t-1} b_{t-i} X_i \\
\varepsilon_t | \mathcal{G}_{t-1} &\sim D(0,1)
\end{aligned}
$$

Interesting features

- Local momentum building: $\omega < 0$
- Local momentum preserving: $\omega > 0$

Basic properties

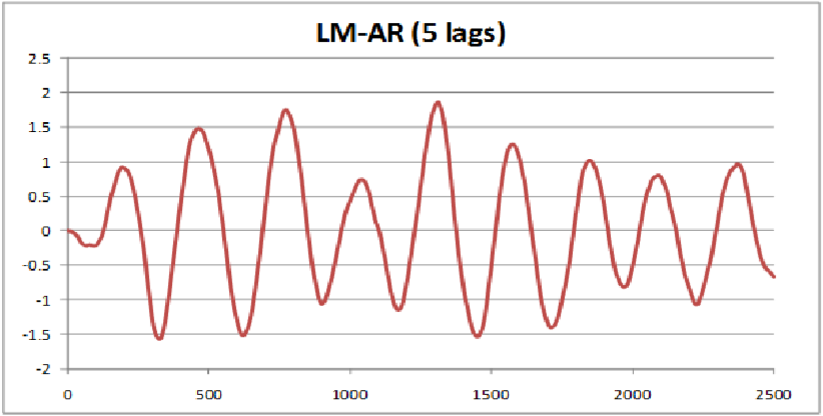
- Stationarity and ergodicity of LM-CTAR can be characterized by recognizing it as ARMA(n,∞). The spectral radius of the AR coefficient matrix less than 1 is both sufficient and necessary, because the MA(∞) coefficients are absolutely summable.
- Easy to verify sufficiency conditions are given in the paper.

LM-CTAR model in a matrix form

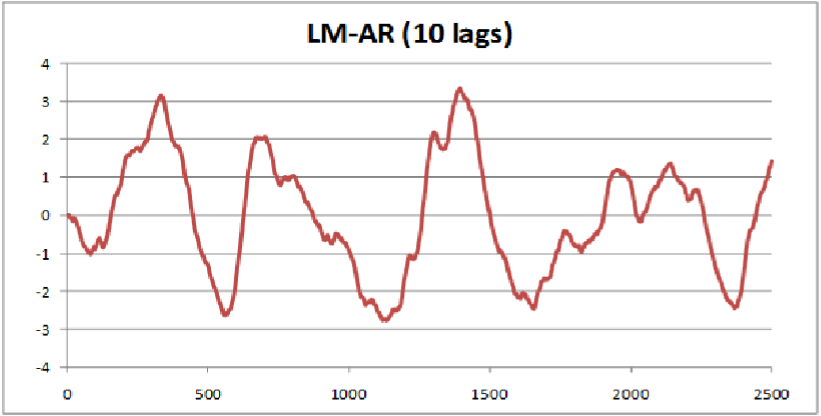
$$\mathbf{X}_t = \mathbf{A} + \mathbf{B}\mathbf{X}_{t-1} + \mathbf{Z}_t$$

$$\mathbf{X}_t = \begin{bmatrix} X_t \\ X_{t-1} \\ \vdots \\ X_{t-n+1} \end{bmatrix} \quad \mathbf{Z}_t = \begin{bmatrix} \kappa_x(\mu_t - \bar{\mu}) + \sigma_x \varepsilon_t \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \mathbf{A} = \begin{bmatrix} \kappa_x \bar{\mu} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} 1 - \kappa_x - \omega(1 - b_1) & \omega b_2 & \ldots & \omega b_{n-1} & \omega b_n \\ 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \end{bmatrix}$$



A simulated sample path for the LM-AR with 5 lags. The parameters are: $\bar{\mu} = 0, \kappa_x = 0.001, \omega = -0.5, \sigma_x = 0.002$ and $\sigma_\mu = 0$.



A simulated sample path for the LM-AR with 10 lags. The parameters are: $\bar{\mu} = 0, \kappa_x = 0.001, \omega = -0.2, \sigma_x = 0.02$ and $\sigma_\mu = 0$.

Continuous-time LM-CTAR

$$
\begin{aligned}
dX_t &= \left[\kappa_x\left(\mu_t - X_t\right) + \omega\left(\bar{X}_t(\tau) - X_t\right)\right] dt + \sigma_x dW_{xt} \\
d\mu_t &= \kappa_\mu\left(\bar{\mu} - \mu_{t-1}\right) dt + \sigma_\mu dW_{\mu t} \\
\bar{X}_t(\tau) &= \int_{t-\tau}^{t} b(t-s) X_s ds
\end{aligned}
$$

where $\kappa_\mu > 0$, $\sigma_\mu > 0$, $\kappa_x \geq 0$, $\sigma_x > 0$, and $\int_0^\tau b(s)ds = 1$ with $b(s) \geq 0$ for $0 \leq s \leq \tau$; and W_{xt} and $W_{\mu t}$ are two independent Wiener processes.

(Note: Again exponentially decaying weights with n being set to ∞ can lead to a 3-dimensional Markov system, i.e., $(X_t, \bar{X}_{t|\infty}, \mu_t)$, with one extra decaying parameter and an additional latent factor, $\bar{X}_{t|\infty}$.)

US Treasury rates (constant maturity, continuously compounded)

Maturity	# of Points	Mean	Median	Standard Deviation
1 month	641	0.015144	0.009636	0.016606
3 months	3080	0.047902	0.047220	0.031239
6 months	2831	0.052203	0.051244	0.031604
1 year	2451	0.057484	0.055555	0.031759
5 years	2674	0.060156	0.058458	0.027835
10 years	2674	0.063239	0.060107	0.025515
20 years	2327	0.063808	0.058458	0.024891
	Skewness	**Excess Kurtosis**	**Maximum**	**Minimum**
1 month	0.942746	-0.397540	0.052626	0
3 months	0.885921	1.416064	0.176280	0
6 months	0.718757	1.086112	0.167857	0.000304
1 year	0.484282	0.759180	0.166050	0.000811
5 years	0.521315	0.396290	0.150745	0.005584
10 years	0.721610	0.391892	0.147040	0.014199
20 years	1.016944	0.596012	0.146522	0.020880

(Every Wednesday in the period of Jan 4, 1954 to December 31, 2013)

Using the weekly series of one interest rate (3-month)

Parameter	AR(1)	CTAR	LM-AR	LM-CTAR
$\bar{\mu}$	0.046387 (0.024706)	0.046832 (0.019745)	0.047097 (0.019325)	0.045395 (0.022350)
k_x	0.002632 (0.000941)	0.004181 (0.001324)	0.003381 (0.000977)	0.105702 (0.025165)
ω			-0.071005 (0.004154)	-0.075354 (0.007143)
σ_x	0.002328 (0.000008)	0.002189 (0.000019)	0.002314 (0.000008)	0.002177 (0.000041)
k_μ		0.295870 (0.039138)		0.003284 (0.001234)
σ_μ		0.134946 (0.048605)		0.002511 (0.000132)
Log-likelihood	14231.12	14247.24	14249.08	14253.61
$\rho(\mathbf{B})$			0.995690	0.825882
Data points	3130			
Missing data	50			

(Both versions of the local-momentum model use a 7-week moving-window average.)

A 3-factor term structure model

The base interest rate dynamic has two components: global driver (X_t) and local variation (v_t):

$$r_t = X_t + v_t$$

(Note: The base rate is driven by 3 latent factors, because X_t is latent, X_t's central tendency is also latent, and v_t is latent.)

The local variation is a standard AR(1) process:

$$\begin{aligned} \Delta v_t &= -\kappa_v v_{t-1} + \sigma_v \xi_t \\ \xi_t | (\mathcal{G}_t \cup v_{t-1}) &\sim N(0,1) \end{aligned}$$

where $0 < \kappa_v < 2$, and $(\mathcal{G}_t \cup v_{t-1})$ denotes the minimum σ-algebra generated by $\mathcal{G}_t$ and v_{t-1}.

The 3-factor model in a matrix form

$$\begin{aligned} r_t &= \mathbf{H}'\mathbf{X}_t^* \\ \mathbf{S}\mathbf{X}_t^* &= \mathbf{C} + \mathbf{D}\mathbf{X}_{t-1}^* + \mathbf{W}_t \end{aligned}$$

where

$$\mathbf{H} = \begin{bmatrix} 1 \\ \mathbf{0}_{n-1} \\ 0 \\ 1 \end{bmatrix} \quad \mathbf{X}_t^* = \begin{bmatrix} \mathbf{X}_t \\ \mu_t - \bar{\mu} \\ v_t \end{bmatrix} \quad \mathbf{W}_t = \begin{bmatrix} \sigma_x \varepsilon_t \\ \mathbf{0}_{n-1} \\ \sigma_\mu \epsilon_t \\ \sigma_v \xi_t \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} \mathbf{A} \\ 0 \\ 0 \end{bmatrix}$$

$$\mathbf{D} = \begin{bmatrix} \mathbf{B} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1-\kappa_\mu & 0 \\ \mathbf{0} & 0 & 1-\kappa_v \end{bmatrix} \quad \mathbf{S} = \begin{bmatrix} 1 & \mathbf{0} & -\kappa_x & 0 \\ \mathbf{0} & \mathbf{I}_{(n-1)\times(n-1)} & \mathbf{0} & \mathbf{0} \\ 0 & \mathbf{0} & 1 & 0 \\ 0 & \mathbf{0} & 0 & 1 \end{bmatrix}$$

$$\mathbf{U} = \begin{bmatrix} \sigma_x^2 & \mathbf{0} & 0 & 0 \\ \mathbf{0} & \mathbf{0}_{(n-1)\times(n-1)} & \mathbf{0} & \mathbf{0} \\ 0 & \mathbf{0} & \sigma_\mu^2 & 0 \\ 0 & \mathbf{0} & 0 & \sigma_v^2 \end{bmatrix}$$

Note: **U** is the covariance matrix for $\mathbf{W}_t$, and $\mathbf{X}_t$, **A** and **B** have been previously defined.

Stochastic discount factor and term structure

h: the length of one period measured as the fraction of a year. The stochastic discount factor from time $t+\tau$ back to time t is assumed to be $\exp[-r_t(\tau)\tau h]M_{t,t+\tau}$ where for $s \geq t$,

$$M_{t,s} = \alpha(t,s)\exp\{\sum_{j=t+1}^{s} \quad (\lambda_0 + \lambda_1 X_{j-1})\varepsilon_j + (\psi_{\mu 0} + \psi_{\mu 1}\mu_{j-1})\epsilon_j + (\psi_{v0} + \psi_{v1} v_{j-1})\xi_j\}$$

Note that $\alpha(t,s)$ is the factor that makes $M_{t,s}$ a martingale for $s \geq t$.

Define a martingale measure $Q_{t,T}$ by setting $dQ_{t,T}/dP = M_{t,T}$.

Forward and spot interest rates

$f_t(\tau)$: the one-period forward rate at time t starting at time $t+\tau$, where each of the τ periods is of length h. $\mathcal{H}_t$: the filtration generated by $\{(X_s, \mu_s, v_s); s \leq t\}$. It follows that, for $\tau \geq 1$,

$$f_t(\tau) = -\frac{\ln E^Q\left(e^{-r_{t+\tau}h} | \mathcal{H}_t\right)}{h}$$

Note: The base interest rate r_t equals $f_t(0)$.

Forward rates for different forward starting times can in turn be used to compute spot interest rate rates such as, for $\tau \geq 1$,

$$r_t(\tau) = \frac{1}{\tau}\sum_{j=0}^{\tau-1} f_t(j).$$

Risk-neutral system

$$
\begin{aligned}
r_t &= \mathbf{H}'\mathbf{X}_t^* \\
\mathbf{S}\mathbf{X}_t^* &= \mathbf{C}^* + \mathbf{D}^*\mathbf{X}_{t-1}^* + \mathbf{W}_t^*
\end{aligned}
$$

where

$$
\mathbf{C}^* = \begin{bmatrix} \kappa_x\bar{\mu} + \sigma_x\lambda_0 \\ \mathbf{0}_{n-1} \\ \sigma_\mu\psi_{\mu 0} \\ \sigma_v\psi_{v0} \end{bmatrix} \quad \mathbf{W}_t^* = \begin{bmatrix} \sigma_x\varepsilon_t^Q \\ \mathbf{0}_{n-1} \\ \sigma_\mu\epsilon_t^Q \\ \sigma_v\xi_t^Q \end{bmatrix}
$$

$$
\mathbf{D}^* = \begin{bmatrix}
d & \omega b_2 & \dots & \omega b_{n-1} & \omega b_n & 0 & 0 \\
1 & 0 & \dots & 0 & 0 & 0 & 0 \\
0 & 1 & \dots & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & \dots & 1 & 0 & 0 & 0 \\
0 & 0 & \dots & 0 & 0 & 1-\kappa_\mu+\sigma_\mu\psi_{\mu 1} & 0 \\
0 & 0 & \dots & 0 & 0 & 0 & 1-\kappa_v+\sigma_v\psi_{v1}
\end{bmatrix}
$$

and $d = 1 - \kappa_x - \omega(1-b_1) + \sigma_x\lambda_1$.

The term structure formula

For $\tau \geq 1$,

$$r_t(\tau) = \Phi_1(\tau) + \Phi_2(\tau)\mathbf{X}_t^*$$

where

$$\begin{aligned}
\Phi_1(\tau) &= \mathbf{H}'(\mathbf{I} - \mathbf{S}^{-1}\mathbf{D}^*)^{-1}\left(\mathbf{I} - \frac{1}{\tau}\sum_{j=0}^{\tau-1}(\mathbf{S}^{-1}\mathbf{D}^*)^j\right)\mathbf{S}^{-1}\mathbf{C}^* \\
&\quad -\mathbf{H}'\left(\frac{h}{2\tau}\sum_{j=0}^{\tau-1}\sum_{i=0}^{j-1}(\mathbf{S}^{-1}\mathbf{D}^*)^i\mathbf{S}^{-1}\mathbf{U}(\mathbf{S}^{-1})'[(\mathbf{S}^{-1}\mathbf{D}^*)^i]'\right)\mathbf{H} \\
\Phi_2(\tau) &= \mathbf{H}'\left(\frac{1}{\tau}\sum_{j=0}^{\tau-1}(\mathbf{S}^{-1}\mathbf{D}^*)^j\right)
\end{aligned}$$

Estimation by the Kalman filter

Facing yields of several maturities gives rise to the measurement equations:

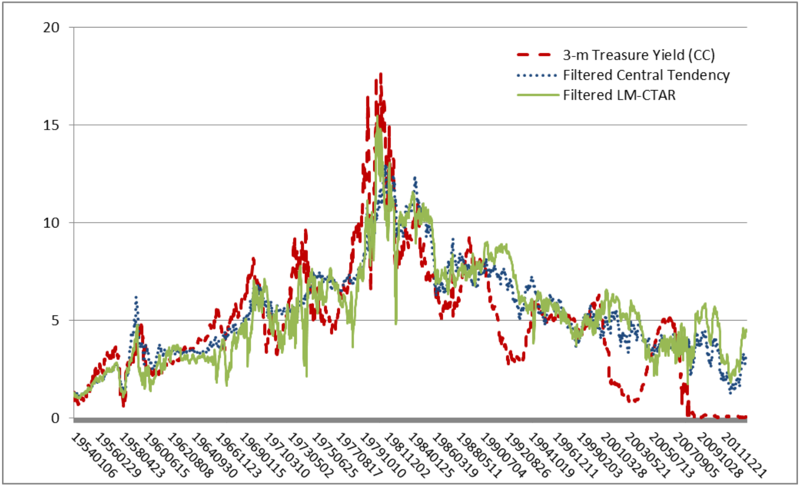
$$
\begin{aligned}
\tilde{r}_t(\tau_1) &= \Phi_1(\tau_1) + \Phi_2(\tau_1)\mathbf{X}_t^* + \epsilon_{1t} \\
\tilde{r}_t(\tau_2) &= \Phi_1(\tau_2) + \Phi_2(\tau_2)\mathbf{X}_t^* + \epsilon_{1t} \\
&\vdots \\
\tilde{r}_t(\tau_k) &= \Phi_1(\tau_k) + \Phi_2(\tau_k)\mathbf{X}_t^* + \epsilon_{kt}
\end{aligned}
$$

The number of identifiable parameters

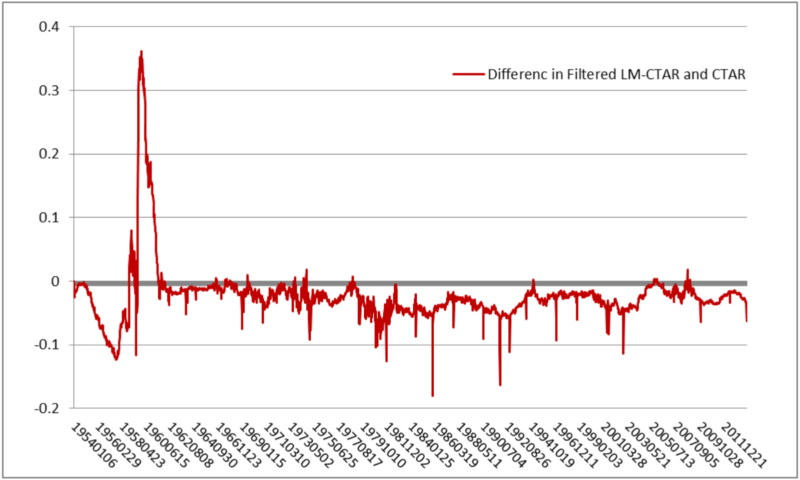
The three latent factor processes are governed by eight parameters under the physical probability, i.e., $\{\bar{\mu}, \kappa_x, \omega, \sigma_x, \kappa_\mu, \sigma_\mu, \kappa_v, \sigma_v\}$. There are six parameters arising from risk-neutralization, i.e., $\{\lambda_0, \lambda_1, \psi_{\mu 0}, \psi_{\mu 1}, \psi_{v0}, \psi_{v1}\}$, but there is only one identifiable parameter among λ_0, $\psi_{\mu 0}$ and ψ_{v0} because these three enter into the same constant in the risk neutral system.

	CTAR+AR(1)	LM-CTAR+AR(1)
Physical process parameters		
$\bar{\mu}$	0.103950 (0.002614)	0.095287 (0.002504)
k_x	0.072262 (0.000993)	0.080571 (0.001112)
ω		-0.007134 (0.000690)
σ_x	0.002649 (0.000028)	0.002591 (0.000027)
k_μ	0.000189 (0.000011)	0.000188 (0.000011)
σ_μ	0.001440 (0.000017)	0.001317 (0.000018)
k_v	0.008227 (0.000087)	0.008176 (0.000086)
σ_v	0.002058 (0.000019)	0.002050 (0.000019)

Measurement errors		
1 month	0.000892 (0.000039)	0.000886 (0.000039)
3 months	0.000576 (0.000013)	0.000579 (0.000013)
6 months	0.000746 (0.000009)	0.000742 (0.000009)
1 year	0.000628 (0.000010)	0.000636 (0.000010)
5 years	0.001767 (0.000022)	0.001770 (0.000022)
10 years	0.000174 (0.000032)	0.000174 (0.000031)
20 years	0.002362 (0.000030)	0.002365 (0.000030)
Log-likelihood	96976.40	97012.06
$\rho(\mathbf{B})$		0.891281
Data points	3130	
Missing data	50	



A 3-factor Gaussian term structure model built on the LM-CTAR process is fitted to the US treasury constant maturity yields (continuously compounded) of seven maturities (1 month, 3 months, 6 months, 1 year, 5 years, 10 years and 20 years). The filtered estimate of the LM-CTAR and central tendency components are plotted along with the 3-month rate, weekly from January 4, 1954 to December 31, 2013. The vertical axis is in percentage points.

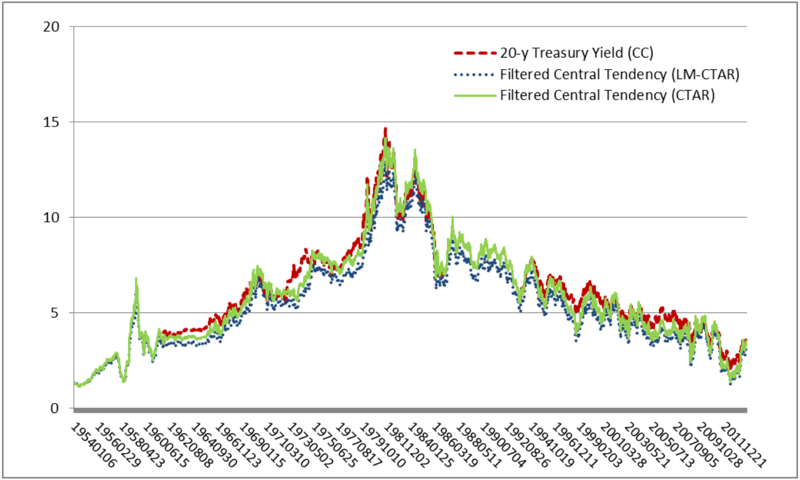


The difference in the filtered LM-CTAR and CTAR factors from two 3-factor term structure models over the sample period from January 4, 1954 to December 31, 2013.

Impacts on yields by the central tendency and local momentum factors

$$\Delta r_t(\tau) = a_\tau + b_\tau \Delta\hat{\mu}_{t|t} + c_\tau \Delta\hat{X}_{t|t} + d_\tau \Delta\hat{v}_{t|t} + \epsilon_t(\tau)$$

Yield Change	**Intercept**	$\Delta\hat{\mu}_{t\|t}$	$\Delta\hat{X}_{t\|t}$	$\Delta\hat{v}_{t\|t}$	R^2
$\Delta r_t(1m)$	0.000109 (0.000083)	0.220220 (0.007196)	0.991148 (0.003920)	0.976043 (0.008784)	0.9148
$\Delta r_t(3m)$	0.000017 (0.000009)	0.351476 (0.000573)	0.732278 (0.000194)	0.952301 (0.000265)	0.9850
$\Delta r_t(6m)$	0.000067 0.000017	0.583060 0.000982	0.413108 0.000325	0.944362 0.000449	0.9511
$\Delta r_t(1y)$	-0.000042 (0.000018)	0.768468 (0.001023)	0.190847 (0.000319)	0.881598 (0.000424)	0.9590
$\Delta r_t(5y)$	-0.000352 (0.000017)	1.030643 (0.001013)	0.039280 (0.000327)	0.384023 (0.000454)	0.9181
$\Delta r_t(10y)$	-0.000011 (0.000001)	1.035115 (0.000074)	0.026693 (0.000024)	0.227487 (0.000033)	0.9995
$\Delta r_t(20y)$	0.000028 (0.000018)	0.902467 (0.001052)	0.038297 (0.000317)	0.191439 (0.000440)	0.9004



The filtered central tendency estimates corresponding to two versions of the 3-factor Gaussian term structure model built on, respectively, the CTAR and LM-CTAR processes from January 4, 1954 to December 31, 2013 on a weekly frequency. Also plotted is the 20-year US Treasury yields (continuously compounded) when they were available. The vertical axis is in percentage points.

Future Research: monetary regimes

- Does it make sense to define regimes as high, average and low rates (or volatilities)? Can one conduct monetary easing while in the low-rate regime?
- How about classifying interest rate regimes as Status Quo, Monetary Easing and Monetary Tightening? Entering Quantitative Easing (QE) is “Monetary Easing” and staying in QE will be “Status Quo”.
- In terms of the local-momentum model, “Status Quo” means using the current parameter values, and entering “Monetary Easing (Tightening)” state means subtracting (adding) a positive constant from (to) $\bar{\mu}$, and this can be done repeatedly. In addition, the local-momentum parameter ω can also be regime-specific.