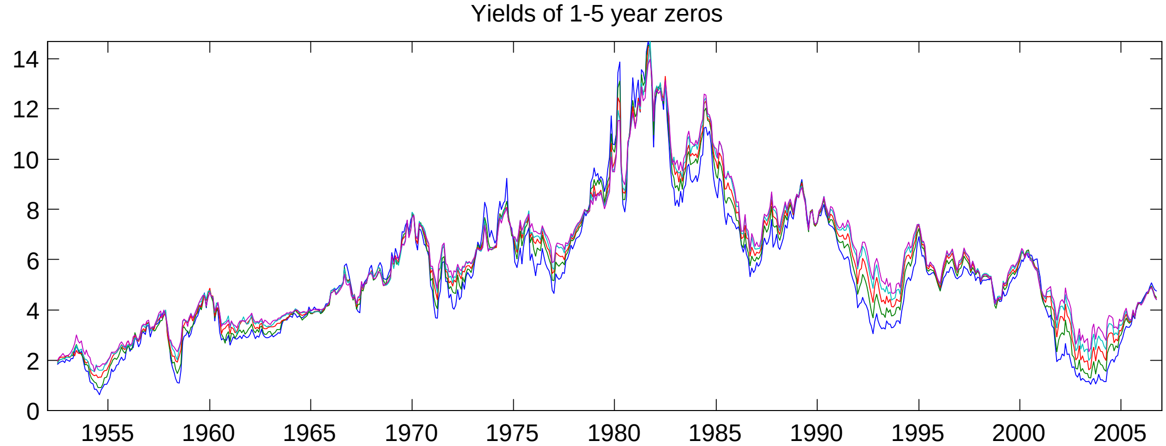
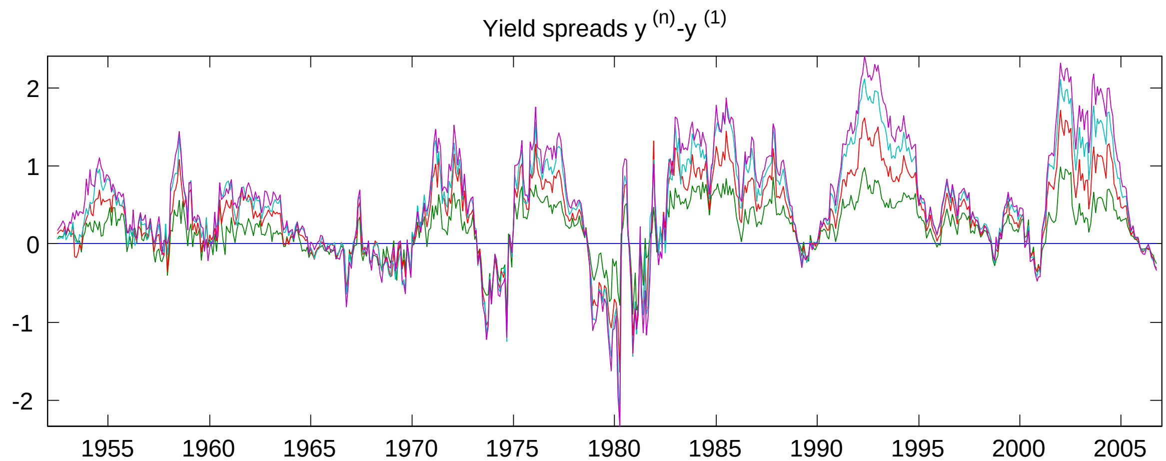
13.4.1 Factor models – eigevnalue principal components

Produces "factor model" What do data say about factor structure in the yield curve?





- Technique: eigenvalue decomposition

$$\begin{aligned} \Sigma &= cov(y) \\ Q\Lambda Q' &= \Sigma; \Lambda \text{ diagonal}, Q\prime Q = QQ\prime = I \end{aligned}$$

$$y_t = Qx_t; cov(xx') = \Lambda \rightarrow cov(yy') = Q\Lambda Q' = \Sigma$$

Thus, we get a factor model. The underlying factors x are uncorrelated, ordered by variance.

1. Λ *give us the variances of the "factors"*
2. *Columns of Q tell us how y loads on x movements, if "factor '1 moves" how much do y move.* $x = Q'y_t$ $(cov(x, x') = Q'\Sigma Q = Q'Q\Lambda Q'Q = \Lambda)$
3. *Colums of Q tell us how x is formed from each y, how to construct each factor.*

4. “Rotation, identification” There are many different ways to write $y = Ax$; $E(xx') = D$. For example, changing to $z_{1t} = x_{1t}/\sigma_1 + x_{2t}/\sigma_2$; $z_{2t} = x_{1t}/\sigma_1 - x_{2t}/\sigma_2$ preserves $cov(z_1, z_2) = 0$. The eigenvalue decomposition solves

$$\begin{aligned} \max var(q'y) \text{ s.t. } q'q &= 1 \\ \max var(q_2'y) \text{ s.t. } q_2'q_2 &= 1, q_2'q_1 = 0 \\ &\ldots \end{aligned}$$

Note: identification will be poor if the variance is about the same; small differences in the sample will produce q that jump between one and the other factor. Often you want to identify in other ways than by variance ordering, i.e. to get interpretable shapes of the loadings. For example, if you do this to the FF 25 portfolios, you get two factors, each of which is a mixture of smb and hml, and nearly the same variance. Rotating to smb and hml gives a more pleasing structure.

5. Unit variance factors

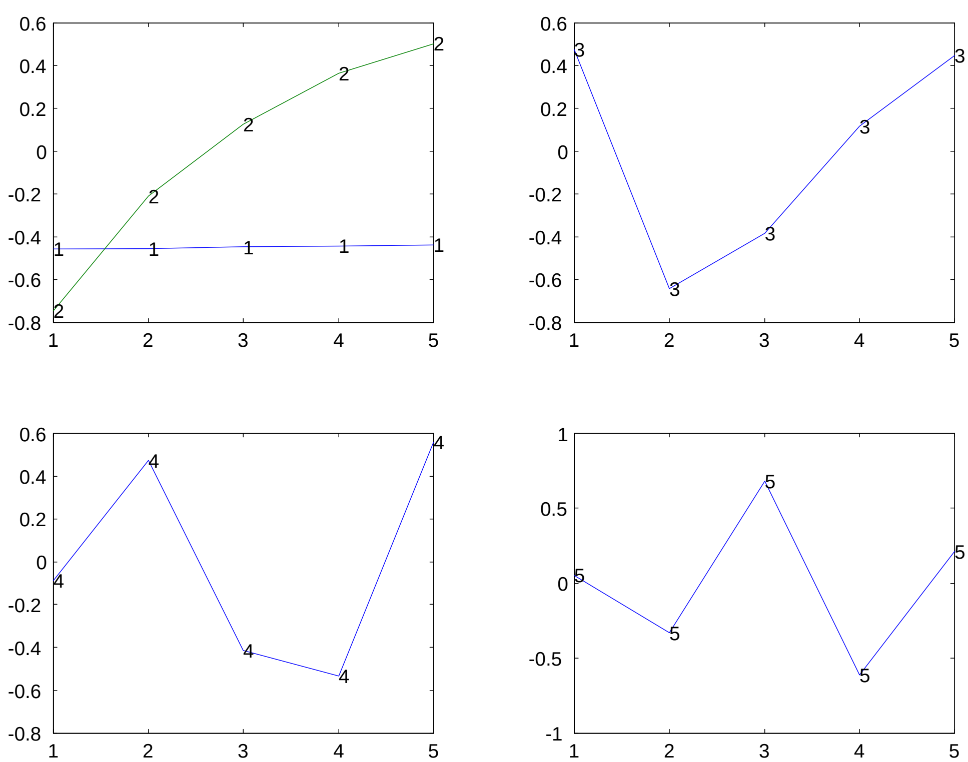
$$\begin{aligned} y_t &= Q\Lambda^{1/2}x_t; cov(x_t x_t') = I \\ x_t &= \Lambda^{-\frac{1}{2}}Q'y \end{aligned}$$

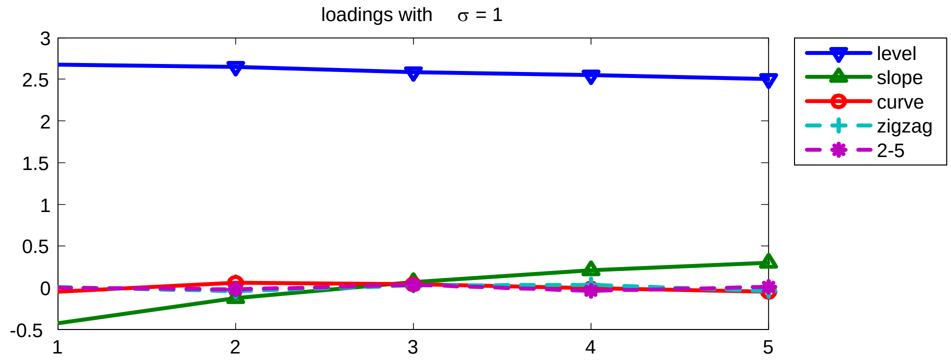
Loadings $Q\Lambda^{1/2}$ are smaller for smaller factors. This is a nice way to show the relative importance as well as the shapes of the factors, and not have your readers spend too much time interpreting the 5th factor.

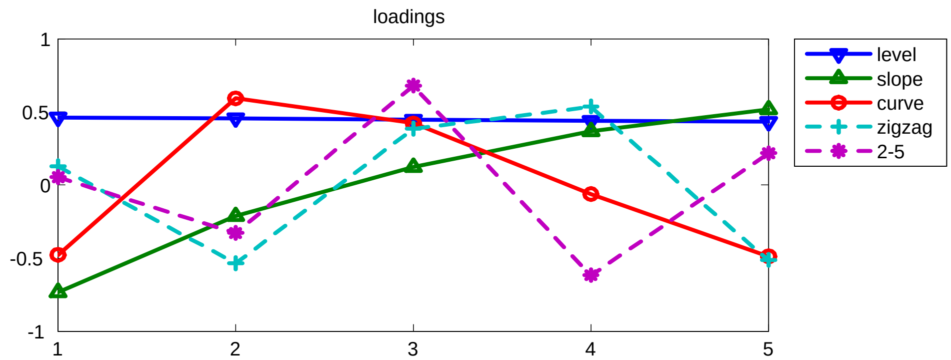
6. $\lambda_i / \sum \lambda_i =$ “fraction of variance explained by ith factor”

- Result, applied to FB yield data

factor x	1	2	3	4	5
$\sigma(x)$	5.75	0.56	0.10	0.08	0.06



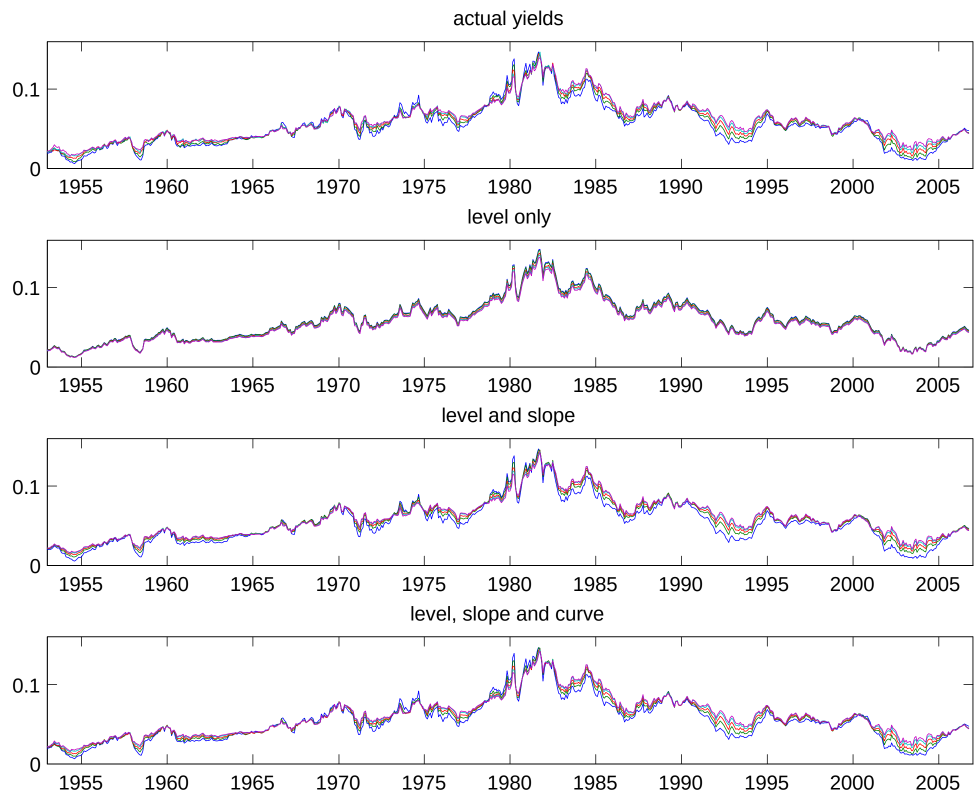


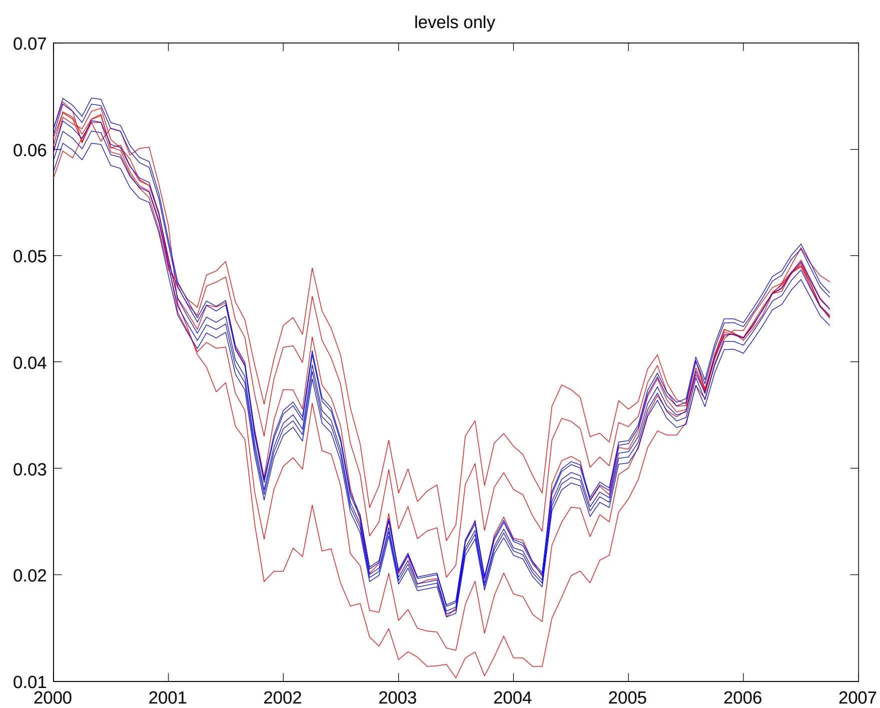


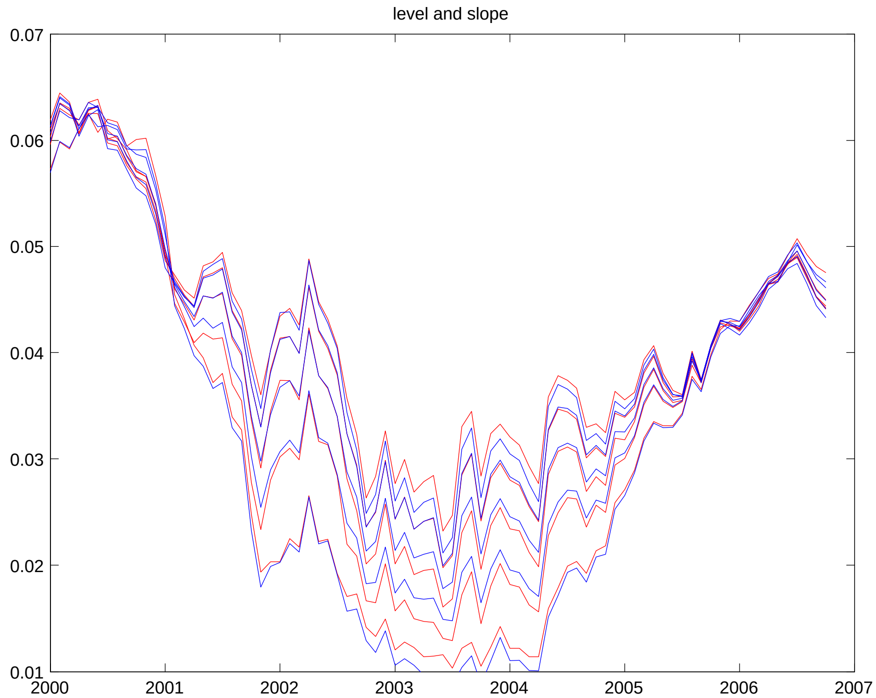
- "Factor models" come from dropping the small eigenvalues. Then a larger number of series are, exactly, driven by a smaller number of factors.

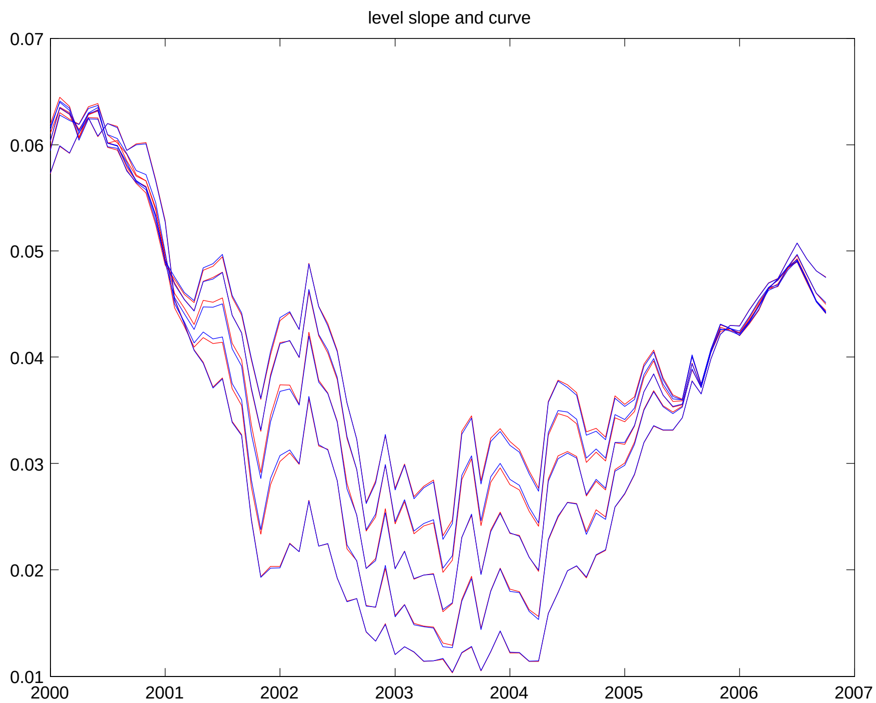
- For example, what if we drop 4 and 5?

$$\begin{bmatrix} y_t^{(1)} \\ y_t^{(2)} \\ y_t^{(3)} \\ y_t^{(4)} \\ y_t^{(5)} \end{bmatrix} \approx q_1 \times \text{level}_t + q_2 \times \text{slope}_t + \text{q}_3 \times \text{curve}_t$$


actual yields
level only
level and slope
level, slope and curve
0.1
0
1955
1960
1965
1970
1975
1980
1985
1990
1995
2000
2005


levels only
0.07
0.06
0.05
0.04
0.03
0.02
0.01
2000
2001
2002
2003
2004
2005
2006
2007





Movements in yields can be captured very well by movements in the first two - three factors alone. ***But not exactly!***

- Dropping factors II. Note that *the factors are uncorrelated with each other.* $cov(xx') = \Lambda$. Thus, the left out factors are uncorrelated with the factors you keep in.

$$y_t^{(n)} \approx q_1^{(n)} \times \text{level}_t + q_2^{(n)} \times \text{slope}_t + q_3^{(n)} \times \text{curve}_t + (\text{left out}_t)$$

Therefore, this is a regression equation! *This is a way of finding a regression model like FF3F when you don't know what to use on the right hand side.*

- Notice the analogy to FF3F: three factors (market, hml, smb) account for almost all return variation (R^2 above 90%). The factors are constructed as weighted combinations of the same securities.

13.5 Market price of risk and expected returns.

Data: expected excess returns *do* vary over time, EH is badly off.

$$hpr_{t+1}^{(n)} - y_t^{(1)} = a + b' y_t + \varepsilon_{t+1}$$

R^2 up to 40% and sign changes over time (Fama Bliss, Campbell Shiller, Cochrane Piazzesi).

Table 20.9. *Forecasts based on forward-spot spread*

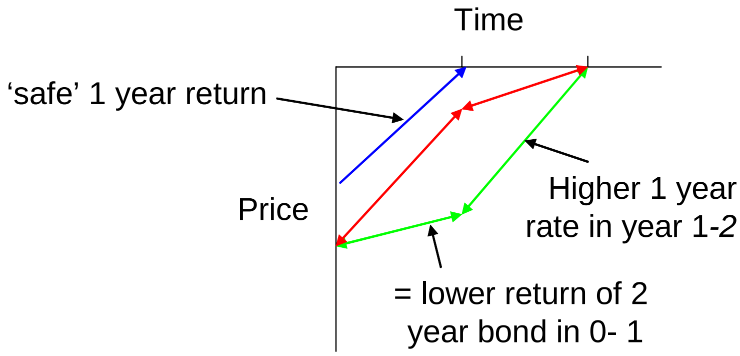
	Change in yields $y_{t+N}^{(1)} - y_t^{(1)} = a + b\left(f_t^{(N\to N+1)} - y_t^{(1)}\right) + \varepsilon_{t+N}$					Holding period returns $\text{hpr}_{t+1}^{(N+1)} - y_t^{(1)} = a + b\left(f_t^{(N\to N+1)} - y_t^{(1)}\right) + \varepsilon_{t+1}$				
N	a	$\sigma(a)$	b	$\sigma(b)$	$\overline{R}^2$	a	$\sigma(a)$	b	$\sigma(b)$	$\overline{R}^2$
1	0.1	0.3	−0.10	0.36	−0.02	−0.1	0.3	1.10	0.36	0.16
2	−0.01	0.4	0.37	0.33	0.005	−0.5	0.5	1.46	0.44	0.19
3	−0.04	0.5	0.41	0.33	0.013	−0.4	0.8	1.30	0.54	0.10
4	−0.3	0.5	0.77	0.31	0.11	−0.5	1.0	1.31	0.63	0.07

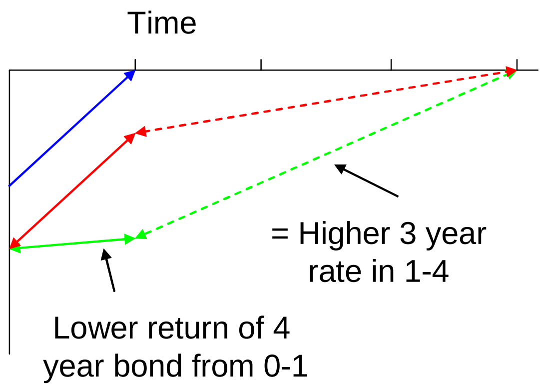
OLS regressions 1953–1997 annual data. Yields and returns in annual percentages. The left-hand panel runs the change in the one-year yield on the forward-spot spread. The right-hand panel runs the one-period excess return on the forward-spot spread

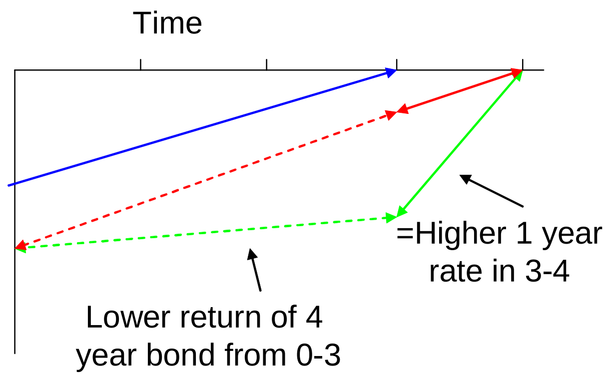
Figure 1:

(Updated 1964-2005)

	$r_{t+1}^{(n)} - y_t^{(1)} = a + b\left(f_t^{(n-1\to n)} - y_t^{(1)}\right) + \varepsilon_{t+1}$					$y_{t+n-1}^{(1)} - y_t^{(1)} = a + b\left(f_t^{(n-1\to n)} - y_t^{(1)}\right) + \varepsilon_{t+1}$				
n	a	b	$\sigma(a)$	$\sigma(b)$	R^2	a	b	$\sigma(a)$	$\sigma(b)$	R^2
2	0.04	0.91	0.28	0.26	0.14	-0.04	0.09	0.28	0.26	0.00
3	-0.15	1.20	0.50	0.35	0.15	-0.34	0.40	0.59	0.29	0.03
4	-0.37	1.41	0.70	0.44	0.16	-0.71	0.66	0.70	0.20	0.09
5	-0.09	1.10	0.95	0.52	0.07	-0.88	0.85	0.79	0.20	0.13
	forecasting *one* year returns on n-year bonds					forecasting *one* year rates n years from now				







Cochrane Piazzesi update

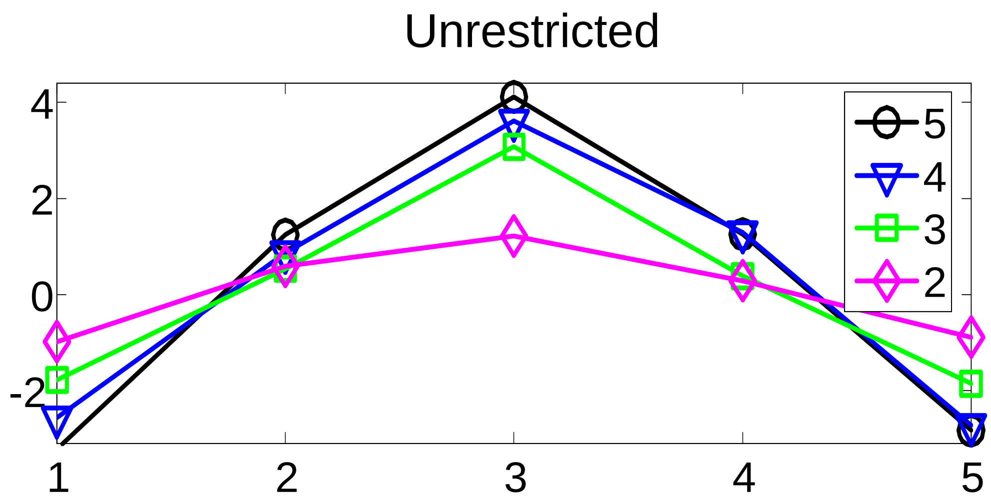
1. Bottom line

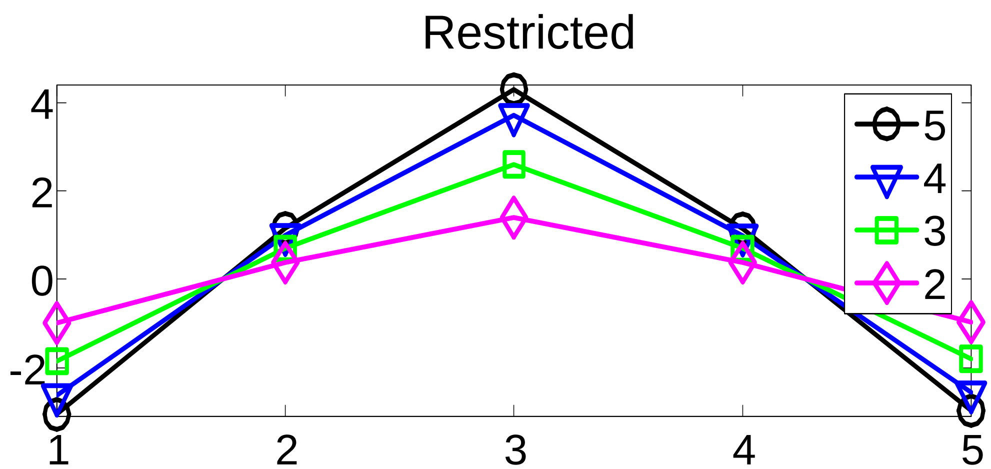
- Forecast 1 year treasury bond returns, over 1 year rate, using *all* forwards:

$$rx_{t+1}^{(n)} = a_n + \beta_n' f_t + \varepsilon_{t+1}^{(n)}$$

- R^2 up to 44%, up from Fama-Bliss / Campbell Shiller 15%
- A *single* "factor" $\gamma' f$ forecasts bonds of *all* maturities. High expected returns in "bad times."
- Tent-shaped factor is correlated with slope but is not slope. Improvement comes because it tells you when to bail out – when rates will rise in an upward-slope environment

Basic regression





$$rx_{t+1}^{(n)} = a_n + b_1 y_t^{(1)} + b_2 f_t^{(2)} + ... + b_5 f_t^{(5)} + \varepsilon_{t+1}^{(n)}$$

•Regressions of bond excess returns on *all* forward rates, not just matched $f - y$ as in Fama-Bliss

•The *same* linear combination of forward rates forecasts *all* maturities' returns. Just stretch the pattern more to get longer term bonds.

CP vs FB. FB want to understand what's in a forward rate, the right hand variable. CP want to understand what's in an expected return – left hand variable.

A single factor for expected bond returns

$$rx_{t+1}^{(n)} = b_n \left(\gamma_0 + \gamma_1 y_t^{(1)} + \gamma_2 f_t^{(1\to 2)} + ... + \gamma_5 f_t^{(4\to 5)} \right) + \varepsilon_{t+1}^{(n)}; \quad \frac{1}{4}\sum_{n=2}^{5} b_n = 1.$$

One common combination of forward rates $\gamma' f$, then stretch up and down more with b_n

•Two step estimation; first γ then b.

$$\overline{rx}_{t+1} = \frac{1}{4}\sum_{n=2}^{5} rx_{t+1}^{(n)} = \gamma_0 + \gamma_1 y_t^{(1)} + \gamma_2 f_t^{(1\to 2)} + ... + \gamma_5 f_t^{(4\to 5)} + \varepsilon_{t+1} = \gamma' f_t + \varepsilon_{t+1}$$

Then

$$rx_{t+1}^{(n)} = b_n\left(\gamma^\top f_t\right) + \varepsilon_{t+1}^{(n)}$$

Results:

Table 1 Estimates of the single-factor model

A. Estimates of the return-forecasting factor, $\overline{rx}_{t+1} = \gamma^\top f_t + \bar{\varepsilon}_{t+1}$

	γ_0	γ_1	γ_2	γ_3	γ_4	γ_5	R^2	$\chi^2(5)$
OLS estimates	-3.24	-2.14	0.81	3.00	0.80	-2.08	0.35	105.5

B. Individual-bond regressions

	Restricted		Unrestricted	
	$rx_{t+1}^{(n)} = b_n\left(\gamma^\top f_t\right) + \varepsilon_{t+1}^{(n)}$		$rx_{t+1}^{(n)} = \beta_n f_t + \varepsilon_{t+1}^{(n)}$	
n	b_n	R^2	R^2	$\chi^2(5)$
2	0.47	0.31	0.32	121.8
3	0.87	0.34	0.34	113.8
4	1.24	0.37	0.37	115.7
5	1.43	0.34	0.35	88.2

•γ capture tent shape.

•b_n increase steadily with maturity, stretch the tent shape out.

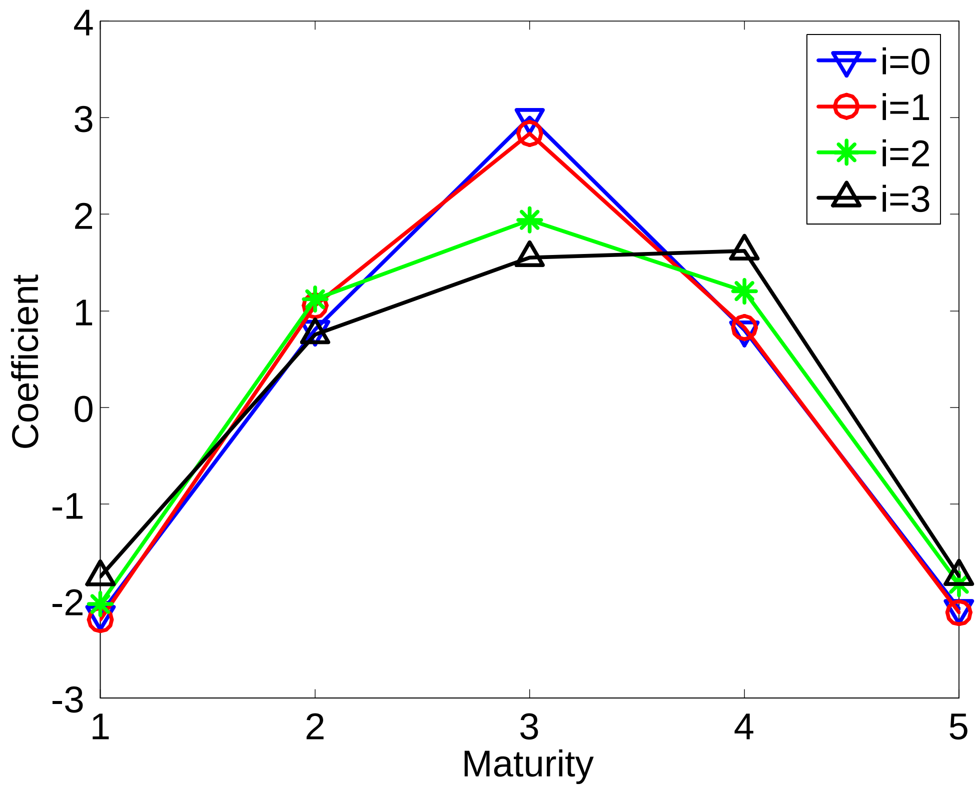
•Restricted model $b_n\gamma$ almost perfectly matches unrestricted coefficients. (well below 1σ)

•$R^2 = 0.34 - 0.37$ up from $0.15 - 0.17$. And we'll get to 0.44! Very significant rejection of $\gamma = 0$

•R^2 almost unaffected by the restriction. Restriction looks good in the graph.

•See paper version of table 1 for standard errors, joint tests including small sample, unit roots, etc. Bottom line: highly significant; EH is rejected, improvement on FB/3 factor models is significant.

More lags



$$rx_{t+1}^{(n)} = a_n + b_n' f_{t-i} + \varepsilon_{t+1}^{(n)}$$

•More lags are significant, with the same pattern.

•Checking individual lags reassures us it's not just measurement error, i.e.

$$p_{t+1} - p_t = a + bp_t + \varepsilon_{t+1}$$

if p_t is measured with error, you'll see something. But

$$p_{t+1} - p_t = a + bp_{t-1/12} + \varepsilon_{t+1}$$

fixes this problem.

•Suggests moving averages

$$rx_{t+1} = a + \gamma' \left(f_t + f_{t-1} + f_{t-2} + ... \right) \varepsilon_{t+1}$$

k	1	2	3	4	6
R^2	0.35	0.41	0.43	0.44	0.43

•Interpretation: Yields at t *should* should carry all information. If the lags enter, there must be a little measurement error. f change slowly over time, so $f_{t-1/12}$ is informative about the true f_t, .

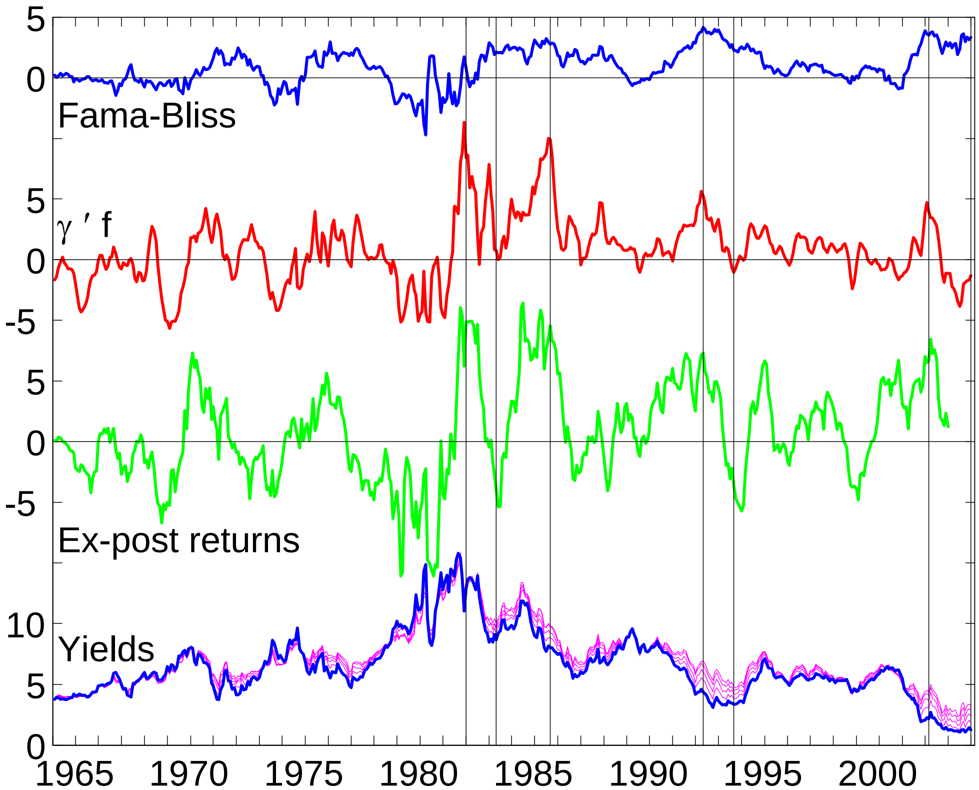
Stock Return Forecasts

Table 3. Forecasts of excess stock returns (VWNYSE)

$$\overline{rx}_{t+1} = a + bx_t + \varepsilon_{t+1}$$

	$\gamma^\top f$	(t)	d/p	(t)	$y^{(5)} - y^{(1)}$	(t)	R^2
	1.73	(2.20)					0.07
			3.56	(1.80)	3.29	(1.48)	0.08
	1.87	(2.38)			−0.58	(−0.20)	0.07
	1.49	(2.17)	2.64	(1.39)			0.10
MA $\gamma^\top f$	2.11	(3.39)					0.12
MA $\gamma^\top f$	2.23	(3.86)	1.95	(1.02)	−1.41	(−0.63)	0.15

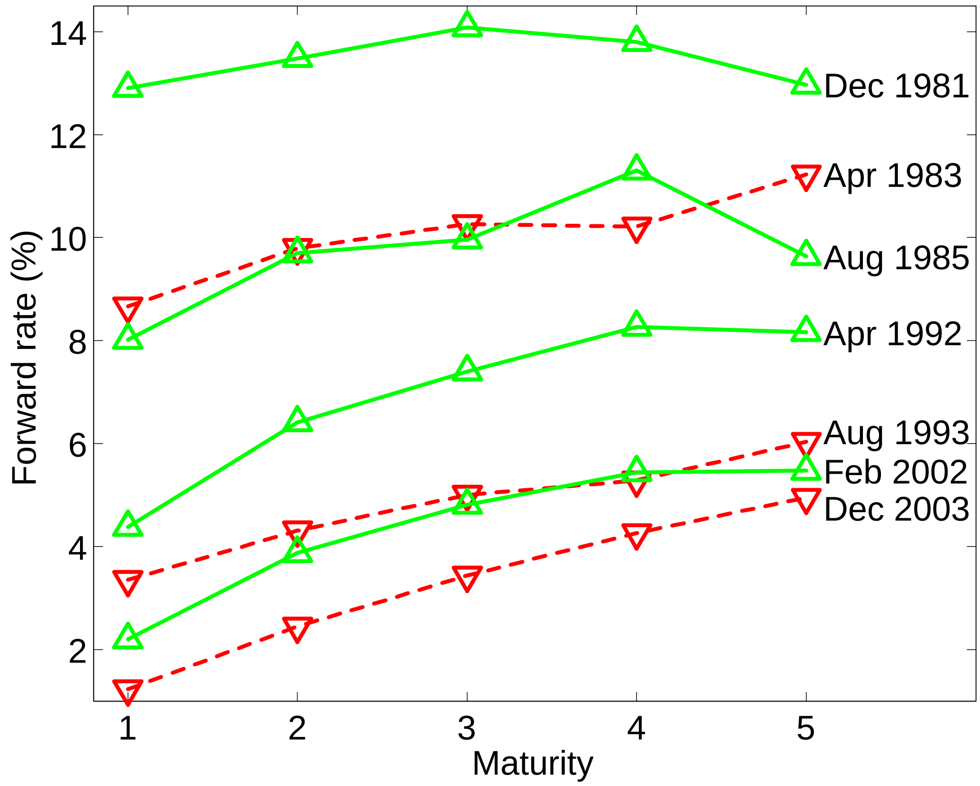
- 5 year bond had $b = 1.43$. Thus, $1.73 - 2.11$ is what you expect for a perpetuity.
- Does better than D/P and spread; Drives out spread; Survives with D/P
- A common term risk premium in stocks, bonds! Reassurance on fads & measurement errors



•Consistent in many episodes

•$\gamma' f$ and slope are correlated. Both show a rising yield curve but no rate rise

•$\gamma' f$ improvement in many episodes. $\gamma' f$ says get out in 1984, 1987, 1994, 2004. What's the signal?

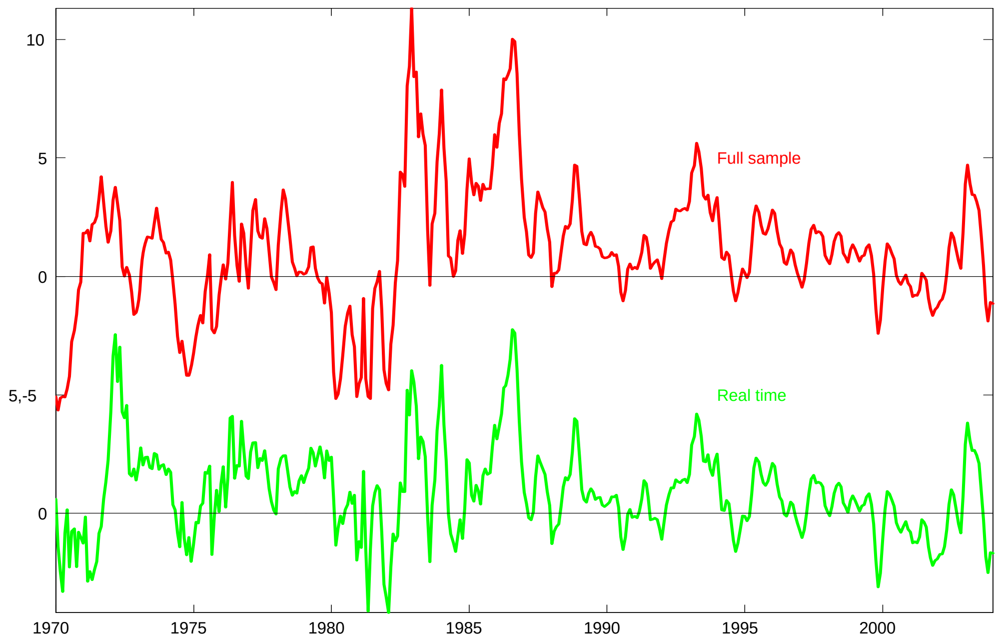


•Green: CP say go. Red: FB say go, CP say no.

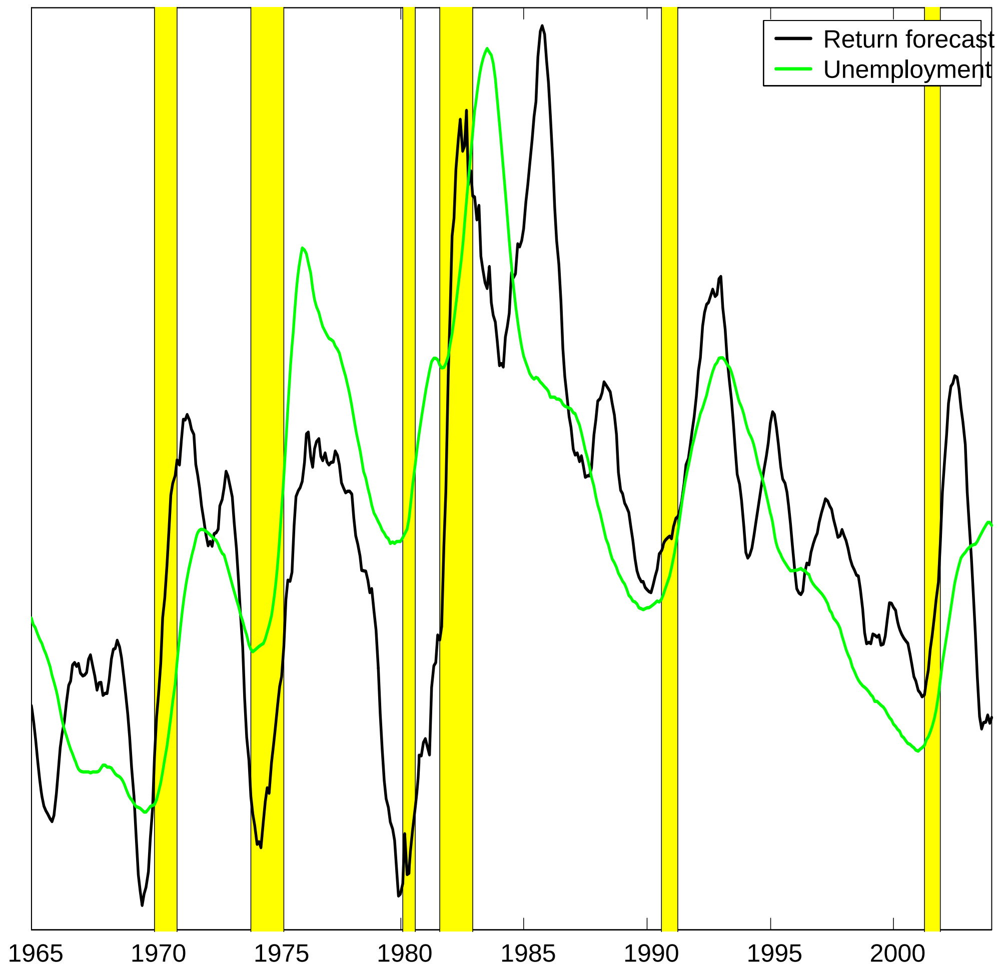
•Tent-shaped coefficients interact with tent-shaped forward curve to produce the signal.

•CP: in the past, tent-shape often came with upward slope. Others saw upward slope, thought that was the signal. But an upward slope without a tent does not work. The tent is the real signal.

Real time



Regression forecasts $\hat{\gamma}^\top f_t$. "Real-time" re-estimates the regression at each t from 1965 to t. (Note: Just because we don't have data before 1964 doesn't mean people don't know what's going on. Out of sample is not crucial, but it is interesting.)



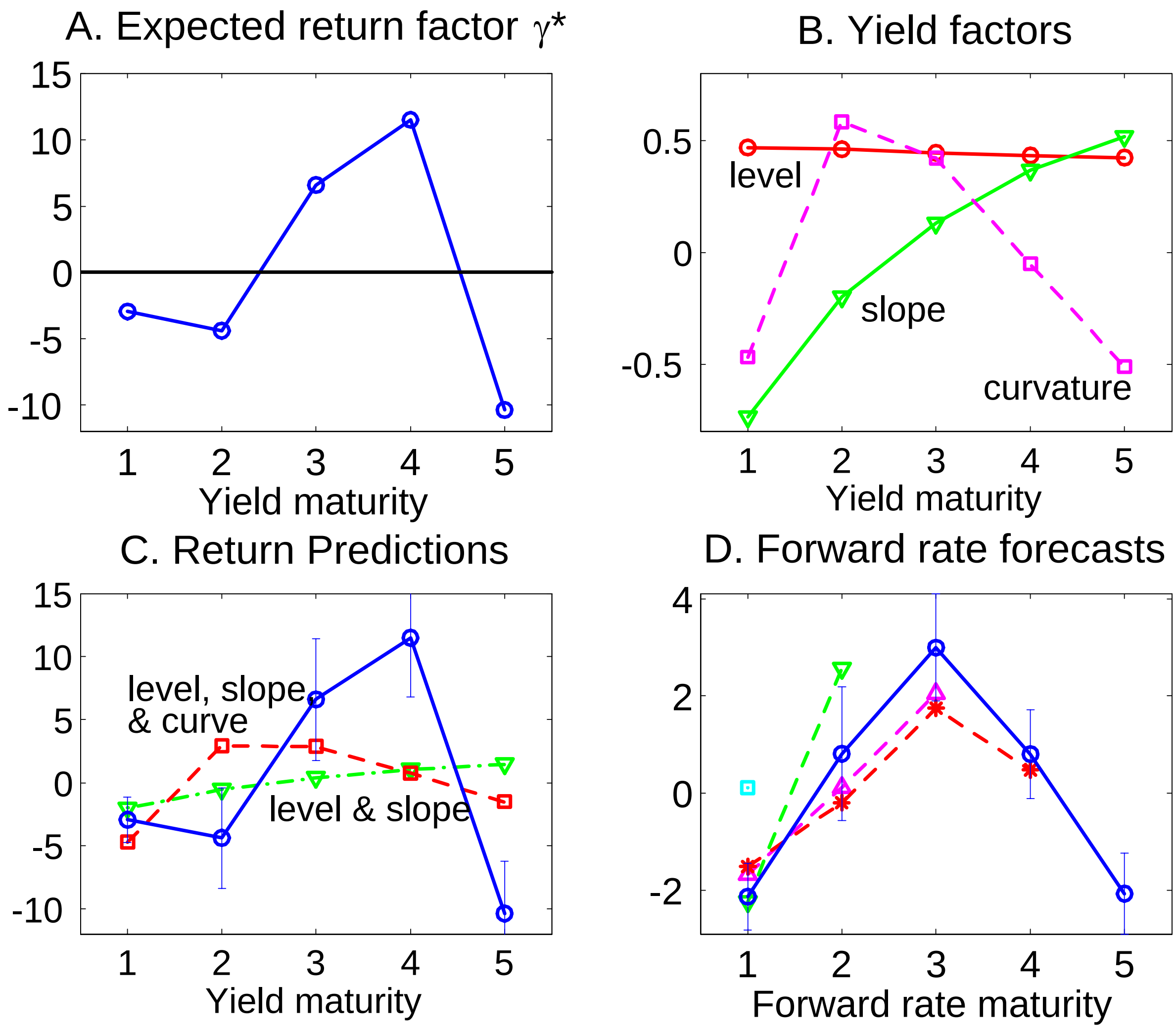
•Is it real, a time-varying risk premium? Or is it some new psychological "effect," an unexploited profit opportunity?

• Here, $\gamma' f$ is correlated with business cycles, and lower frequency. (Level, not growth.) Suggests a "business cycle related risk premium."

• Also significant that the same signal predicts all bonds, and predicts stocks. If "overlooked" it is common to a lot of markets!

Relation to factor models (why is this news?)

(Postpone until we understand factor models)



Panel A: We can express $\gamma' f$ as a function of yields too. $\gamma' f = \gamma^{*\prime} y$; What *yield* curve signals high returns on long term bonds (not *forward* curve)? A: γ^* $\approx$ Slope plus 4-5 spread.

Panel B: $\gamma' f$ has *nothing* to do with slope (symmetry: γ' linear $= 0$) and curvature (curved at the long, not short end). (These are the linear combinations of yields you use to create the factors. More next week.)

Panel C: $rx_{t+1} = a + b \times level_t + c \times slope_t + \varepsilon_{t+1}$. Then, from the construction $level_t = q' y_t$, you can figure out the implied regression coefficients of rx_{t+1} on y_t. The graph plots those implied coefficients. Moral: You can't approximate $\gamma' f$ well with level, slope, and curvature factors.

•Moral 1 Term structure models need L, S, C to get yield behavior *and* $\gamma' f$ to get expected returns.

•Adding $\gamma' f$ will not help much to hit yields (pricing errors) but it will help to get *tran-*

sition dynamics right (i.e. expected returns, yield *differences*)

•Moral 2. You can't *first* reduce to L, S, C, *then* examine $E_t rx_{t+1}$ →Reason #1 this was missed.

Panel D: The pattern is stable as we add forward rates.

Failures and spread trades

•What this *is* about (so far): when, overall is there a risk premium (high expected returns) in long term vs. short term bonds. "Trade" is just betting on long vs. short maturity, "betting on interest rate movements."

•What this is *not* about (so far). Much fixed income "arbitrage" involves relative pricing, small deviations from the yield curve. "Trade" might be short 30 year, long 29.5 year.

A hint of spread trades

If the one-factor model is exactly right, then *deviations from the single-factor model should not be predictable.*

$$rx_{t+1}^{(2)} - b_2 \overline{rx}_{t+1} = a^{(2)} + 0' f_t + \varepsilon_{t+1} = a^{(2)} + 0' y_t + \varepsilon_{t+1}$$

(Why?

$$\begin{aligned} rx_{t+1}^{(2)} &= \alpha^{(2)} + b_2 \left(\gamma' f_t\right) + \varepsilon_{t+1}^{(2)} \\ \overline{rx}_{t+1} &= \alpha + \gamma' f_t + \varepsilon_{t+1}^{(2)} \end{aligned}$$

multiply the second by b_2 and subtract.)

Table 7. Forecasting the failures of the single-factor model
A. Coefficients and t-statistics

		Right hand variable				
Left hand var.	const.	$y_t^{(1)}$	$y_t^{(2)}$	$y_t^{(3)}$	$y_t^{(4)}$	$y_t^{(5)}$
$rx_{t+1}^{(2)} - b_2\overline{rx}_{t+1}$	−0.11	−0.20	**0.80**	−0.30	−0.66	0.40
(t-stat)	(−0.75)	(−1.43)	**(2.19)**	(−0.90)	(−1.94)	(1.68)
$rx_{t+1}^{(3)} - b_3\overline{rx}_{t+1}$	0.14	0.23	−1.28	**2.36**	−1.01	−0.30
(t-stat)	(1.62)	(2.22)	(−5.29)	**(11.24)**	(−4.97)	(−2.26)
$rx_{t+1}^{(4)} - b_4\overline{rx}_{t+1}$	0.21	0.20	−0.06	−1.18	**1.84**	−0.82
(t-stat)	(2.33)	(2.39)	(−0.33)	(−8.45)	**(9.13)**	(−5.48)
$rx_{t+1}^{(5)} - b_5\overline{rx}_{t+1}$	−0.24	−0.23	0.55	−0.88	−0.17	**0.72**
(t-stat)	(−1.14)	(−1.06)	(1.14)	(−2.01)	(−0.42)	**(2.61)**

B. Regression statistics

Left hand var.	R^2	$\chi^2(5)$	$\sigma(\tilde{\gamma}^\top y)$	$\sigma(\text{lhs})$	$\sigma(b^{(n)}\gamma^\top y)$	$\sigma(rx_{t+1}^{(n)})$
$rx_{t+1}^{(2)} - b_2\overline{rx}_{t+1}$	0.15	41	0.17	0.43	1.12	1.93
$rx_{t+1}^{(3)} - b_3\overline{rx}_{t+1}$	0.37	151	0.21	0.34	2.09	3.53
$rx_{t+1}^{(4)} - b_4\overline{rx}_{t+1}$	0.33	193	0.18	0.30	2.98	4.90
$rx_{t+1}^{(5)} - b_5\overline{rx}_{t+1}$	0.12	32	0.21	0.61	3.45	6.00

• Pattern: if $y^{(n)}$ is a little out of line with the others (low price), then $r^{(n)}$ is good *relative* to all the others.

• No common factor. Bond-specific mean-reversion.

•This is *tiny.* 17-21 bp, compare to 200-600 bp returns.

•The single-factor $\gamma' f$ accounts for all the *economically important* variation in expected returns

• But the left hand side is tiny too, so tiny/tiny = good R^2

• Tiny isn't so tiny if you leverage up like crazy!

• But... measurement error looks the same.

13.6 Term structure models to capture risk premia

•How do we construct model to do this? λ_t but how...Idea:

$$\frac{E_t(R^e_{t+1})}{\sigma_t(R^e_{t+1})} = \frac{\sigma_t(m)}{E_t(m)}$$

we need conditionally heteroskedastic discount rates to produce time vayring risk premia.

•Framework: Multifactor discrete time Vasicek

$$X_{t+1} = \mu + \phi X_t + v_{t+1};\ v_{t+1} \sim \mathcal{N}(0, V).$$

$$\begin{aligned} M_{t+1} &= \exp\left(-\delta_0 - \delta_1^\top X_t - \frac{1}{2}\lambda_t^\top V \lambda_t - \lambda_t^\top v_{t+1}\right) \\ \lambda_t &= \lambda_0 + \lambda_1 X_t. \end{aligned}$$

$$\begin{aligned} P_t^{(1)} &= E_t(M_{t+1}) = \exp\left(-\delta_0 - \delta_1^\top X_t\right) \\ y_T^{(1)} &= \delta_0 + \delta_1^\top X_t \end{aligned}$$

Fact:

$$p_t^{(n)} = A_n - B_n' X_t$$

Just as before, guess then

$$e^{p_t^{(n)}} = E_t\left(M_{t+1} e^{p_{t+1}^{(n-1)}}\right)$$

....

Simplest answer

$$\begin{aligned} \phi^* &\equiv \phi - V\lambda_1 \\ \mu^* &\equiv \mu - V\lambda_0 \end{aligned}$$

$$\begin{aligned} B_n^\top &= -\delta_1^\top (I + \phi^* + \phi^{*2} + .. + \phi^{*n-1}) = -\delta_1^\top \left(I - \phi^{*n}\right)\left(I - \phi^*\right)^{-1} \\ f_t^{(n-1\to n)} &= \left(\delta_0 - B_{n-1}^\top \mu^* - \frac{1}{2} B_{n-1}^\top V B_{n-1}\right) + \left(\delta_1^\top \phi^{*n-1}\right) X_t. \end{aligned}$$

Compare to the earlier discrete time Vasicek – $\lambda_1 = 0$. There had only a *constant* distortion. Now a time-varying distortion as well. "risk neutral" transition matrix

(long algebra)

$$E_t r_{t+1}^{(n)} - y_t^{(1)} = B_{n-1}^{\top} V \left(\lambda_0 - \frac{1}{2} B_{n-1} \right) + \left(B_{n-1}^{\top} V \lambda_1 \right) X_t$$

.

•*How do we pick* λ_0, λ_1 to match CP or other facts? (Said to be a big puzzle!)

"find a discount factor to price a set of assets with given mean, covariance matrix" – child's play!

$$R_{t+1}^e = a + bX_t + \varepsilon_{t+1}$$
$$M_{t+1} = 1/R^f \left\{ 1 - E_t \left(R_{t+1}^{e\prime} \right) \Sigma^{-1} \left[R_{t+1}^e - E_t(R_{t+1}^e) \right] \right\}$$
$$E_t(MR^f) = 1$$
$$E_t \left(M_t R^{e\prime} \right) = 1/R^f \left(E_t(R^{e\prime}) - E_t \left(R^{e\prime} \right) \Sigma^{-1} \Sigma \right) = 0$$
$$M_{t+1} = 1/R^f \left\{ 1 - (a + bX_t)' \, \Sigma^{-1} \varepsilon_{t+1} \right\}$$

i.e.

$$\lambda_t = \Sigma^{-1} (a + bX_t) \, .$$

All we need to do is adapt this idea to the lognormal world.

•Step 1: 1 = E(MR) in this lognormal world

$$E_t \left[rx_{t+1}^{(n)} \right] + \frac{1}{2} \sigma_t^2 (rx_{t+1}^{(n)}) = -cov_t(rx_{t+1}^{(n)}, m_{t+1}),$$

Proof:

$$1 = E_t \left[M_{t+1} R_{t+1}^{(n)} \right] = E_t \left[e^{m_{t+1} + r_{t+1}^{(n)}} \right]$$
$$0 = E_t \left[m_{t+1} \right] + E_t \left[r_{t+1}^{(n)} \right] + \frac{1}{2} \sigma_t^2 (m_{t+1}) + \frac{1}{2} \sigma_t^2 (r_{t+1}^{(n)}) + cov_t(r_{t+1}^{(n)}, m_{t+1})$$

$$0 = E_t \left[m_{t+1} \right] + y_t^{(1)} + \frac{1}{2} \sigma_t^2 (m_{t+1}),$$
$$E_t \left[rx_{t+1}^{(n)} \right] + \frac{1}{2} \sigma_t^2 (rx_{t+1}^{(n)}) = -cov_t(rx_{t+1}^{(n)}, m_{t+1}),$$

•Fact we want to match:

$$rx_{t+1} = \beta f_t + \varepsilon_{t+1}; \;\; cov(\varepsilon_{t+1} \varepsilon_{t+1}^{\top}) = \Sigma.$$

•Idea. Load on ε shocks,

$$M_{t+1} = \exp \left(-y_t^{(1)} - \frac{1}{2} \lambda_t^{\top} \Sigma \lambda_t - \lambda_t^{\top} \varepsilon_{t+1} \right) ?$$

$$\begin{aligned} E_t \left[rx_{t+1} \right] + \frac{1}{2} \sigma_t^2 (rx_{t+1}) &= \beta f_t + diag(\Sigma) \\ &= -cov_t(rx_{t+1}^{(n)}, m_{t+1}) = cov_t(rx_{t+1}^{(n)}, \varepsilon_{t+1}') \lambda_t = \Sigma \lambda_t \end{aligned}$$

$$\lambda_t = \Sigma^{-1}\left[\beta f_t + \frac{1}{2}diag\left(\Sigma\right)\right]$$

• *We're done!* Answer:

$$rx_{t+1} = \beta f_t + \varepsilon_{t+1};\ cov(\varepsilon_{t+1}\varepsilon_{t+1}^{\top}) = \Sigma.$$

$$M_{t+1} = \exp\left(-y_t^{(1)} - \frac{1}{2}\lambda_t^{\top}\Sigma\lambda_t - \lambda_t^{\top}\varepsilon_{t+1}\right),$$

$$\lambda_t = \Sigma^{-1}\left[\beta f_t + \frac{1}{2}diag\left(\Sigma\right)\right].$$

Wait, what about the state variable transition? The first equation isn't in standard form. Oh yes it is! Recall

$$rx_{t+1}^{(n)} = p_{t+1}^{(n-1)} - p_t^{(n)} + p_t^{(1)}$$

so this rx forecasting equation IS

$$f_{t+1} = \mu + \phi f_t + Q\varepsilon_{t+1}$$

or yields or prices.

• *Check.* Before, I had the idea of using $y_t^{(1)}$ as a "state" variable and then checking that $P_t^{(1)} = E(M_{t+1})$ generates the $y_t^{(1)}$ we started with. We need to do that again here...

$$P_t^{(N)} = E_t(M_{t+1}...M_{t+N})$$

(algebra). Yes, this works, the model is "self-consistent."

• This is exsistence... much more to "what it looks like" see CPII. Interesting stuff, like "the risk premium is earned for covariance with level shocks."

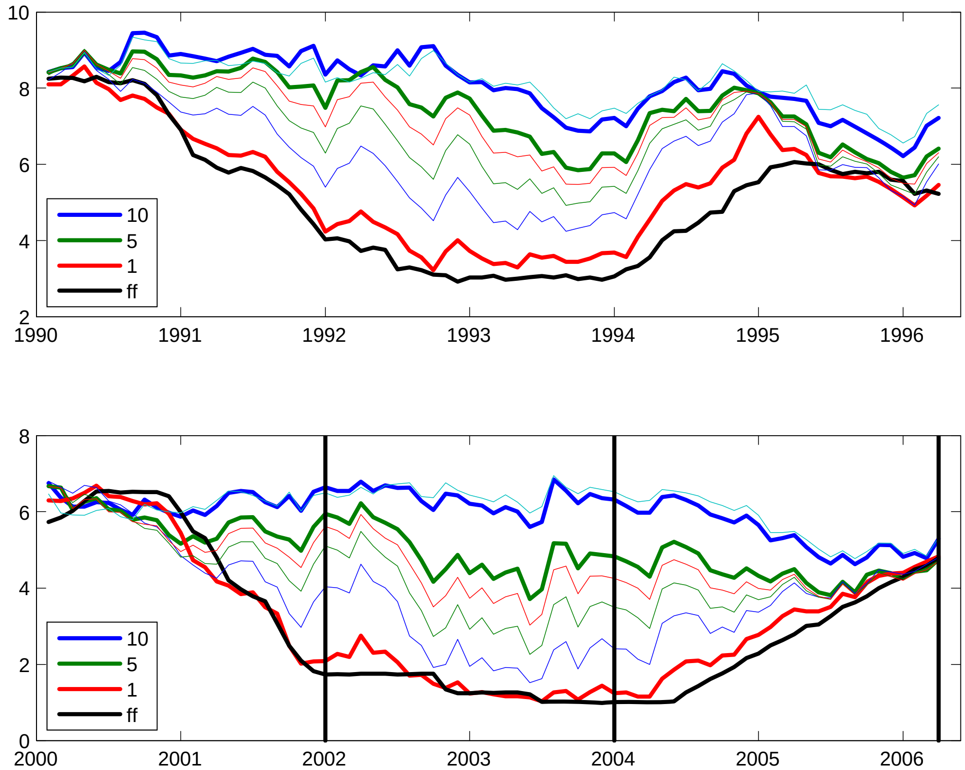
•What is $E_t\left(\gamma' f_{t+1}\right) = E_t rx_{t+2}^{(n)}$. What is the *term structure of risk premia?*

$$f_t^{(n)} = E_t\left(y_{t+n-1}^{(1)}\right) + E_t\left(rx_{t+1}^{(n)} - rx_{t+1}^{(n-1)}\right) + E_t\left(rx_{t+2}^{(n-1)} - rx_{t+2}^{(n-2)}\right) + \ldots + E_t\left(rx_{t+n-1}^{(2)}\right)$$

$$f_t^{(n)} = E_t\left(y_{t+n-1}^{(1)}\right) + \left(b_n - b_{n-1}\right)\left(\gamma' f_t\right) + \left(b_{n-1} - b_{n-2}\right)E_t\left(\gamma' f_{t+1}\right)\ldots$$

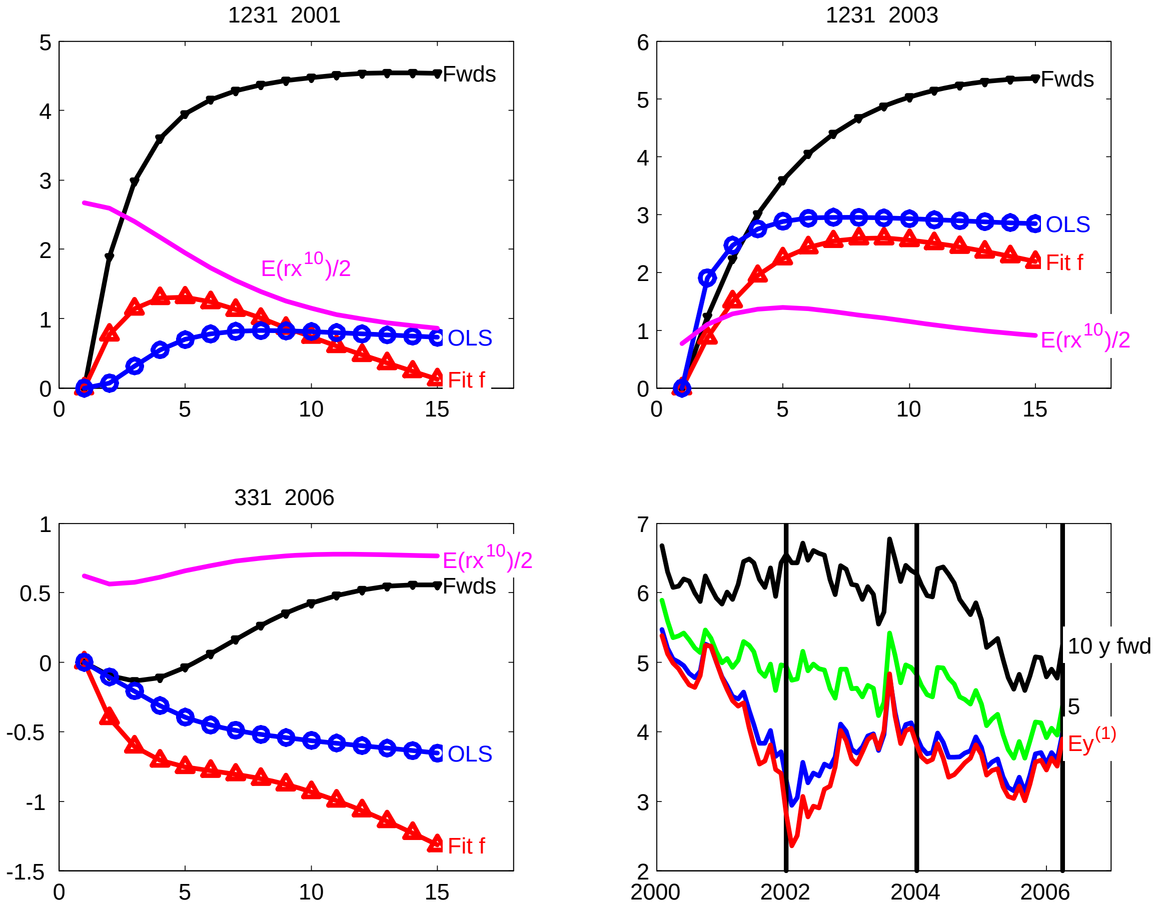
Geometric decay or interesting dynamics? Is it all the first year premium, or is premium expected to stay high "random walk"? "forward rate = expectations plus risk premium" sure, but current or future risk premium? CP II starts answer, much more to do !

Task: Understand recent interest rates. Conundrum, or replay the past?



Forward rates in two recessions. The federal funds rate, 1-5, 10 and 15 year forward rates are plotted. Federal funds, 1, 5 and 10 year forwards are emphasized. The vertical lines in the lower panel highlight specific dates that we analyze more closely below.

Two guesses: A cointegrated estimate (no "yields revert to uncdonditional mean") a) risk premium high, then declining (zero if graphs parallel)



Forward curve decompositions by the affine model in the spread-difference specification.

13.7 Evolution of term structure models

•Vasicek

$$
\begin{aligned}
dr &= \phi(\bar{r} - r)dt + \sigma dz \\
\frac{d\Lambda}{\Lambda} &= -rdt - \sigma_\Lambda dz
\end{aligned}
$$

$$p(N, r) = -A(N) - B(N)r$$

$$
\begin{aligned}
y_t^{(1)} - \delta &= \phi\left(y_{t-1}^{(1)} - \delta\right) + \varepsilon_t \\
m_t &= \log M_t = -y_t^{(1)} - \frac{1}{2}\lambda^2\sigma_\varepsilon^2 - \lambda\varepsilon_t
\end{aligned}
$$

B(N) is just expectations, no time varying risk premium. Single factor. Constant volatility

•CIR

$$
\begin{aligned}
dr &= \phi(\bar{r} - r)dt + \sigma\sqrt{r}dz \\
\frac{d\Lambda}{\Lambda} &= -rdt - \sigma_\Lambda\sqrt{r}dz
\end{aligned}
$$

Keeps $r > 0$. Gives some (but not enough) TV risk premium. Allows some TV volatility, but tied to level

•Multifactor affine

$$\begin{aligned}
dy &= \phi(\bar{y} - y)dt + \Sigma dw \\
r &= \delta_0 - \delta' y \\
\frac{d\Lambda}{\Lambda} &= -rdt - \sigma'_\Lambda dw \\
dw_i &= \sqrt{\alpha_i + \beta'_i y dz_i}
\end{aligned}$$

Result,

$$P(N, y) = e^{-A(N) - B(N)'y}$$

Allows more time-varying risk premium (but not enough). Allows volatility to vary as a function of rate spreads too.

Current issues

• Better volatility,

• Fact 1 : there is more volatility when r high. (80s?) fact 2: there is a lot of transient r volatility (weeks), seen in realized volatility , and short-dated term strucutre options move a lot. fact 3: transient r volatility does not affect bond prices much. Also, r^2_{t+1} is not well forecast by y (RB).

• Like stochatic volatility in options

$$\begin{aligned}
dr &= \mu_r(r)dt + \sigma_r(v)dz \\
dv &= \mu_v dt + \sigma_v dw \\
\frac{d\Lambda}{\Lambda} &= -rdt - \sigma_{\Lambda r} dz - \sigma_{\Lambda v} dw \\
& \quad p(r, v, n)
\end{aligned}$$

a) puzzle, – p should be a function of any state variable – we *should* be able to see volatility in bond prices!

b) In general $p(r, v, n)$ is not a strong function of v. And cases can be worked out where $dp/dv = 0$ – exact “unspanned stochastic volatility”.

c)

$$P = Ee^{-\int_{s=0}^{n} r_{t+s} ds} = e^{-\int_{s=0}^{n} Er_{t+s} ds + \frac{1}{2}\sigma^2 \left[\int_{s=0}^{n} r_{t+s}\right]}$$

$1/2\sigma^2$ isn’t that big anyway, and *short lasting* changes aren’t going to affect bond prices much.

Hedging volatility sensitive term structure options with treasury zeros doesn’t work that well.

Solution: calibrate model to options directly.

Is it an arbitrage, in the payoff space? Same issue as BS. 1) If it’s *really* a factor model, and *really* not unspanned exactly, then yes. 2) If not... Just because it’s “close” for bonds does not mean “close” for options!

But then.... why are we doing an affine model and fitting to bond/swap data?

• Better mean? –CP

In discrete-time multifactor vasicek, we saw how to *construct* $\lambda_t = \lambda_0 + \lambda_1 f_t$ to match return regressions.

• Merge better mean and better volatility?

• "observable macro factors z_t" But then why do bond prices not reveal the macro factors and render them useless? $p(..z_t)$? Could this be generalized to have unspanned macro factors as well?

• economic models of risk premia,

$$M_t = \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma}$$

with money, etc.? "new keynesian affine term structure models.'