

23 Week 9. Term structure I Expectations Hypothesis and Bond Risk Premia – Overheads

23.0.1 Definitions

1. *Discount (zero-coupon) bond*: Promise to pay \$1 at time $t + n$.

2. *Price*

$$P_t^{(n)} = \text{price of } n \text{ year discount bond at time } t$$

$$\text{Example. } P_t^{(2)} = 0.9 \iff 10\% \text{ interest rate}$$

$$p_t^{(n)} = \log \text{ price of } n\text{-year discount bond at time } t.$$

$$\text{Example: } \ln(0.9) = -0.10536 \approx -0.1 \text{ or "10\% discount"}$$

3. *Yield*. For discount bonds,

$$P_t^{(n)} = \frac{1}{[Y_t^{(n)}]^n} \leftrightarrow Y_t^{(n)} \equiv [P_t^{(n)}]^{-\frac{1}{n}};$$

$$p_t^{(n)} = -ny_t^{(n)} \leftrightarrow y_t^{(n)} \equiv -\frac{1}{n}p_t^{(n)}.$$

$$\text{Example: } p_t^{(2)} = -0.1 \rightarrow y_t^{(2)} = \frac{0.1}{2} = 0.05, \text{ "5\% discount per year."}$$

4. *Forward rate*.

(a) Definition: The rate at which you can contract today to borrow from time $t + n - 1$ and pay back at time $t + n$.

(b) Fact:

$$f_t^{(n)} = f_t^{(n-1 \rightarrow n)} \equiv p_t^{(n-1)} - p_t^{(n)}$$

(c) Example.

$$p_t^{(3)} = -0.15; p_t^{(2)} = -0.10 \rightarrow f_t^{(2 \rightarrow 3)} = 0.05 = 5\%$$

5. *Holding period return*.

(a) Definition (words). Buy an n -year bond at time t and sell it – now an $n - 1$ year bond – at time $t + 1$.

(b) Definition (equations)

$$R_{t+1}^{(n)} = \frac{P_{t+1}^{(n-1)}}{P_t^{(n)}}$$

$$r_{t+1}^{(n)} \equiv p_{t+1}^{(n-1)} - p_t^{(n)}.$$

(c) Excess log returns (over the risk free rate)

$$rx_{t+1}^{(n)} \equiv r_{t+1}^{(n)} - y_t^{(1)}$$

6. Summary:

$$\begin{aligned}
 p_t^{(n)} &= \log(P_t^{(n)}) \\
 y_t^{(n)} &= -\frac{1}{n}p_t^{(n)} \\
 f_t^{(n)} &= p_t^{(n-1)} - p_t^{(n)} \\
 r_t^{(n)} &= p_{t+1}^{(n-1)} - p_t^{(n)} \\
 rx_{t+1}^{(n)} &= r_{t+1}^{(n)} - y_t^{(1)}.
 \end{aligned}$$

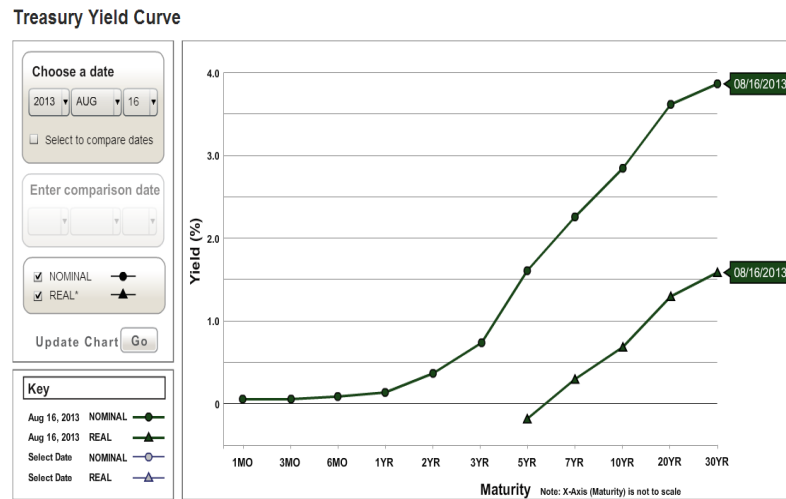
7. Notation: (n) = maturity. t = time, when observed. To be clear,

$$r_{t+1}^{(n)} = r_{t \rightarrow t+1}^{(n \rightarrow n-1)}$$

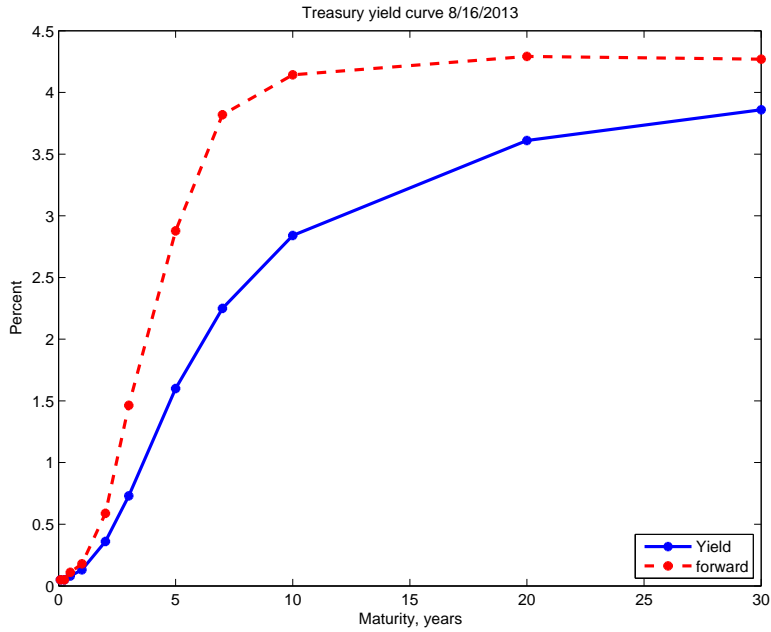
23.1 Expectations hypothesis

- Question: Why does the yield curve sometimes slope up, sometimes slope down?

1. Yield curves¹⁵



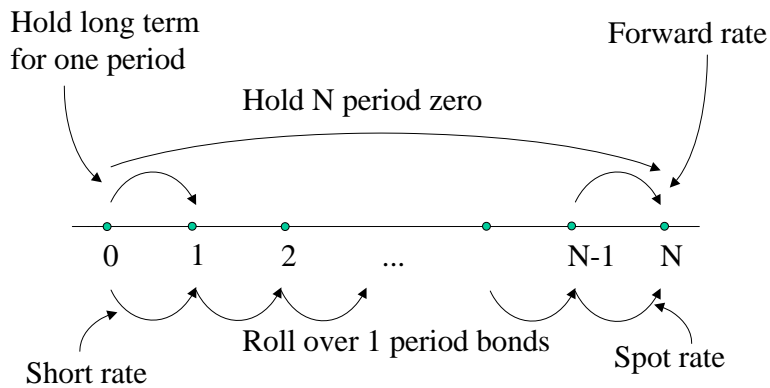
¹⁵It's a lot of fun to play with the website <http://www.treasury.gov/resource-center/data-chart-center/interest-rates/Pages/Historic-Yield-Data-Visualization.aspx>



2. Why are yields different for bonds with different maturity? Three equivalent statements:

- (a) *Long maturity yield = average of expected future short rates (plus risk premium)*
- (b) *Forward rate = expected future spot rate (plus risk premium)*
- (c) *Expected holding period returns should be equal across maturities (plus risk premium)*

3. Big picture



(a) (0) to N Long yield = average of expected future short rates

$$\begin{aligned}
 &\text{Buy } N \text{ year zero} && \text{Rollover} \\
 r_{0 \rightarrow N} = -p^{(N)} = N y^{(N)} &&& y_0^{(1)} + y_1^{(1)} + y_2^{(1)} + \dots + y_{N-1}^{(1)} \\
 y_0^{(N)} = \frac{1}{N} E \left(y_0^{(1)} + y_1^{(1)} + y_2^{(1)} + \dots + y_{N-1}^{(1)} \right) &&& (+\text{risk premium})
 \end{aligned}$$

(b) $N-1$ to N : Forward rate = expected future spot rate

lock in = $E(\text{wait and use spot})$ (+risk premium)

$$f_t^{(N)} = E_t \left[y_{t+N-1}^{(1)} \right] \text{ (+risk premium)}$$

(c) 0 to 1 Expected returns are the same at all maturities

hold N period bond 1 year = hold 1 period bond for 1 year

$$E_t \left[r_{t \rightarrow t+1}^{(N)} \right] = E_t \left[r_{t \rightarrow t+1}^{(1)} \right] = y_t^{(1)} \text{ (+risk premium)}$$

$$E_t \left[r x_{t \rightarrow t+1}^{(N)} \right] = E_t \left[r x_{t \rightarrow t+1}^{(M)} \right] = 0 \text{ (+risk premium)}$$

4. Example: $y_t^{(1)} = 5\%$, $y_t^{(2)} = 10\%$. Find and understand

(a) $p_t^{(1)} = -0.05$ $p_t^{(2)} = -0.20$

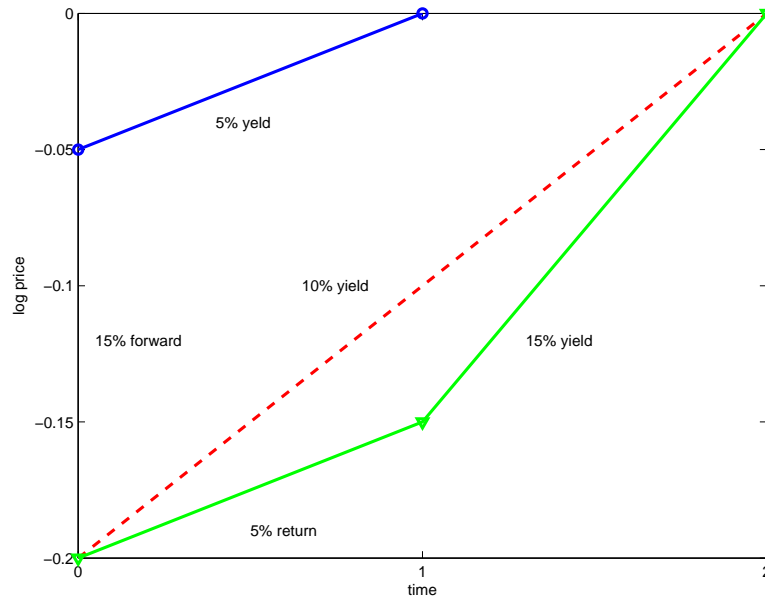
(b) $f_t^{(2)} = 15\%$

(c) $E_t r_{t+1}^{(2)} = y_t^{(1)} = 5\%$

(d) $E_t y_{t+1}^{(1)} = 15\%$

(e) Yield, forward, and return statements of the EH are all equivalent.

(f) This is a “hypothesis” about expectations, not a truth about how things will come out or about actual expectations!



5. Exchange rates and interest spreads.

(a) Realized return to US investor = foreign interest + exchange rate depreciation ($s_t = \log$

$$r_{t \rightarrow t+1}^{US} = r_{t \rightarrow t+1}^{Euro} + s_t - s_{t+1}$$

(b) Expectations hypothesis:

$$E_t(r_{t+1}^{US}) = E_t(r_{t+1}^{Euro}) + s_t - E_t(s_{t+1})$$

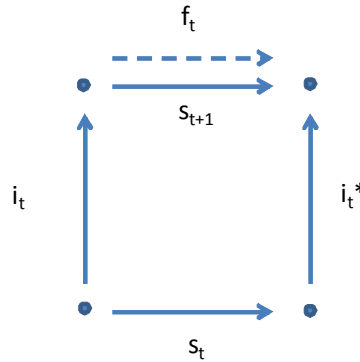
(c) For one-period rates, $y_t^{(1)} = E_t(r_{t+1})$ so

$$\begin{aligned} y_t^{(1)Eu} - y_t^{(1)US} &= E_t(s_{t+1}) - s_t \\ i_t^* - i_t &= E_t(s_{t+1}) - s_t \end{aligned}$$

Interest rate spread = expected depreciation of the euro

6. Exchange rates and forward rates.

(a) Arbitrage, “Covered interest parity”



$$i_t + f_t - i_t^* - s_t = 0$$

$$i_t^* - i_t = f_t - s_t$$

(b) “Uncovered interest parity” = expectations hypothesis

$$E_t(s_{t+1}) - s_t = i_t^* - i_t = f_t - s_t$$

23.2 Risk premia

1. What risk premium do we expect? Which end of these ways to get money from x to y is riskier?

Summary: The sign of risk premium can go either way; it depends on investor’s horizon relative to supply of bonds, and whether real interest rates or inflation are the source of interest rate risk.

2. Theory II: as always,

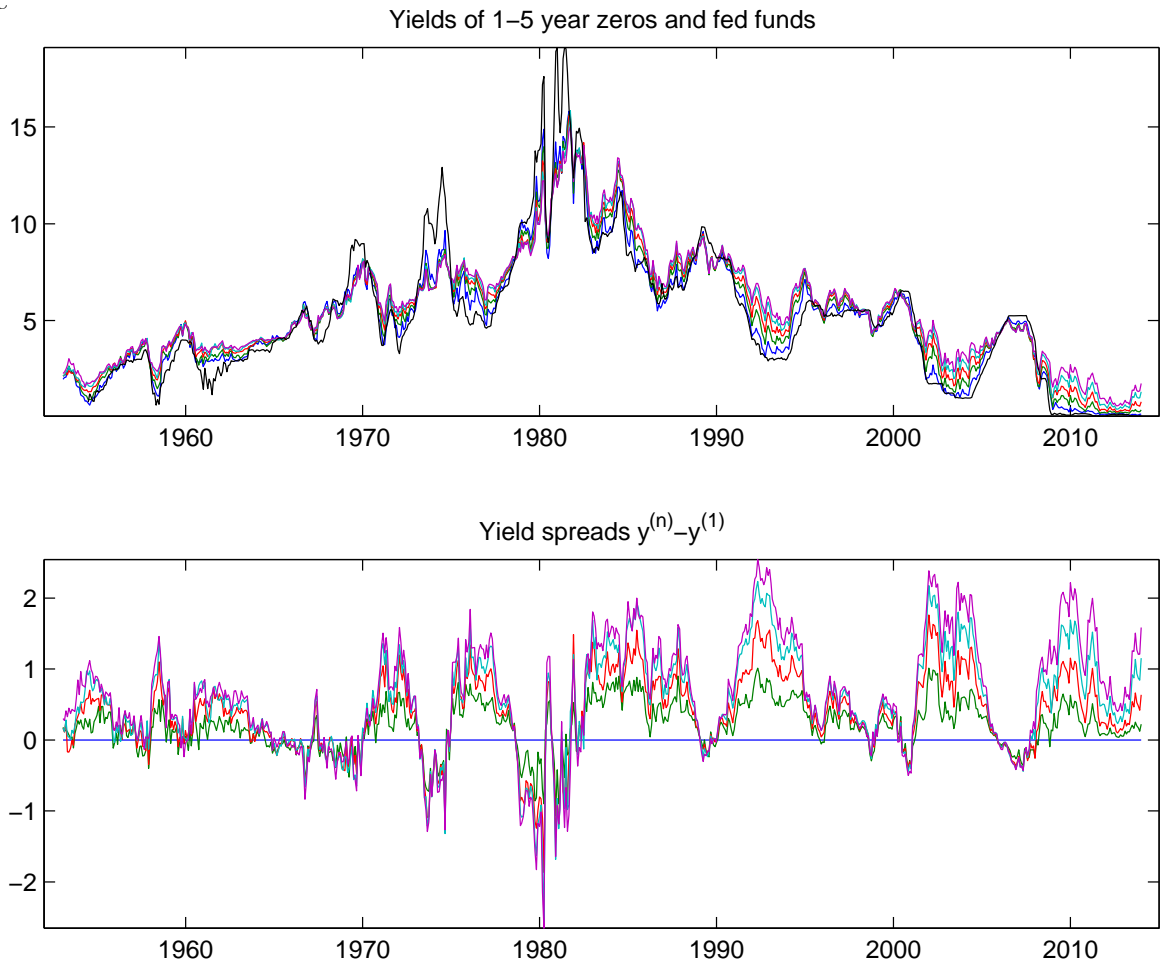
$$\begin{aligned} E_t(R_{t+1}^e) &= cov_t(R_{t+1}^e, m_{t+1}) \\ &= \gamma_t cov_t(R_{t+1}^e, \Delta c_{t+1}) \\ &\approx -\gamma_t n cov_t(y_{t+1}^{(n)}, \Delta c_{t+1}) \end{aligned}$$

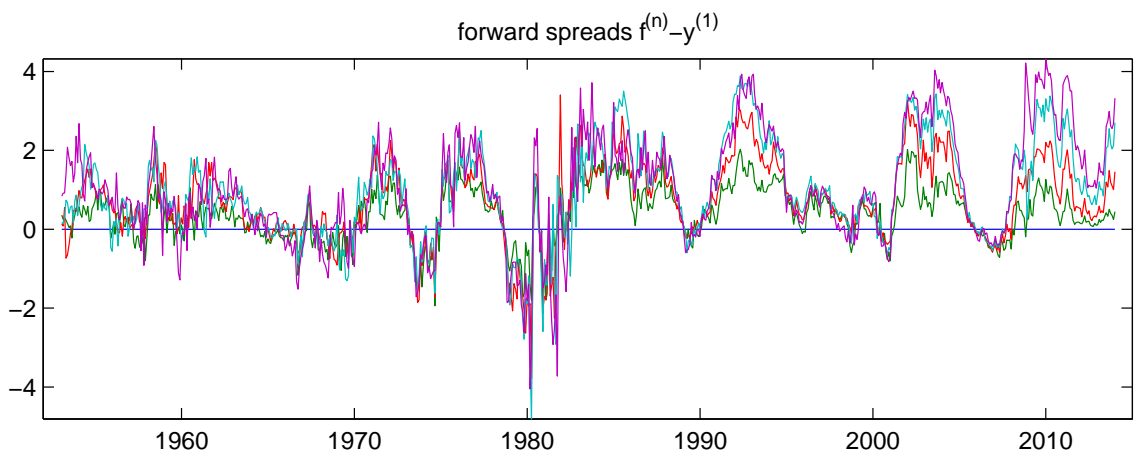
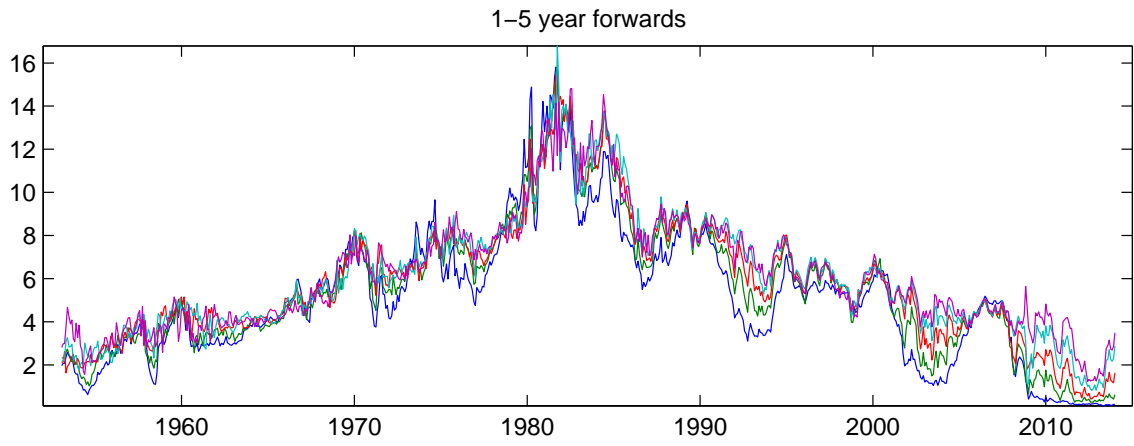
Do interest rate surprises come in good times or bad times? Typically, we think higher rates come with higher consumption growth, so the covariance is positive and the risk premium should be negative!

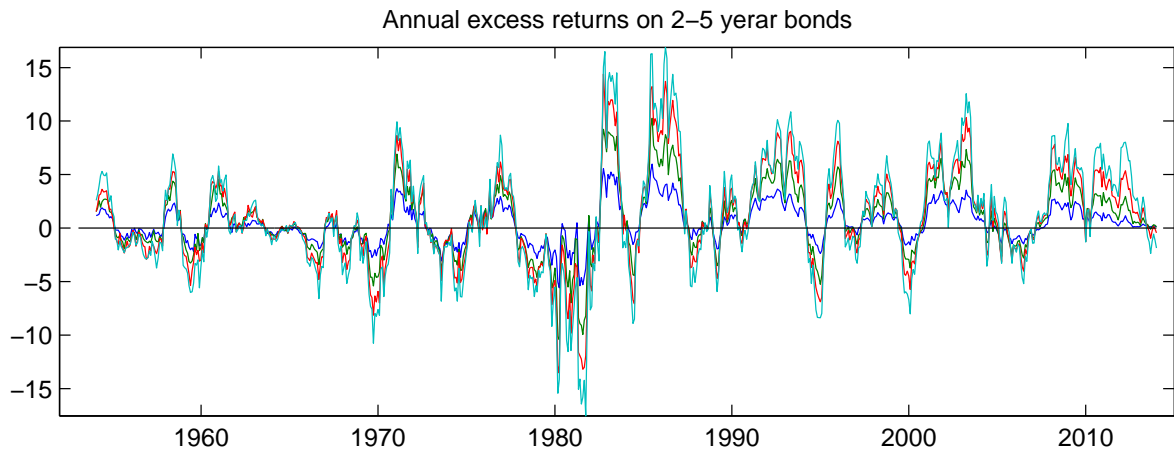
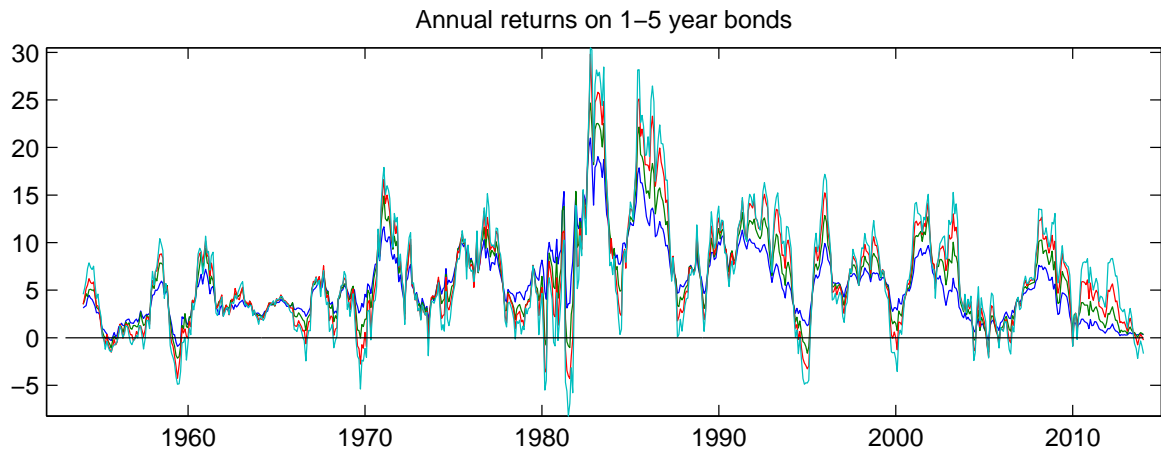
3. US: We'll just look at it empirically.
4. Terminology: The "strict expectations hypothesis" means no risk premium. The "expectations hypothesis" alone means that the risk premium is small and *constant over time*.

23.3 Empirical evaluation of yield curves and risk premia

1. Fac







- (a) EH does not seem awful: When the curve slopes up, yields subsequently rise!
- (b) Risk premium *on average*?

Interest rate data 1964:01-2012:12

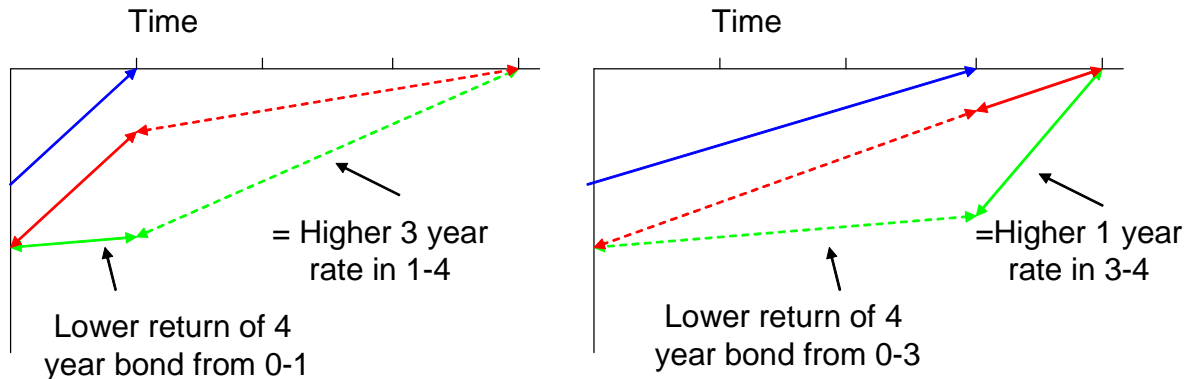
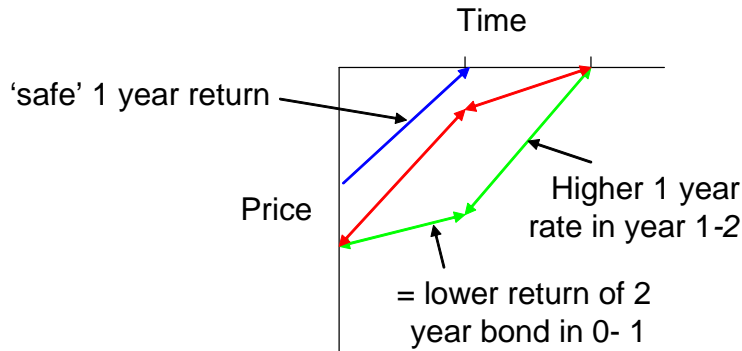
Maturity n	1	2	3	4	5
$E[y^{(n)}]$	5.68	5.90	6.08	6.23	6.34
$E[y^{(n)} - y^{(1)}]$	-	0.21	0.39	0.54	0.65
$E[r^{(n)} - y^{(1)}]$	-	0.50	0.91	1.24	1.36
t statistic	-	(1.94)	(1.91)	(1.86)	(1.66)
$\sigma[r^{(n)} - y^{(1)}]$	-	1.80	3.31	4.60	5.67
“Sharpe”	-	0.28	0.28	0.27	0.24

Slight upward slope. Long bonds seem terrible by one-period measures. (Why hold them? A: you don't care about one period.)

2. Expectations failures–Fama Bliss. (Updated 1964-2012)

n	$rx_{t+1}^{(n)} = a + b(f_t^{(n)} - y_t^{(1)}) + \varepsilon_{t+1}$			$y_{t+n-1}^{(1)} - y_t^{(1)} = a + b(f_t^{(n)} - y_t^{(1)}) + \varepsilon_{t+1}$			
	b	$\sigma(b)$	R^2	a	b	$\sigma(b)$	R^2
2	0.83	0.27	0.11	0.17	0.27	0.01	
3	1.14	0.35	0.13	0.53	0.33	0.05	
4	1.38	0.43	0.15	0.84	0.26	0.14	
5	1.05	0.49	0.07	0.92	0.17	0.17	
	forecasting <i>one</i> year returns on <i>n</i> -year bonds			forecasting <i>one</i> year rates <i>n</i> years from now			

- (a) Two year forward-spot spread translates (0.83) almost 1-1 to expected return. The EH is exactly wrong!
- (b) Two-year forward-spot spread does not forecast one-year rate change at all (0.17).
- (c) The EH does start to work well at 4-5 year horizon. $f^{(5)}$ does forecast that $y^{(1)}$ will be higher in 4 years.
- (d) First row adds to 1. Higher rows do not add up to 1. Why?



Motivated by the left hand graph, break the return from t to $t + n$ into the first year and the rest,

$$0 - p_t^{(n)} = [0 - p_{t+1}^{(n-1)}] + [p_{t+1}^{(n-1)} - p_t^{(n)}]$$

$$f_t^{(n)} - y_t^{(1)} = p_t^{(n-1)} - p_t^{(n)} - y_t^{(1)} = [p_t^{(n-1)} - p_{t+1}^{(n-1)}] + [r_{t+1}^{(n)} - y_t^{(1)}]$$

$$f_t^{(n)} - y_t^{(1)} = (n-1)[y_{t+1}^{(n-1)} - y_t^{(n-1)}] + [r_{t+1}^{(n)} - y_t^{(1)}]$$

$$f_t^{(n)} - y_t^{(1)} = (n-1)[y_{t+1}^{(n-1)} - y_t^{(n-1)}] + rx_{t+1}^{(n)}$$

Motivated by the right hand graph, break the return from t to $t+n$ into the last year and the rest,

$$0 - p_t^{(n)} = [0 - p_{t+n-1}^{(1)}] + [p_{t+n-1}^{(1)} - p_t^{(n)}]$$

$$f_t^{(n)} = p_t^{(n-1)} - p_t^{(n)} = [p_t^{(n-1)} - p_{t+n-1}^{(1)}] + [p_{t+n-1}^{(1)} - p_t^{(n)}]$$

$$f_t^{(n)} - y_t^{(1)} = [y_{t+n-1}^{(1)} - y_t^{(1)}] + [p_t^{(n-1)} + p_{t+n-1}^{(1)} - p_t^{(n)}]$$

$$f_t^{(n)} - y_t^{(1)} = [y_{t+n-1}^{(1)} - y_t^{(1)}] + [p_{t+n-1}^{(1)} - p_t^{(n)} - (n-1)y_t^{(n-1)}]$$

3. Q: Why do we run

$$y_{t+1}^{(1)} - y_t^{(1)} = a + b(f_t^{(2)} - y_t^{(1)}) + \varepsilon_{t+1}?$$

The expectations hypothesis says $f_t^{(2)} = E_t[y_{t+1}^{(1)}]$, so why not run

$$y_{t+1}^{(1)} = a + bf_t^{(2)} + \varepsilon_{t+1}?$$

A1: Temperature analogy

$$T_{t+1} = 0 + 1 \times T_t + \varepsilon_{t+1}$$

$$T_{t+1} = 0 + 1 \times F_t + \varepsilon_{t+1}$$

it's much better to run

$$T_{t+1} - T_t = 0 + 1 \times (F_t - T_t) + \varepsilon_{t+1}$$

A2: “Does a forward *spread* forecast a rate *rise*?” is more powerful than, “does a high forward rate level forecast a high interest rate level?”

4. Interpretation.

- (a) Dividend yield “should” forecast dividend growth and not returns. Dividend yield *does* forecast excess returns and not dividend growth. The two forecasts add up mechanically.
- (b) Yield spread “should” forecast short rates, and not excess returns. Yield spreads *do* forecast excess returns and not yield changes. The two forecasts add up mechanically.
- (c) “Sluggish adjustment.” “Buy Yield”
- (d) The question is really “what sets bond prices and forward rates today?”
- (e) There are big risk premia and they vary over time.

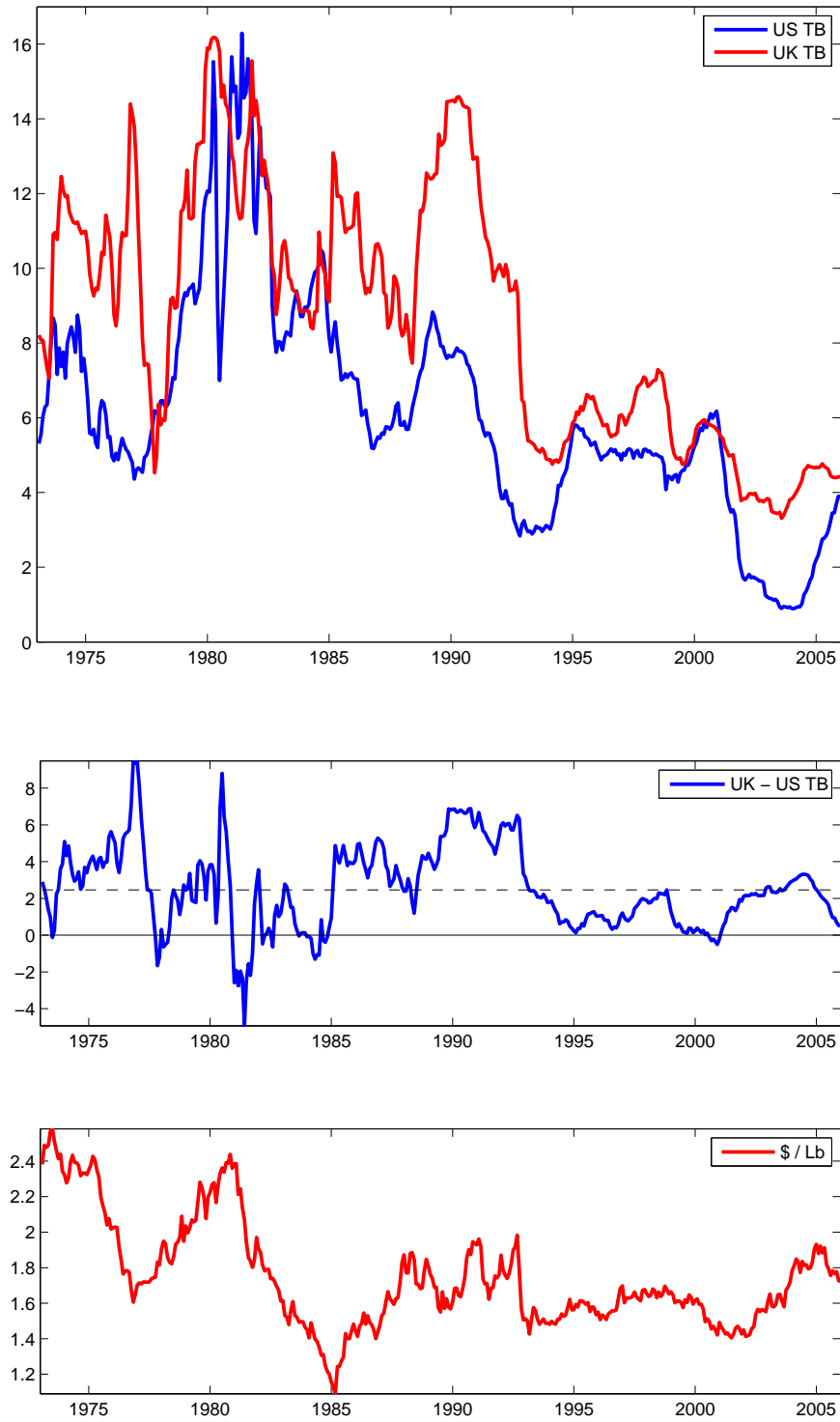
5. Exchange rates: “Carry Trade.” “Buy yield” again.

- (a) From “Asset pricing”

	DM	£	¥	SF
Mean appreciation	-1.8	3.6	-5.0	-3.0
Mean interest differential	-3.9	2.1	-3.7	-5.9
b , 1975-1989	-3.1	-2.0	-2.1	-2.6
R^2	.026	.033	.034	.033
b , 1976-1996	-0.7	-1.8	-2.4	-1.3

$$s_{t+1} - s_t = a + b(f_t - s_t) + \varepsilon_{t+1} = a + b(r_t^f - r_t^d) + \varepsilon_{t+1}$$

6. Picture: as above for yields.



(a) UK > US rates mean stronger pound

- (b) UK > US rates mean pound depreciates
- (c) But not soon enough. Another “sluggish adjustment.”

23.4 Cochrane-Piazzesi update

Bottom line

1. Forecast 1 year horizon returns on treasury bonds, using *all* forwards

- (a) FB:

$$rx_{t+1}^{(n)} = a_n + b_n(f_t^{(n)} - y_t^{(1)}) + \varepsilon_{t+1}^{(n)}$$

- (b) CP:

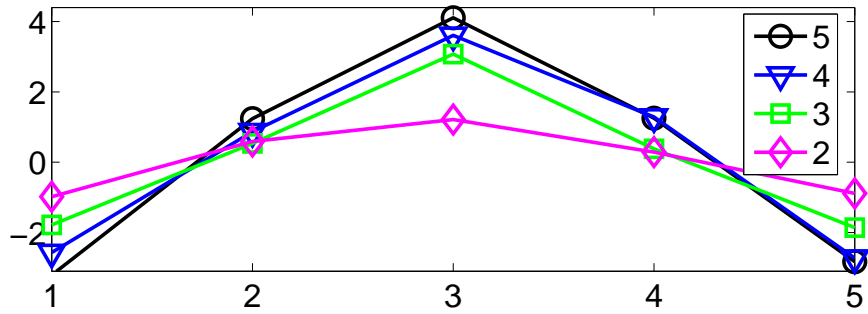
$$\begin{aligned} rx_{t+1}^{(n)} &= a_n + \beta_{n,1}y_t^{(1)} + \beta_{n,2}f_t^{(2)} + \beta_{n,3}f_t^{(3)} + \beta_{n,4}f_t^{(4)} + \beta_{n,5}f_t^{(5)} + \varepsilon_{t+1}^{(n)} \\ &= a_n + \beta'_n f_t + \varepsilon_{t+1}^{(n)} \end{aligned}$$

2. Results:

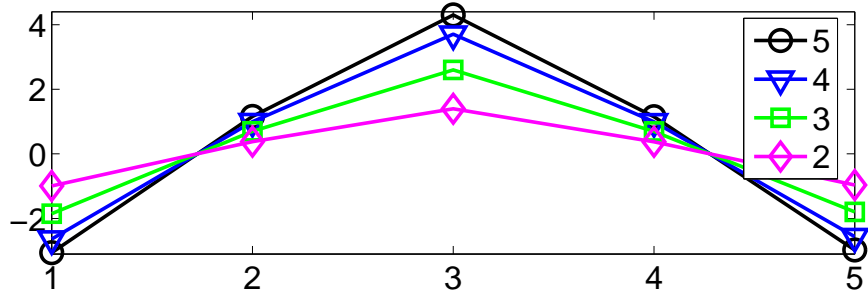
- (a) R^2 rises up to 44%, up from Fama-Bliss 15%
- (b) The pattern of the $\beta_{n,1}.. \beta_{n,5}$ are the *same* when forecasting each maturity n
- (c) A *single* “factor” $\gamma' f_t$ forecasts bond returns rx_{t+} of *all* maturities. High expected returns in “bad times.”
- (d) The factor tells you when to bail out – when rates will rise in an upward-slope environment

Basic regression

Unrestricted



Restricted



$$rx_{t+1}^{(n)} = a_n + \beta_{n,1}y_t^{(1)} + \beta_{n,2}f_t^{(2)} + \dots + \beta_{n,5}f_t^{(5)} + \varepsilon_{t+1}^{(n)}$$

A single factor for expected bond returns

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

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

$$\overline{rx}_{t+1} = \frac{1}{4} \sum_{n=2}^5 rx_{t+1}^{(n)} = \gamma' f_t + \varepsilon_{t+1}$$

Table 1 Estimates of the single-factor model

A. Estimates of the return-forecasting factor, $\overline{rx}_{t+1} = \gamma' f_t + \bar{\varepsilon}_{t+1}$

	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

n	Restricted		Unrestricted
	b_n	R^2	R^2
2	0.47	0.31	0.32
3	0.87	0.34	0.34
4	1.24	0.37	0.37
5	1.43	0.34	0.35

Stock Return Forecasts

Table 3. Forecasts of excess stock returns (VWNYSE)

$$\overline{r}x_{t+1} = a + bx_t + \varepsilon_{t+1}$$

	$\gamma^\top f$	(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

- The coefficient is about right.
- 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.

What is and is not important about CP

- What is important, and very robust: *A single factor model for expected returns*. What is not important: tent vs slope, or the ability of additional variables to forecast returns.
- CP suggests more: A common factor across stocks, bonds, fx?
- I'll come back to these statements after talking about "factor models" in some detail.

23.5 FX update

1. Lustig, Roussanov, Verdelhan. *Let's form portfolios like Fama and French!*

(a) Table 1: Form 6 portfolios based on f-s. Mean returns etc. by portfolio

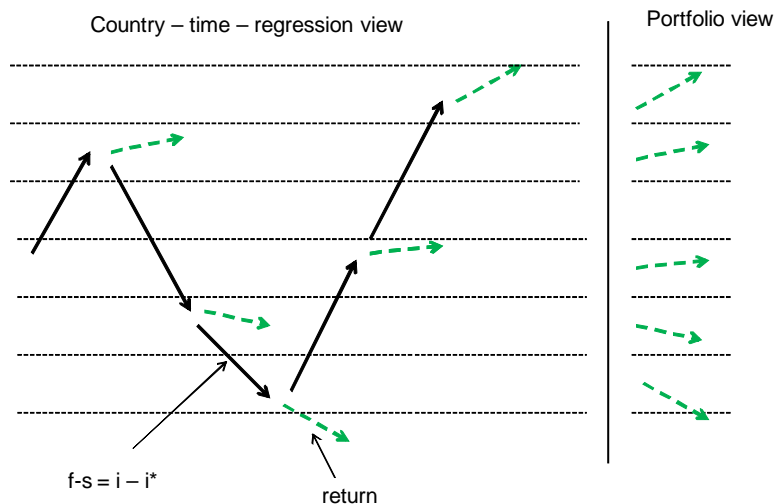


Table 1
Currency portfolios—U.S. investor

Portfolio	1	2	3	4	5	6
Panel I: All Countries						
	Spot change: Δs^j					
Mean	-0.64	-0.92	-0.95	-2.57	-0.60	2.82
Std	8.15	7.37	7.63	7.50	8.49	9.72
	Forward Discount: $f^j - s^j$					
Mean	-2.97	-1.23	-0.09	1.00	2.67	9.01
Std	0.54	0.48	0.47	0.52	0.64	1.89
	Excess Return: rx^j (without b-a)					
Mean	-2.33	-0.31	0.86	3.57	3.27	6.20
Std	8.23	7.44	7.66	7.59	8.56	9.73
SR	-0.28	-0.04	0.11	0.47	0.38	0.64
	Net Excess Return: rx_{net}^j (with b-a)					
Mean	-1.17	-1.27	-0.39	2.26	1.74	3.38
Std	8.24	7.44	7.63	7.55	8.58	9.72
SR	-0.14	-0.17	-0.05	0.30	0.20	0.35

(b) Table 3: how you form “factors” like rrmf and hml

Table 3
Principal components

Panel I: All Countries						
<i>Portfolio</i>	1	2	3	4	5	6
1	0.42	0.43	0.18	-0.15	0.74	0.20
2	0.38	0.24	0.15	-0.27	-0.61	0.58
3	0.38	0.29	0.42	0.12	-0.28	-0.71
4	0.38	0.04	-0.35	0.83	-0.03	0.18
5	0.43	-0.08	-0.72	-0.44	-0.03	-0.30
6	0.45	-0.81	0.35	-0.03	0.11	0.06
% Var.	71.95	11.82	5.55	4.00	3.51	3.16

Panel II: Developed Countries					
<i>Portfolio</i>	1	2	3	4	5
1	0.44	0.66	-0.54	-0.25	0.12
2	0.45	0.25	0.75	0.01	0.41
3	0.46	0.02	0.19	0.04	-0.86
4	0.44	-0.27	-0.29	0.78	0.20
5	0.45	-0.66	-0.14	-0.57	0.17
% Var.	78.23	10.11	4.97	3.49	3.20

This table reports the principal component coefficients of the currency portfolios presented in Table 1. In each panel, the last row reports (in %) the share of the total variance explained by each common factor. Data are monthly, from Barclays and Reuters (Datastream). The sample period is 11/1983–12/2009.

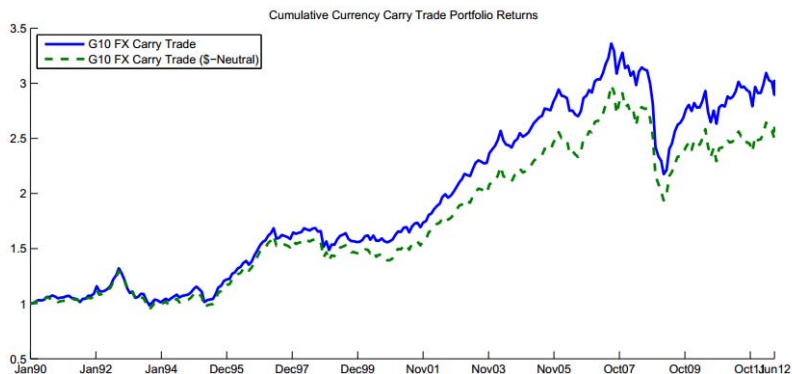
(c) Table 4: Like FF Table 1, high average returns have high betas on the hml factor.

Table 4
Continued

Portfolio	All Countries			
	α_0^j	$\beta_{HML_{EX}}^j$	β_{RX}^j	R^2
1	-0.10 [0.50]	-0.39 [0.02]	1.05 [0.03]	91.64
2	-1.55 [0.73]	-0.11 [0.03]	0.94 [0.04]	77.74
3	-0.54 [0.74]	-0.14 [0.03]	0.96 [0.04]	76.72
4	1.51 [0.77]	-0.01 [0.03]	0.95 [0.05]	75.36
5	0.78 [0.82]	0.04 [0.03]	1.06 [0.05]	76.41
6	-0.10 [0.50]	0.61 [0.02]	1.05 [0.03]	93.84
...				

2. The “Peso problem” and fat tails. Jurek Figure 1

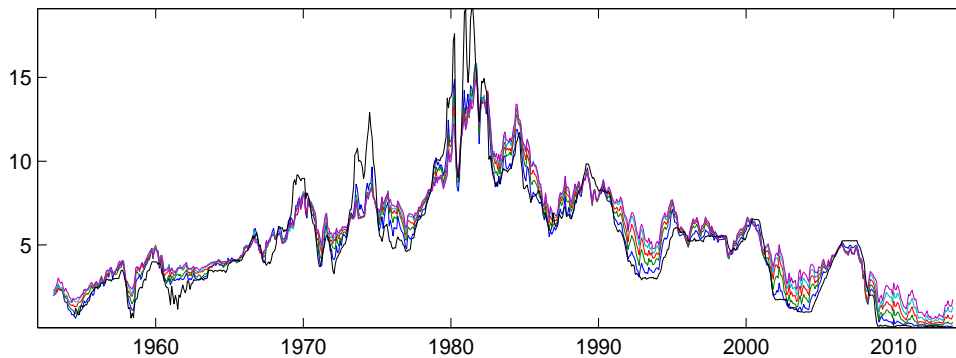
Figure 1. Total Returns to G10 Currency Carry Trades. The top panel of this figure illustrates the total return indices for portfolios of G10 currency carry trades over the period from January 1990 to June 2012. The portfolio composition is determined by sorting currencies on the basis of their prevailing 1-month LIBOR rate, and going long (short) currencies with the highest (lowest) interest rates. Portfolios are rebalanced monthly, and allocations to individual currencies are spread-weighted. The graphs plots the returns to dollar-neutral and non-dollar-neutral portfolios. The bottom panel plots the time series of instantaneous volatility for the spread-weighted, non-dollar neutral currency carry trade portfolio. The plot reports values based on an EGARCH(1, 1) model estimated in-sample using daily, historical realized returns (1990:1-2012:6), and a measure of portfolio volatility based on FX option-implied volatilities (1999:1-2012:6). The *FX Option-implied* volatility measure is obtained by reconstructing the portfolio variance-covariance matrix using the full cross-section of 45 G10 cross-rate variance swap quotes.



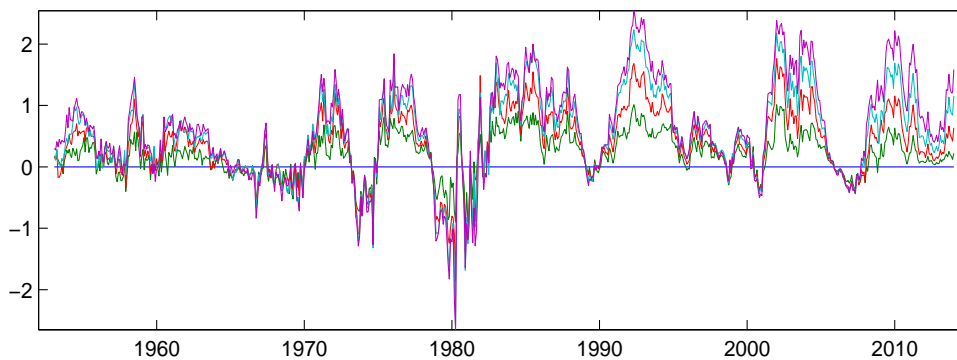
24 Term Structure II: Interest Rates Factor Models Overheads

24.1 Motivation and idea

Yields of 1–5 year zeros and fed funds



Yield spreads $y^{(n)} - y^{(1)}$



- “Level, slope and curvature movements seem to capture most variation in yields”
- A model? How about

$$y_t^{(n)} = a_n + b_n \text{level}_t + c_n \text{slope}_t + (\text{other stuff})$$

$$\begin{bmatrix} y_t^{(1)} \\ y_t^{(2)} \\ y_t^{(3)} \\ y_t^{(4)} \\ y_t^{(5)} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \text{level}_t + \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{bmatrix} \text{slope}_t + (\text{other stuff})$$

- A lot like the FF3F model.

$$R_{t+1}^{ei} = \alpha_i + b_i r_{mr} f_{t+1} + h_i h_{ml} m_{t+1} + s_i s_{mb} b_{t+1} + \varepsilon_{t+1}^i$$

As there, we’re describing *variance*, not yet *expected returns*. R^2 matters, intercepts don’t.

- A simple approach to a factor model

1. Like Fama and French, let’s guess some good factors:

$$\text{level}_t \equiv \frac{1}{5} [y_t^{(1)} + y_t^{(2)} + y_t^{(3)} + y_t^{(4)} + y_t^{(5)}]$$

$$\text{slope}_t = \frac{1}{2} [y_t^{(4)} + y_t^{(5)}] - \frac{1}{2} [y_t^{(1)} + y_t^{(2)}].$$

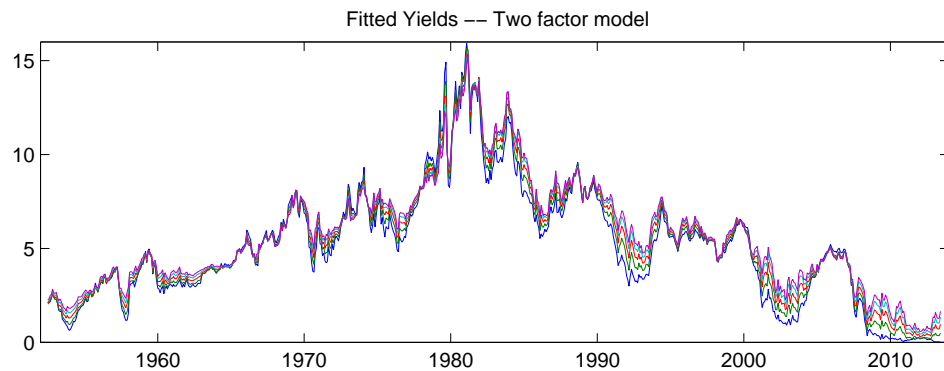
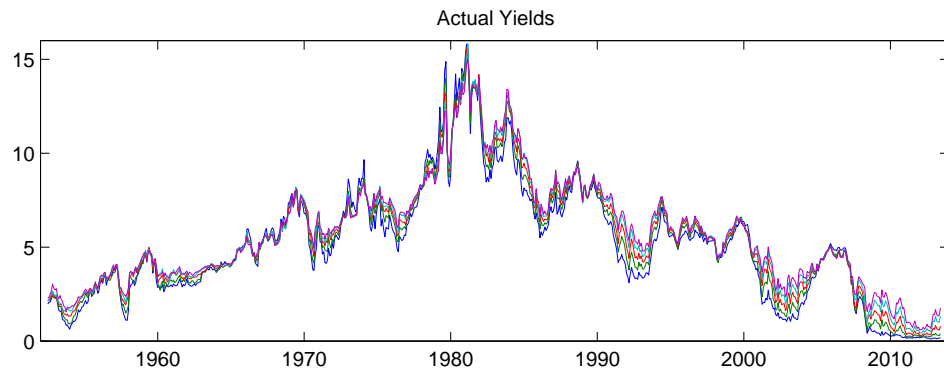
2. Run regressions to find the “betas.”

$$y_t^{(n)} = a_n + b_n \times \text{level}_t + s_n \times \text{slope}_t + \varepsilon_t^{(n)}$$

	a_n	b_n	s_n	R^2
$y^{(1)}$	0.09	0.9998	-0.77	0.97
$y^{(2)}$	-0.08	1.0003	-0.28	0.97
$y^{(3)}$	-0.04	0.9999	0.10	0.97
$y^{(4)}$	-0.02	1.0008	0.38	0.98
$y^{(5)}$	0.04	0.9993	0.56	0.98

3. R^2 is very high! The two-factor model is a very good approximation

$$\text{fit } y_t^{(n)} = b_n \times \text{level}_t + s_n \times \text{slope}_t$$



• *Two steps (which is in everything we do)*

1. *Weights.* How do you add up yields to form the factors. Form the factors (level, slope) from yields (returns, etc) $\text{level}_t = \dots$, $\text{slope}_t = \dots$
2. *Loadings or betas,* how do yields change when one of the factors changes $y_t^{(n)} = \dots \text{level}_t + \dots \text{slope}_t \dots$

24.2 Factor analysis

- A way to automate step 1 and 2 that maximizes the R^2 of the factor model.

- Procedure

1. Form the covariance matrix of yields. $\Sigma = \text{cov}(100 \cdot \text{yields})$; Using the notation

$$y_t = \begin{bmatrix} y_t^{(1)} \\ y_t^{(2)} \\ y_t^{(3)} \\ \vdots \end{bmatrix}$$

$$\Sigma = \text{cov}(y_t, y_t')$$

2. Take its eigenvalue decomposition. $[Q, \Lambda] = \text{eig}(\Sigma)$.

$$\Sigma = Q\Lambda Q'$$

Λ is diagonal,

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$

3. $Q'Q = QQ' = I$.

$$Q = \begin{bmatrix} | & | & | \\ q_1 & q_2 & q_3 \\ | & | & | \end{bmatrix}$$

$$\begin{bmatrix} - & q'_1 & - \\ - & q'_2 & - \\ - & q'_3 & - \end{bmatrix} \begin{bmatrix} | & | & | \\ q_1 & q_2 & q_3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

$$\Sigma = Q\Lambda Q'$$

$$\begin{bmatrix} | & | & | \\ q_1 & q_2 & q_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \begin{bmatrix} - & q'_1 & - \\ - & q'_2 & - \\ - & q'_3 & - \end{bmatrix}$$

4. The columns of Q tell you how to construct the factors x_t from the yields y_t . You can form the factors by

$$x_t = Q'y_t.$$

This means

$$\begin{bmatrix} x_t^{(1)} \\ x_t^{(2)} \\ x_t^{(3)} \end{bmatrix} = \begin{bmatrix} - & q'_1 & - \\ - & q'_2 & - \\ - & q'_3 & - \end{bmatrix} \begin{bmatrix} y_t^{(1)} \\ y_t^{(2)} \\ y_t^{(3)} \end{bmatrix}$$

for instance,

$$x_t^{(1)} = \begin{bmatrix} - & q'_1 & - \end{bmatrix} \begin{bmatrix} y_t^{(1)} \\ y_t^{(2)} \\ y_t^{(3)} \end{bmatrix}.$$

For example: if

$$q'_1 = \begin{bmatrix} 1/3 & 1/3 & 1/3 \end{bmatrix}$$

then this operation constructs the “level” factor.

$$x_t^{(1)} = \begin{bmatrix} 1/3 & 1/3 & 1/3 \end{bmatrix} \begin{bmatrix} y_t^{(1)} \\ y_t^{(2)} \\ y_t^{(3)} \end{bmatrix} = \frac{1}{3} (y_t^{(1)} + y_t^{(2)} + y_t^{(3)}).$$

5. The columns of Q are also the **loadings**, they tell you how much each y moves when a factor x moves.

$$y_t = Qx_t.$$

This means

$$\begin{bmatrix} y_t^{(1)} \\ y_t^{(2)} \\ y_t^{(3)} \end{bmatrix} = \begin{bmatrix} | & | & | \\ q_1 & q_2 & q_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} x_t^{(1)} \\ x_t^{(2)} \\ x_t^{(3)} \end{bmatrix} = \begin{bmatrix} | \\ q_1 \\ | \end{bmatrix} x_t^{(1)} + \begin{bmatrix} | \\ q_2 \\ | \end{bmatrix} x_t^{(2)} + \begin{bmatrix} | \\ q_3 \\ | \end{bmatrix} x_t^{(3)}$$

or

$$y_t^{(n)} = q_1^{(n)} x_t^{(1)} + q_2^{(n)} x_t^{(2)} + q_3^{(n)} x_t^{(3)}$$

6. The factors are uncorrelated with each other, and the Λ diagonals tell you the variance of each factor

$$\text{cov}(x_t, x_t') = \Lambda$$

this means

$$\text{cov} \left(\begin{bmatrix} x_t^{(1)} \\ x_t^{(2)} \\ x_t^{(3)} \end{bmatrix} \begin{bmatrix} x_t^{(1)} & x_t^{(2)} & x_t^{(3)} \end{bmatrix} \right) = \begin{bmatrix} \sigma^2(x^{(1)}) & \text{cov}(x^{(1)}, x^{(2)}) & \text{cov}(x^{(1)}, x^{(3)}) \\ \cdot & \sigma^2(x^{(2)}) & \text{cov}(x^{(2)}, x^{(3)}) \\ \cdot & \cdot & \sigma^2(x^{(3)}) \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$

7. $y = Qx$ is a factor model—the columns of Q are the “loadings” or the “betas”; the b_n and c_n of the first equation.
8. To form a factor model, *leave out the factors with really small eigenvalues*, either ignoring them or treating them as small errors For example,

$$\begin{aligned} y_t^{(n)} &= q_1^{(n)} x_t^{(1)} + q_2^{(n)} x_t^{(2)} + \varepsilon_t \\ y_t^{(n)} &\doteq q_1^{(n)} x_t^{(1)} + q_2^{(n)} x_t^{(2)} \end{aligned}$$

Since the factors are uncorrelated with each other, ε_t is uncorrelated with $x_t^{(1)}$ and $x_t^{(2)}$ so q are still regression coefficients.

24.2.1 A real example: yields

- A real example. Yields! This also shows you that all this talk is only about 2-3 lines of matlab

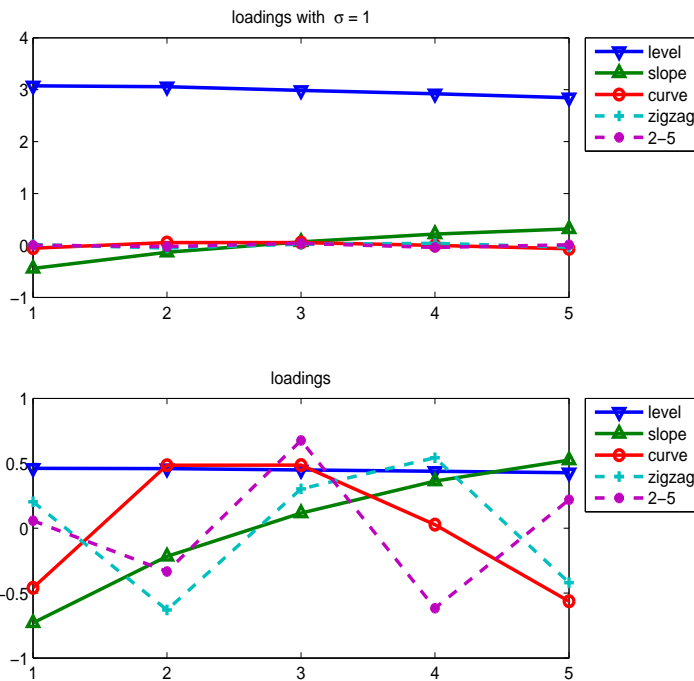
```

Sigma = cov(100*yields);
[Q,L] = eig(Sigma);
loads = Q*L^0.5;
disp('standard deviation of factors');
disp(diag(L)'.^0.5);
(my names)2-5          zigzag      curve      slope      level
      0.06          0.07          0.10          0.58          5.80

disp(Q)
      0.06          0.15          -0.47          -0.74          0.46
     -0.35          -0.55          0.56          -0.21          0.46
      0.70          0.32          0.44          0.12          0.45
     -0.59          0.57          -0.03          0.36          0.44
      0.19          -0.49          -0.52          0.51          0.43

plot(loads)
plot(Q)

```



1. Interpretation: *If factor x moves, how much do yields of maturity 1,2,3,4,5 move?*
2. Interpretation: *What combination of $y^{(1)} \dots y^{(n)}$ produces each factor? Sensibly, "level" is an*

average of all yields, slope is long - short, etc.

$$\text{Interpretation 1: } \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ y^{(3)} \end{bmatrix} = \begin{bmatrix} | & | & | \\ q_1 & q_2 & q_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} x^{(1)} \\ x^{(2)} \\ x^{(3)} \end{bmatrix}$$

$$\text{Interpretation 2: } \begin{bmatrix} x^{(1)} \\ x^{(2)} \\ x^{(3)} \end{bmatrix} = \begin{bmatrix} - & q_1 & - \\ - & q_2 & - \\ - & q_3 & - \end{bmatrix} \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ y^{(3)} \end{bmatrix}$$

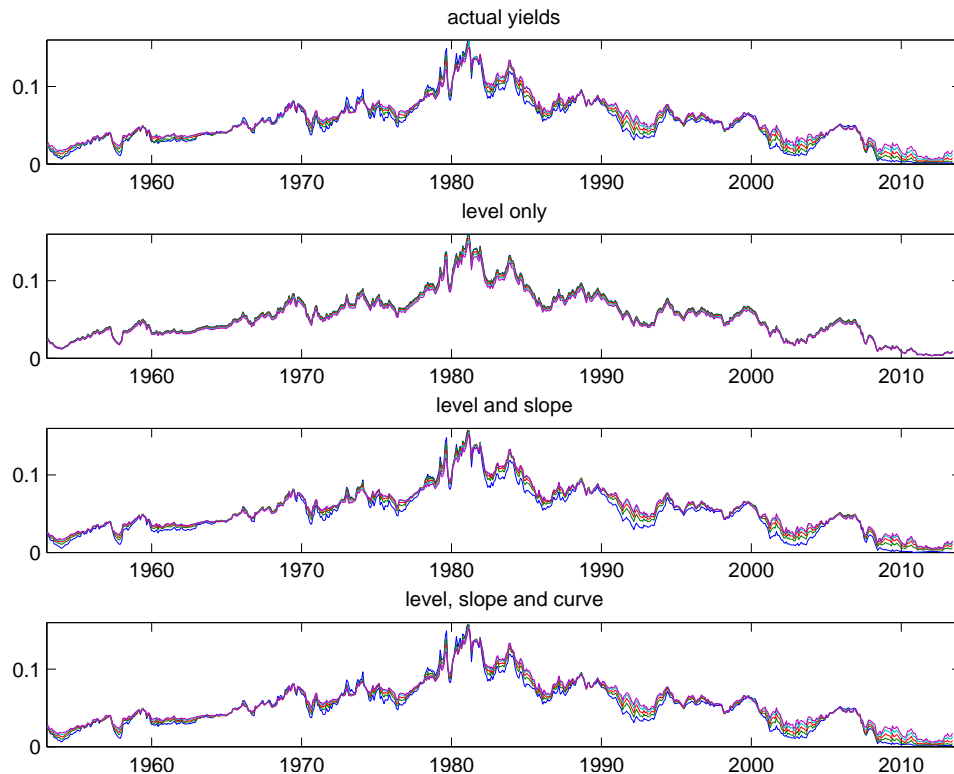
The names “level” “slope” and “curvature” come entirely as verbal descriptions of the shapes of these graphs.

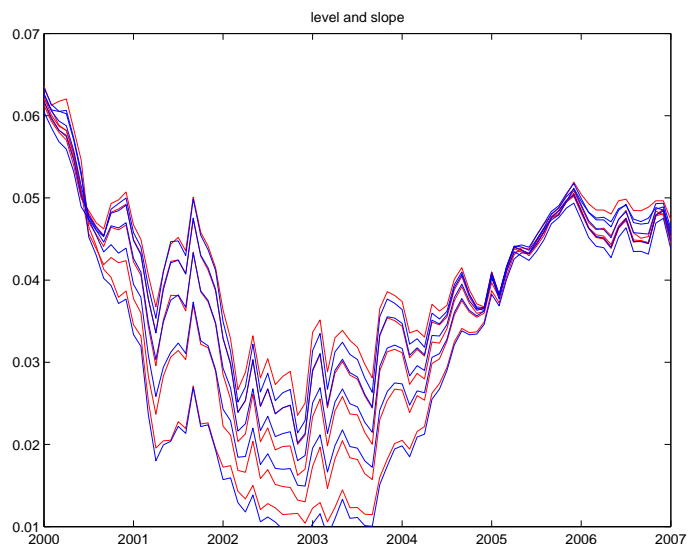
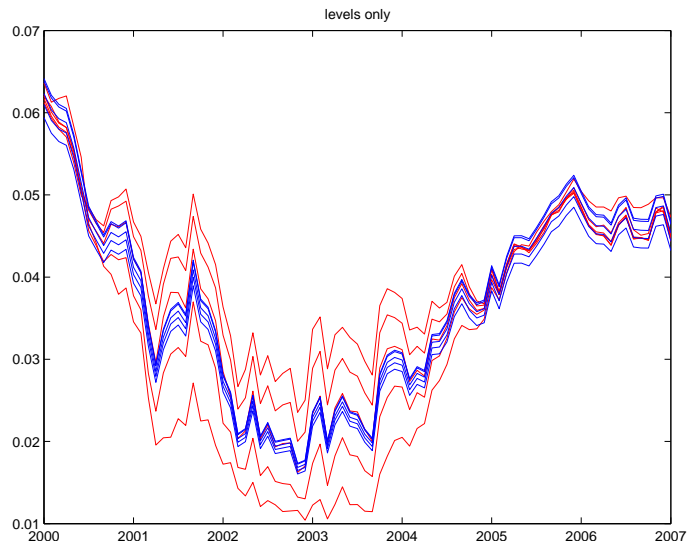
- I also plot “loads” $Q\Lambda^{\frac{1}{2}}$ where I scale each column of q by the standard deviation of the corresponding factor. This lets you see the effect of a typical 1σ move of that factor, and shows visually how the factor with more variance totally dominates.

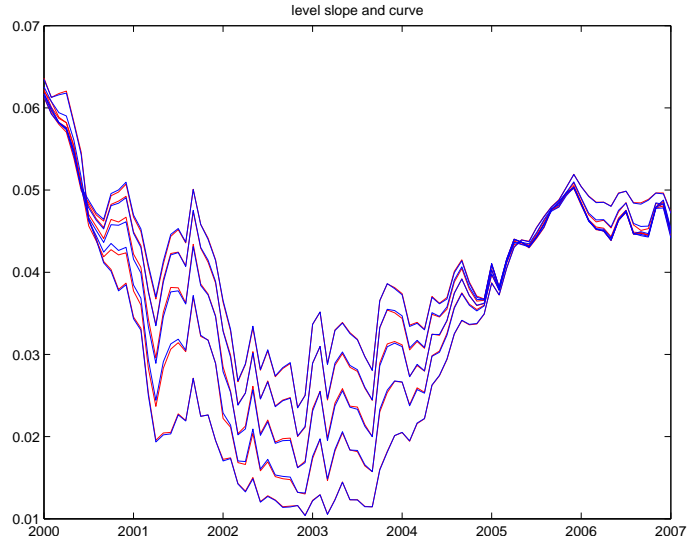
24.2.2 Dropping factors

- Dropping some factors.* So, what happens if we only keep the factors with big variance?

$$\begin{bmatrix} y_t^{(1)} \\ y_t^{(2)} \\ y_t^{(3)} \\ y_t^{(4)} \\ y_t^{(5)} \end{bmatrix} \approx q_1 \times \text{level}_t + q_2 \times \text{slope}_t + q_3 \times \text{curve}_t$$







- Uncorrelated factors:

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

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

- You can just as easily do the same thing for stocks and for returns.

$$\begin{aligned} \Sigma &= \text{cov}(R, R') \\ \Sigma &= Q\Lambda Q' \\ x &= Q'R \end{aligned}$$

now we have the factors

$$R_t = Qx_t \approx Q(:, 1 : 3)x_t + \text{small error}$$

This is the heart of Barra, commercial risk management models.

24.3 Simplest term structure model – expectations hypothesis

- “Derive” a term structure model. Later, form “arbitrage free” term structure models that price derivatives.
- Assume

$$(y_{t+1}^{(1)} - \delta) = \rho(y_t^{(1)} - \delta) + \varepsilon_{t+1}.$$

- Find other rates from EH:

$$\begin{aligned} f_t^{(2)} &= E_t(y_{t+1}^{(1)}) = \delta + \rho(y_t^{(1)} - \delta) \\ f_t^{(3)} &= E_t(y_{t+2}^{(1)}) = \delta + \rho^2(y_t^{(1)} - \delta) \\ f_t^{(N)} &= E_t(y_{t+N-1}^{(1)}) = \delta + \rho^{N-1}(y_t^{(1)} - \delta) \end{aligned}$$

- Then prices, yields are easy

$$y_t^{(2)} = \frac{1}{2} \left(y_t^{(1)} + f_t^{(2)} \right) = \delta + \frac{1}{2} (1 + \rho) \left(y_t^{(1)} - \delta \right)$$

$$y_t^{(n)} = \frac{1}{n} \left(y_t^{(1)} + f_t^{(2)} + \dots + f_t^{(n)} \right) = \delta + \frac{1}{n} \left(1 + \rho + \rho^2 + \dots + \rho^{n-1} \right) \left(y_t^{(1)} - \delta \right)$$

- Look what we have

1. An exact (no error) one-factor model
2. Lots of different shapes – some slope and curvature

- More complex shapes, more factors? Move past the simple AR(1)! Example

$$z_t = z_{t-1} + \varepsilon_t$$

$$x_t = \phi x_{t-1} + \delta_t$$

$$y_t^{(1)} = z_t + x_t$$

$$f_t^{(2)} = E_t y_{t+1}^{(1)} = z_t + \phi x_t$$

$$f_t^{(n)} = E_t y_{t+n-1}^{(1)} = z_t + \phi^{n-1} x_t$$

z_t is a “level” factor and x_t is a “slope” factor.

$$f_t^{(n)} - y_t^{(1)} = \left(\phi^{n-1} - 1 \right) x_t$$

- Real “arbitrage free” term structure models work the same way. In place of the expectations hypothesis, however, those models start with $P_t^{(n)} = E_t(m_{t+1}m_{t+2}\dots m_{t+n} \times 1)$. Example: “Single factor Vasicek”

$$\left(y_{t+1}^{(1)} - \delta \right) = \rho \left(y_t^{(1)} - \delta \right) + \varepsilon_{t+1}.$$

$$\left(f_t^{(2)} - \delta \right) = \rho \left(y_t^{(1)} - \delta \right) - \left[\frac{1}{2} + \lambda \right] \sigma_\varepsilon^2$$

$$\left(f_t^{(3)} - \delta \right) = \rho^2 \left(y_t^{(1)} - \delta \right) - \left[\frac{1}{2} (1 + \rho)^2 + \lambda(1 + \rho) \right] \sigma_\varepsilon^2$$

- Use? Since we price by arbitrage, and all prices are functions of the single factor, we can price bond options “by arbitrage” just as Black-Scholes price equity options.

24.4 Liquidity, on-off run spreads

Crazy treasury yield curve in Dec 2008 just like everything else (Courtesy Aaron Pancost)

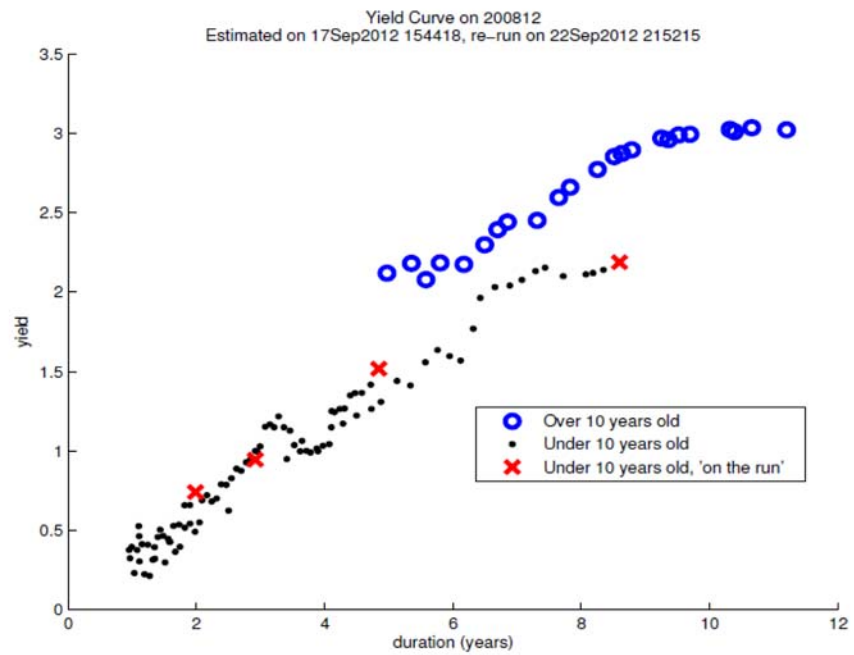
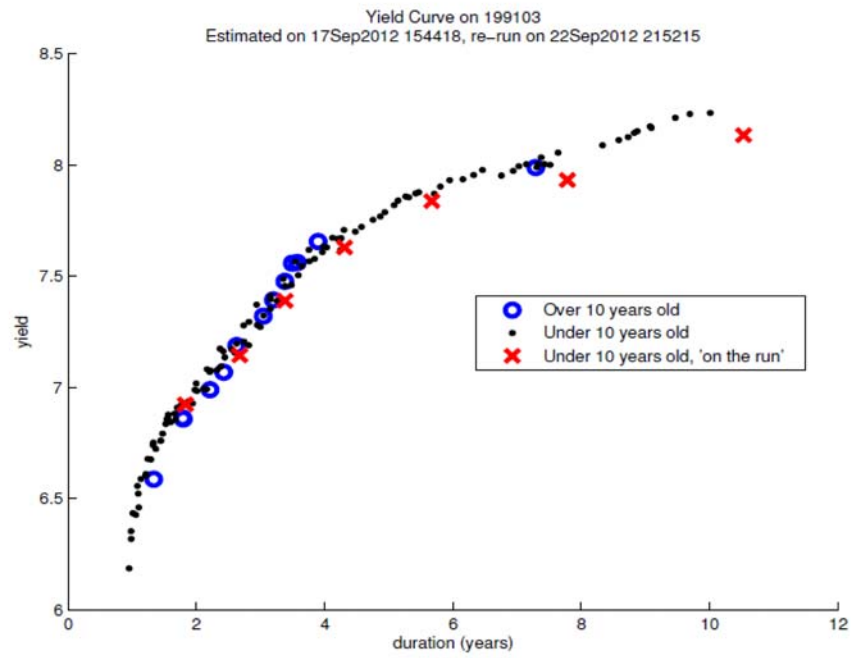


Figure 6: Yield Curves in 1991 and 2008