Bonds and FX

Fama-Bliss, Asset Pricing

1. Preview: The basic picture is still "1/p forecasts excess returns" (If you're not familiar with yields etc., read discrete time bonds chapter)

2. Expectations hypothesis:

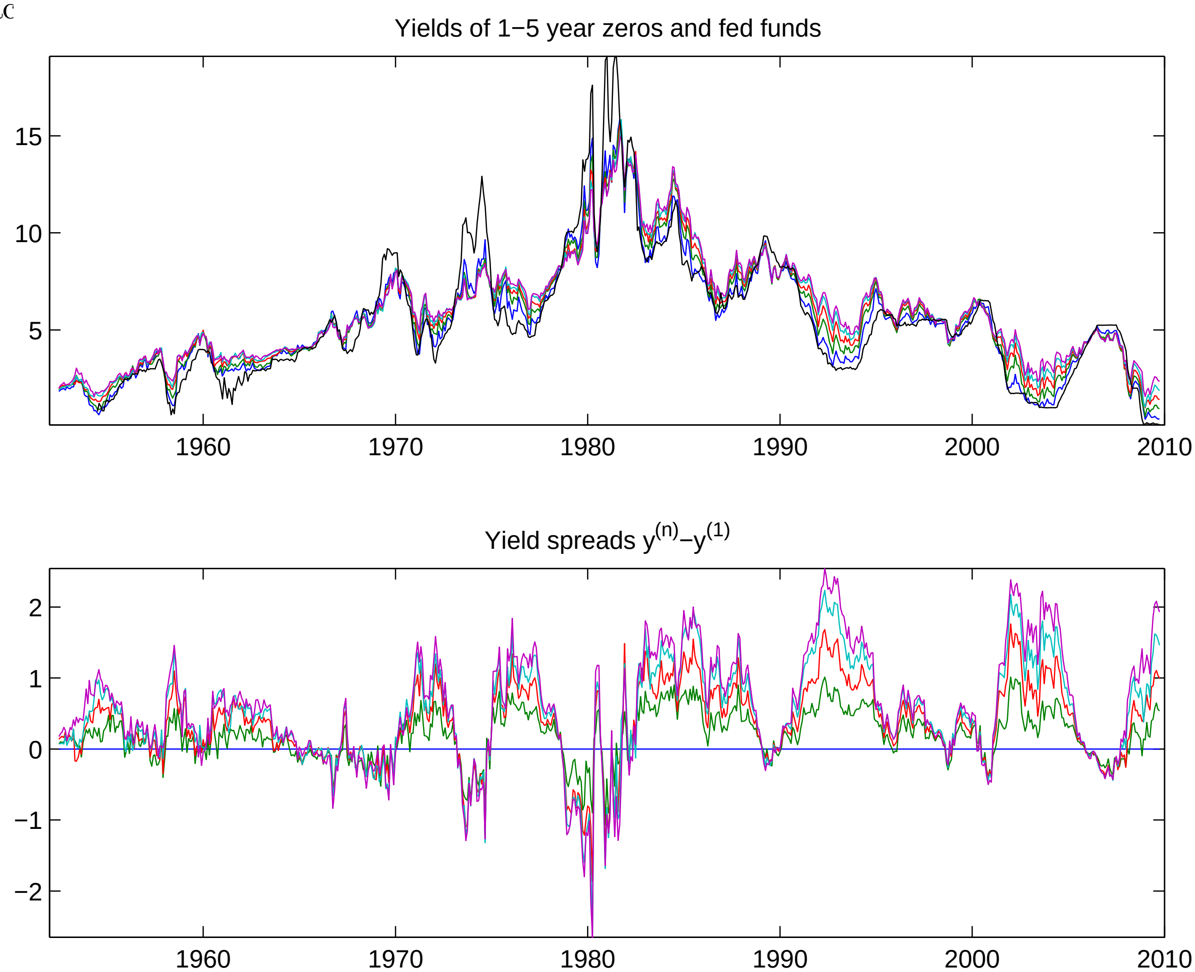
1) $y_t^{(N)} = E_t \sum_{j=0}^{N-1} y_{t+j}^{(1)}$ (+rp, constant over time)

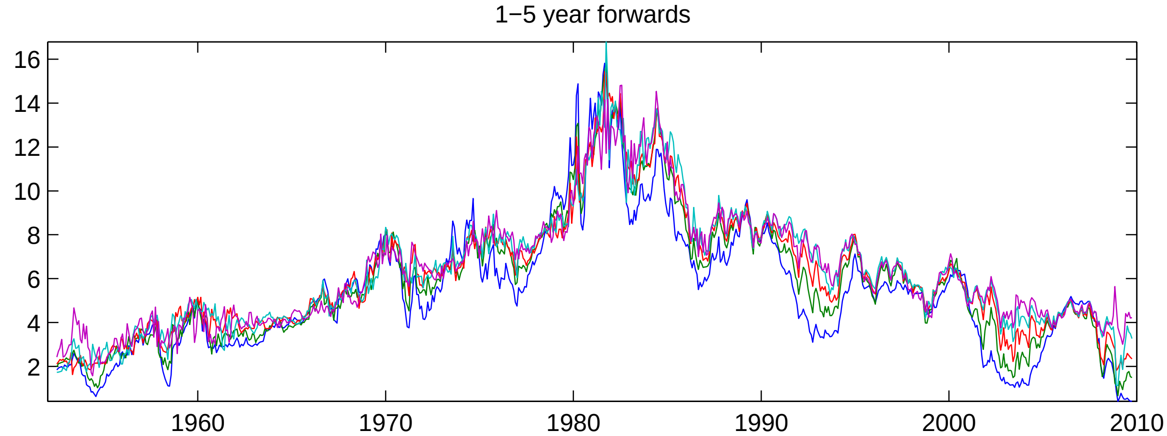
2) $f_t^{N-1 \to N} = E_t(y_{t+N-1}^{(1)})$ (+rp, constant over time)

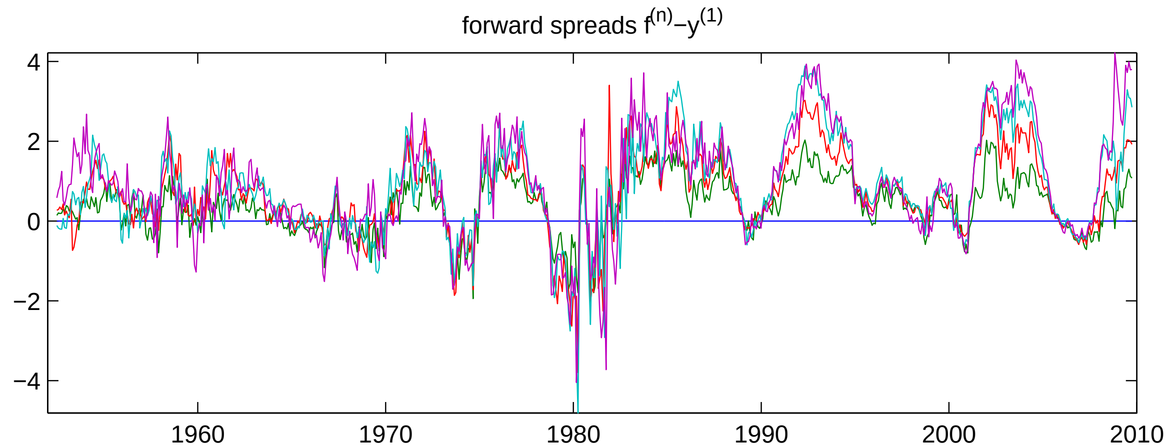
3) $y_t^{(1)} = E_t r_{t+1}^{N \to N-1}$.(+rp, constant over time)

You can see a rough truth in the data. But you can also see "sluggish adjustment" which is the core of the predictability result.

1. Fac







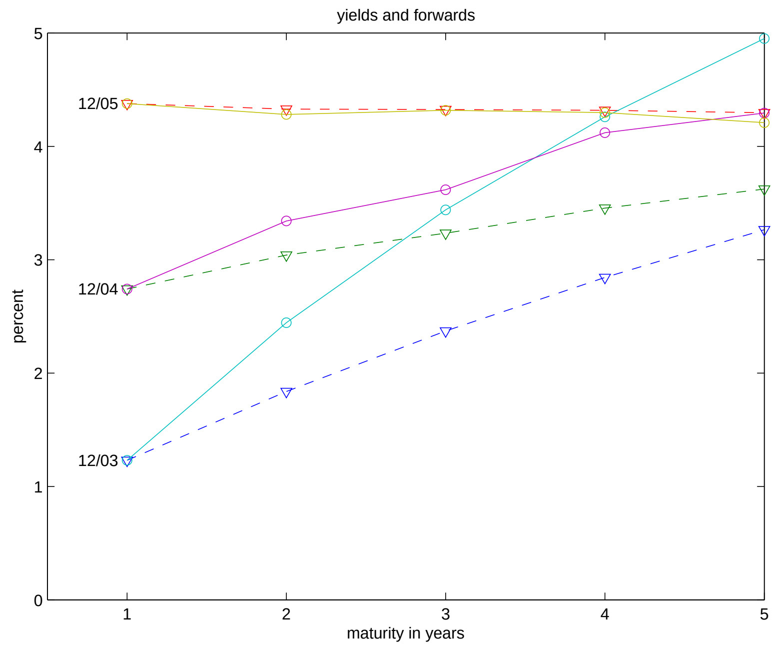
(a) The dominant movement is "level shift" – all yields up and down together.

(b) The yield curve does change shape. Sometimes rising, sometimes falling. This is the "slope" factor. There is a business cycle pattern: inverted at peaks, rising at troughs.

(c) The pattern since 1987 has been remarkably stable. Poeple who think in terms of monetary policy see a new "stable" regime here.

(d) When the yield/forward curve is upward sloping, interest rates subsequently rise; when yield curve is inverted, interest rates do subsequently fall. The EH looks fairly reasonable.

(e) In 03-04, EH forecast big increases in rates. In 04 and 05 the 1 year rate did rise! Actually 1 year rates rose more than forecast. 5 year rate is dead on the forecast.



Solid = forward curve. Dash = yield curve.

(f) Risk premium *on average?* Table 20.8 Update

Interest rate data 1964:01-2008:12

Maturity n	1	2	3	4	5
$E\left[y^{(n)}\right]$	6.16	6.37	6.54	6.68	6.76
$E\left[y^{(n)} - y^{(1)}\right]$	0	0.21	0.38	0.51	0.59
$E\left[r^{(n)} - y^{(1)}\right]$	0	0.46	0.80	1.38	1.05
$\sigma\left[r^{(n)} - y^{(n)}\right]$	0	1.87	3.42	4.73	5.79
"Sharpe"	0	0.25	0.23	0.22	0.18

There does seem to be a small average risk premium for long term bonds. But it seems small on average. The expectations hypothesis usually means "plus a constant risk premium" and this is that constant risk premium.

(g) Compare Sharpe to 0.5 of the market portfolio. Long term bonds are *way* inside the mean-variance frontier, and $\beta \approx 0$. This makes it a bit of a puzzle why people hold them. Actually long-term bonds are the "safe" asset for long term investors, and this is a warning about one-period mean-variance thinking.

(h) Warning: it's easy to find huge sharpe ratios and "arbitrage" in the term structure; very similar securities with slightly different prices. On/off the run spreads, liquidity spreads, corporate spreads. A bit of a warning, these are a) hard to trade b) hide earthquake risk. Here we're looking at the basic term spread in treasuries. This may be a pretty boring part! I'll review these aspects at the end.

2. Expectations failures–Fama Bliss. (Updated 1964-2008), table 20.9 update

	$rx^{(n)}_{t+1} = a + b\left(f^{(n)}_t - y^{(1)}_t\right) + \varepsilon_{t+1}$			$y^{(1)}_{t+n-1} - y^{(1)}_t = a + b\left(f^{(n)}_t - y^{(1)}_t\right) + \varepsilon_{t+1}$			
n	b	$\sigma(b)$	R^2	a	b	$\sigma(b)$	R^2
2	0.83	0.27	0.12		0.17	0.27	0.01
3	1.12	0.36	0.13		0.47	0.31	0.04
4	1.34	0.45	0.14		0.75	0.23	0.12
5	1.02	0.52	0.06		0.87	0.16	0.16
	forecasting *one* year returns on n-year bonds			forecasting *one* year rates n years from now			

3. Wake up. This is *the central table.*

(a) Left hand panel: "Expected returns on all bonds should be the same" $\rightarrow$ Run a standard regression to see if anything forecasts the difference in return. Run

$$r^{(n)}_{t+1} - y^{(1)}_t = a + bX_t + \varepsilon_{t+1}$$

We should see $b = 0$.

1. Over one year, expected excess returns move *one for one* with $f - s$ spread! (Should be zero.) The "fallacy of yield" turns out to be correct (at a one year horizon).
2. How do we reconcile this with the first table? Long bonds don't earn much more *on average*. But there are *times* when long bonds expect to earn more, and other times when they expect to earn less. $f - s$ is sometimes positive, sometimes negative.

(b) Right hand panel, row 1.

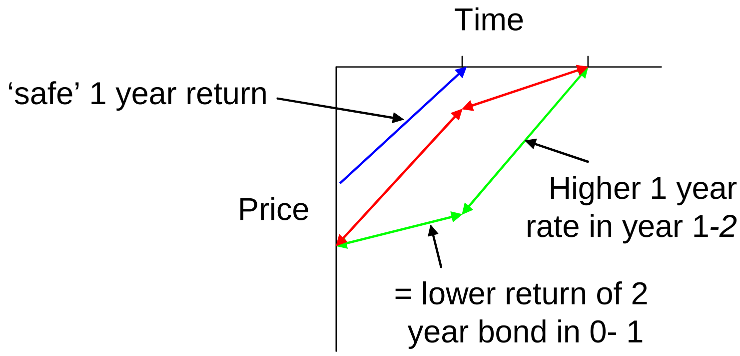
1. If $f^{(1\rightarrow 2)}$ is 1% higher than $y^{(1)}$, we should see $y^{(1)}_{t+1}$ rise 1% higher. We should see $b = 1$. Instead, we see $b = 0$! *A forward rate 1% higher than the spot rate should mean the spot rate rises. Instead, it means that the 2 year bond earns 1% more over the next year on average.*

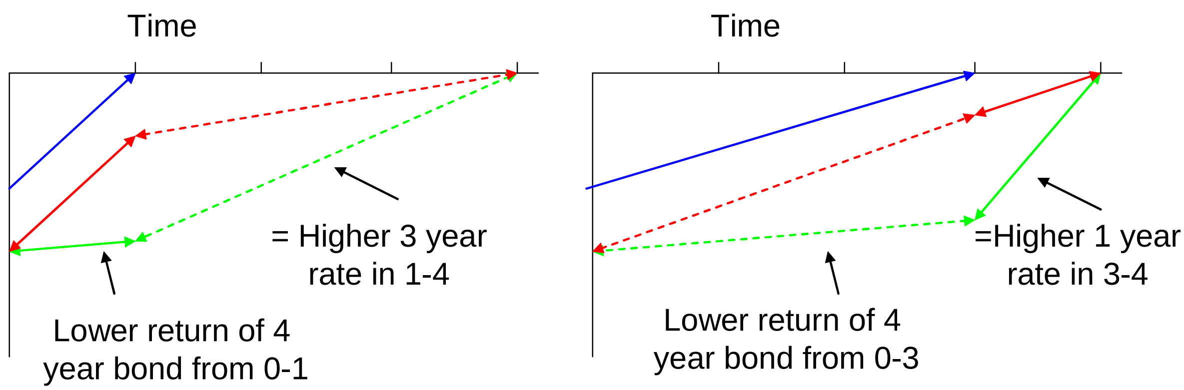
(c) Right hand panel, higher rows

1. At longer horizons, $f - y$ spread *does* start to forecast n year changes in yields. (Correspondingly, it ceases to forecast n year returns – not shown.)
2. Once again, expectations does seem to work "in the long run."

(d) $0.83 + 0.17 = 1$ is not by chance. *Mechanically*, the two coefficients in the first row add up. If $f - y$ *does not* forecast Δy, it *must* forecast returns. See the plot of bonds over time for intuition. If we move the $p^{(1)}_{t+1}$ up we increase expected returns $r^{(2)}_{t+1}$ and decrease the future yield $y^{(1)}_{t+1}$.

(e) Note: higher rows do not add up to 1. Why? There are "complementary" regressions that add up to 1, I just didn't show them.





(f) The algebra: As we did before, start with an identity (exact this time)

$$\left(r_{t+1}^{(2)} - y_t^{(1)}\right) + \left(y_{t+1}^{(1)} - y_t^{(1)}\right) = f_t^{(2)} - y_t^{(1)}$$

Proof: Note that the forward-spot spread equals the change in yield plus the excess return on the two year bond.

$$\begin{aligned} f_t^{(1\to 2)} - y_t^{(1)} &= p_t^{(1)} - p_t^{(2)} + p_t^{(1)} \\ &= \left(p_{t+1}^{(1)} - p_t^{(2)} + p_t^{(1)}\right) + \left(-p_{t+1}^{(1)} + p_t^{(1)}\right) \\ &= \left(r_{t+1}^{(2)} - y_t^{(1)}\right) + \left(y_{t+1}^{(1)} - y_t^{(1)}\right) \end{aligned}$$

Hence, when we project down on to $f^{(2)} - y^{(1)}$, and inventing some notation,

$$b_r^{(2)} + b_y^{(1)} = 1$$

In this precise way, a forward rate higher than the spot rate *must* imply a high return on two year bonds or an increase in the on year rate.

(g) Notice how this is all so much like our dividend yield friends, D/P: (If p.d varies) Δd *should* be forecastable so that R not. Fact: Δd is not forecastable, so R is. Here, (if the yield curve varies) Δy *should* be forecastable so r not. Fact:

Δy is *not* forecastable, so r is. We also split variation in the dividend yiled to two terms *by identity* and deduced that two regression coefficients must add up,

$$
\begin{aligned}
p_t - d_t &= \sum_{j=1}^{\infty} \rho^{j-1} \left(\Delta d_{t+j} - r_{t+j} \right) \\
b_r^{lr} - b_d^{lr} &= 1
\end{aligned}
$$

Here too we had a "complementary regressions" *this is just like the D column of return forecasts.*

4. Q: Why do we run $y_{t+1}^{(1)} - y_t^{(1)}$on $f_t^{(2)} - y_t^{(1)}$? The EH says $f_t^{(1\to 2)} = E_t \left[y_{t+1}^{(1)} \right]$, so why not run $y_{t+1}^{(1)}$ on $f_t^{(1\to 2)}$?

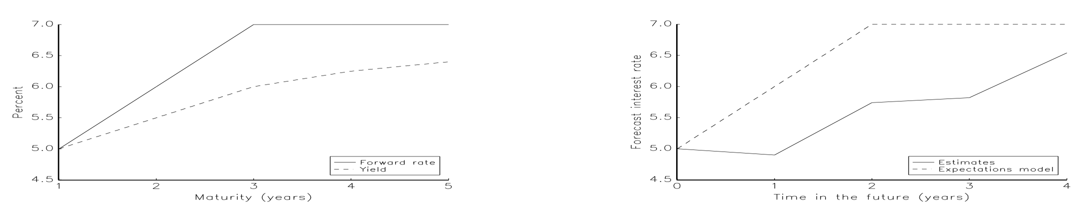
   A: That's valid but not as strong a test. An analogy: $\text{Temp}_{t+1} = 0 + 1 \times \text{Temp}_t + \varepsilon_{t+1}$ works pretty well, so just reporting Temp_t makes you look like a good forecaster! But this won't work for $(\text{Temp}_{t+1} - \text{Temp}_t)$. Being able to forecast *changes* is a more powerful test than being able to forecast *levels* of a slow-moving series like temp, or yield. In short, there's nothing *wrong* with doing it with levels, but differences is a more *powerful* test.

5. Interpretation.

   (a) This is like D/P that forecasts returns, not dividend growth. ($f - y$ forecasts returns, not yield growth.) Variation in D/P *must* be due to returns of dividend growth. ($f - y$ *must* either forecast yield changes or returns.) Surprise in both cases: it's due to returns.

   (b) "Sluggish adjustment." For stocks on D/P; bonds on F-S and exchange rates on domestic-foreign interest rate, an expected offsetting adjustment doesn't happen.

      1. Stocks: P/D high? D should rise. It doesn't. "Buy yield" ("value").
      2. Bonds: $f - y$ high? y should rise. It doesn't (at least for a few years). "Buy yield"



If the current yield curve is as plotted in the left hand panel, the right hand panel gives the forecast of future one year interest rates. This is based on the right hand panel of the above table. The dashed line in the right hand panel gives the forecast from the expectations hypothesis, in which case forward rates today are the forecast of future spot rates.

(c) There are big risk premia and they vary over time. There are times when you expect much better returns on long term bonds, and other times when you expect much better returns on short term bonds. (An upward sloping yield curve results in rising rates, but not soon enough – you earn the yield in the meantime). When are these times? Look at the plot! High expected returns in recessions (level recessions, not growth rate recessions) Interpretation: Business cycle related variation in risk, expected returns. (γ varies over time). Who wants to hold long-term bonds in the bottom of a recession?

6. ADDITIONS/PROBLEM SETS.

   (a) Note Fama Bliss with coefficient = 1 is NOT consistent, since if you iterate you don't get one year regresisons.

   (b) Note the one factor Vasicek does not produce Fama Bliss coefficient.

   (c) Result. FB coefficients should be less than one.

Cochrane-Piazzesi 1

Let's move the question.

Now, let's ask the question, "what best forecasts returns?" Let's put all available forwards on the right hand side, not just the maturity-matched pair.

It was *right* for FF not to do this. Their question was "does the forward spread forecast Δy?" "What causes variation in the right hand side?" (Fama regressions are really about this – *why does D/P, f-s move?*) Also, we're not really asking, "is the expectations hypothesis right?" That was FB's motivation. We're asking, "given it's wrong, what's the best characterization of bond expected returns?"

Motivation 2. (More important) Now we have many right and left hand variables. *We need to look for common factors.*

$$\begin{aligned} r^1_{t+1} &= a + bx^1_t + \varepsilon^1_{t+1} \\ r^2_{t+1} &= a + bx^2_t + \varepsilon^2_{t+1} \end{aligned}$$

$x = f - s$ for fama bliss, dp vs. $f - s$ looking across stocks and bonds.

a) Does x^1 enter equation 2?

b) Maybe

$$\begin{aligned} r^1_{t+1} &= a + bx^1_t + cx^2_t + \varepsilon^1_{t+1} \\ r^2_{t+1} &= a + 2bx^2_t + 2cx^2_t + \varepsilon^2_{t+1} \end{aligned}$$

If so the expected returns are perfectly correlated.

What is the *factor structure in expected returns?*

We want to ask this question across all asset classes. Asking across bond maturities is the start.

Cochrane-Piazzesi update

Bottom line

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

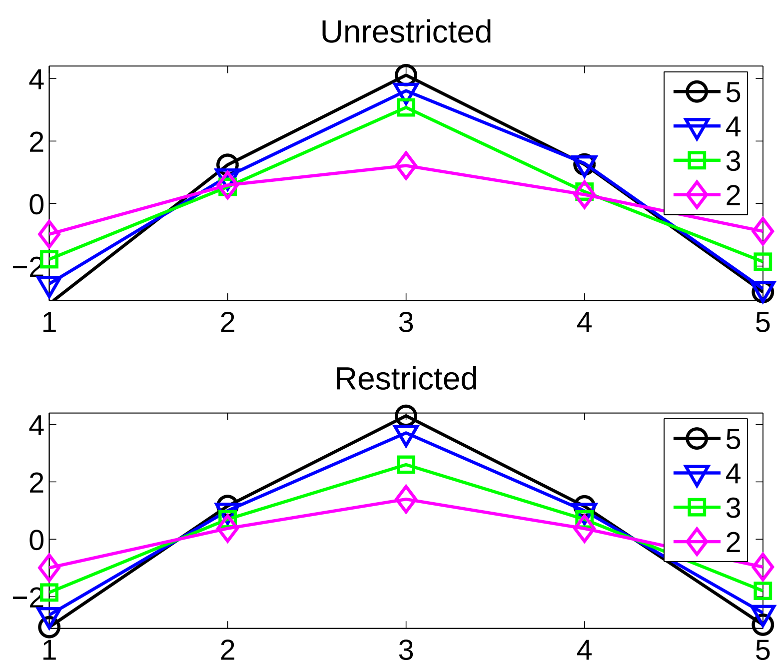
$$rx_{t+1}^{(n)} = a_n + \beta_n' f_t + \varepsilon_{t+1}^{(n)}$$

2. R^2 up to 44%, up from Fama-Bliss / Campbell Shiller 15%

3. A *single* "factor" $\gamma' f$ forecasts bonds of *all* maturities. High expected returns in "bad times."

4. A 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!

• To see the point, what would Fama-Bliss coefficients look like?

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 $\gamma' f$ tells you where all expected returns are going at any date. Then stretch it up more for long term bonds 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.

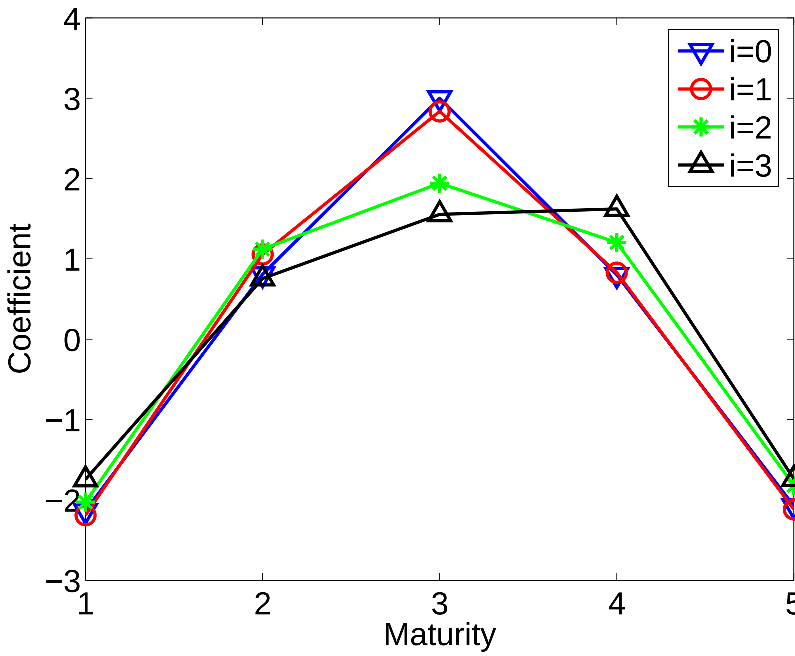
•Restricted model $b_n\gamma$ almost perfectly matches unrestricted coefficients. (well below 1σ) (This is the point of the graph)

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

$\bullet R^2$ almost unaffected by the single-factor restriction. The restriction looks good in the graph.

$\bullet$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



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

$\bullet$More lags are significant, with the same pattern.

$\bullet$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.

$\bullet$The pattern 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 about the future. 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. Moving averages are a good way to enhance a slow-moving "signal" buried in high-frequency "noise".

Stock Return Forecasts

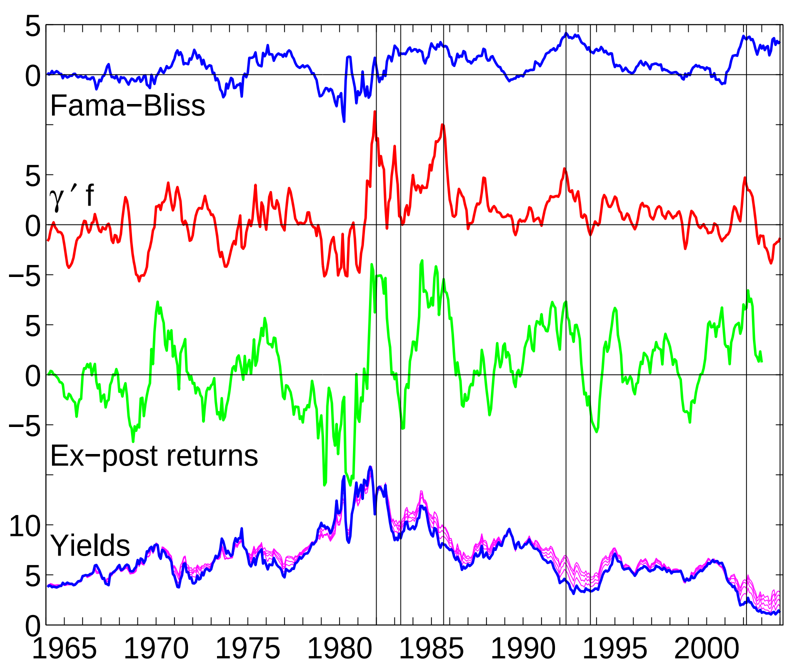
Table 3. Forecasts of excess stock returns (VWNYSE)

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

	$\gamma^\top f$	(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

- The 5 year bond had $b = 1.43$. Thus, $1.73 - 2.11$ is what you expect for a perpetuity.
- It does better than D/P and term spread: Drives out spread; It survives with D/P
- A common term risk premium in stocks, bonds! Reassurance on fads & measurement errors
- *Need to follow up – do impulse responses like I did for cay*

History

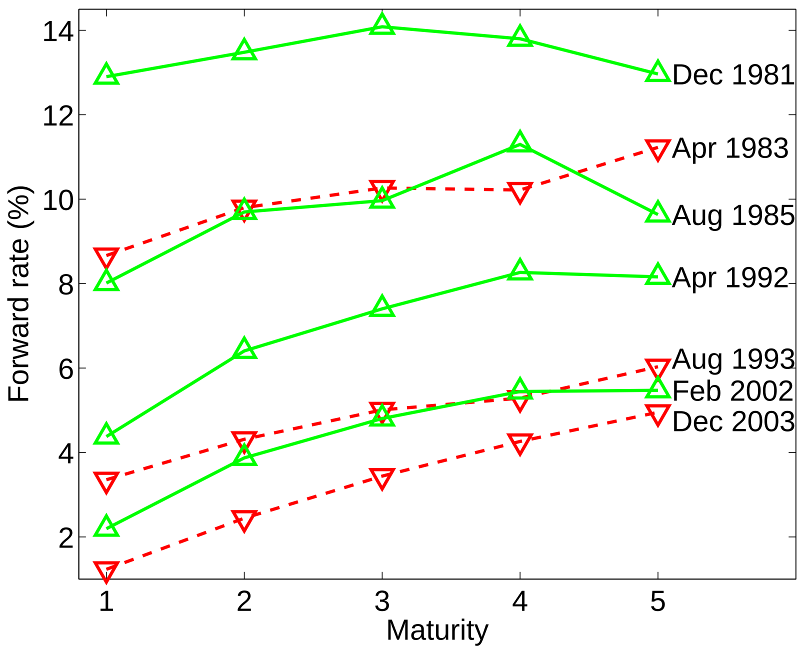


• Is it real or just a few data points? What is the story?

•Consistent in many episodes

•$\gamma' f$ and slope are correlated. Many episodes are interpreted the same way.

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

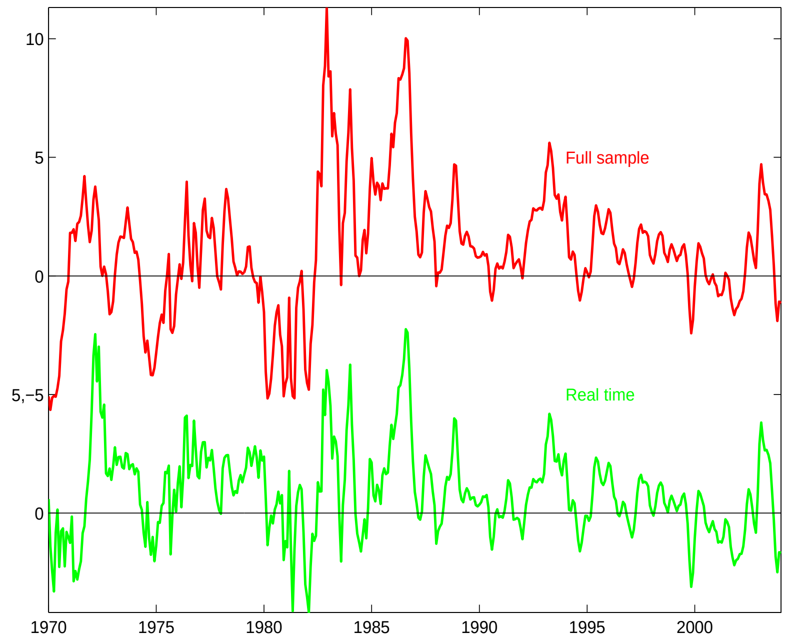


•Green: CP say buy. Red: FB say buy, CP say don't.

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

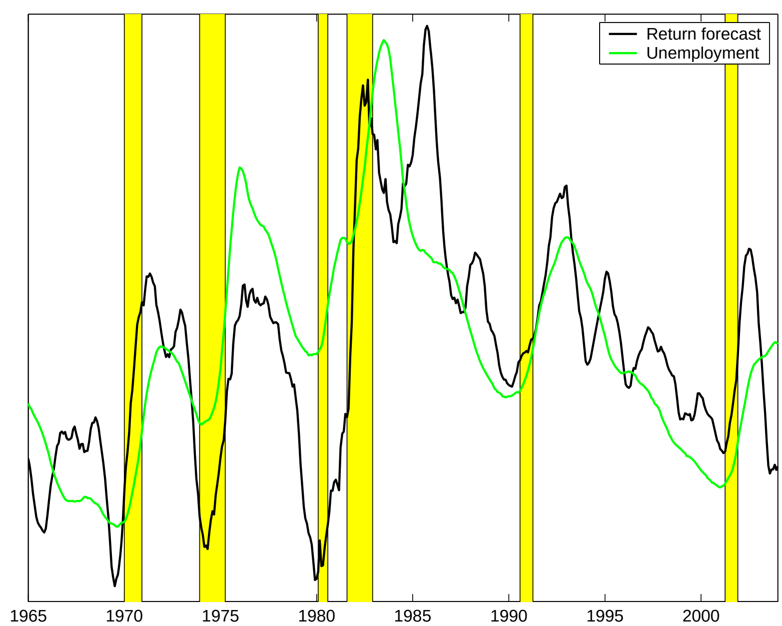
•CP: in the past, tent-shape often came with upward slope. Others saw upward slope, thought that was the signal. But an upward slope without a tent does not work. The tent is the real signal.

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.)

Macro



•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!

Yield curve factor models.

The usual approach: By eigevalue decomposition/factor analysis reduce the yield curve to "level" "slope" and "curvature" factors which explain 99.9% of yield curve variance.

$$\begin{aligned} l_t &= q_l' y_t; s_t = q_s' y_t; c_t = q_c' y_t \\ y_t &= q_l l_t + q_s s_t + q_c c_t \ (+\varepsilon_t) \end{aligned}$$

Theorem,

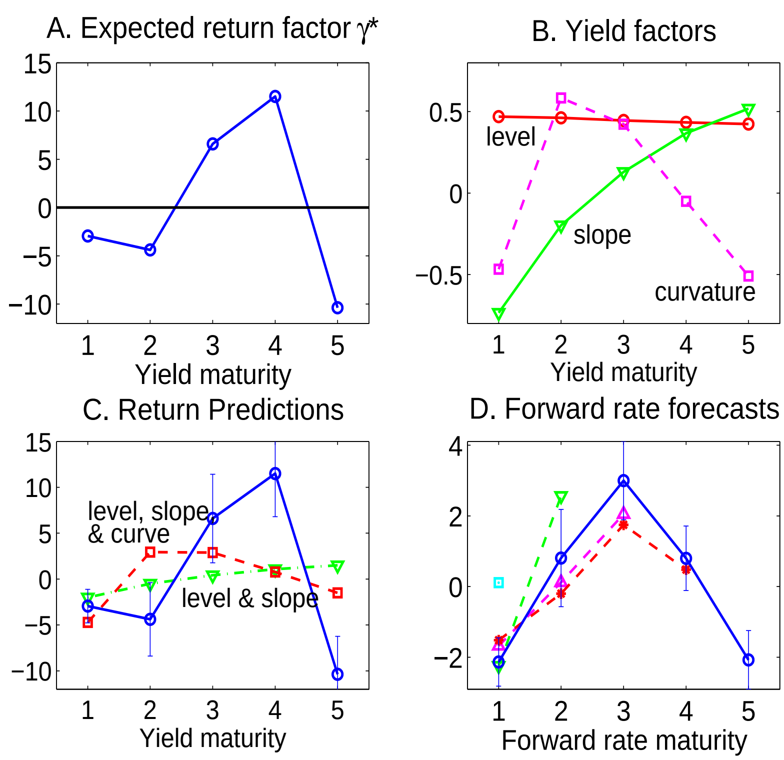
$$Q \Lambda Q' = cov(y_t y_t')$$

the columns of Q provide these numbers.

In this approach, we usually find that s_t forecasts returns (Fama-Bliss slopes) but with about 15% R^2

Similar structures come from affine models, which we'll look at in a bit.

How does the CP $\gamma' f_t = \gamma^{*\prime} y_t$ compare to this approach?



• Top left: γ expressed as an operator for *yields* not *forwards*. (Not so pretty, eh?)

- Top right: q expressed as a function of yields – how to *form* factors from yields, and how yield *load* on factors.

- Bottom left: what function of level, slope and curve best forecasts returns?

$$rx_{t+1} = a + b(l_t) + c(s_t) + d(c_t) + \varepsilon_{t+1}$$

  Then unwind the factors to find the overall function of yields that best forecasts returns.

Lesson.

1. *First* smooth, *then* forecast can throw out the baby with the bathwater. This contradicts the usual practice (Stock and Watson) to first find principal components then forecast. But there was no theory that usual practice was right; maybe the parts that forecast do not explain much contemporaneous variance.

2. "Maximize variance of yields, returns, etc." is not "maximize variance of expected returns"

3. Expected returns are (nearly) "unspanned state variables" a big new issue in term structure models.

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

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)**

Right hand variable spans columns $y_t^{(1)}$ through $y_t^{(5)}$.

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

•This is *statistically significant*, and this is why GMM rejects the single factor model.

• 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.

Measurement error, ϕ^{12} problems, etc.

(See paper.)

Latest data, and treasury curves during the crash

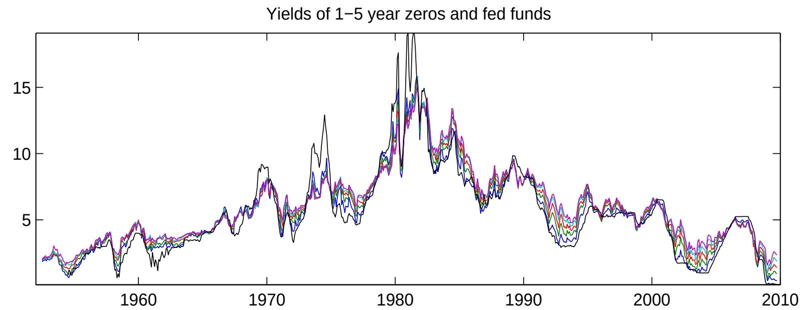
See overheads. This is a real problem. *The Fama Bliss data extraction method no longer works. We need to go back to basic treasury data and do it right.*

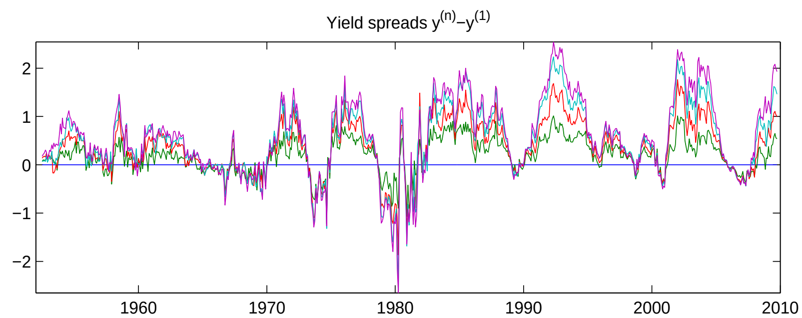
Affine models, background for CP II

Statistical factor model approach

Produces "factor model."

The yield curve is crying for factor structure





- 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. (Note "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

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

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. (Note: Unit variance factors

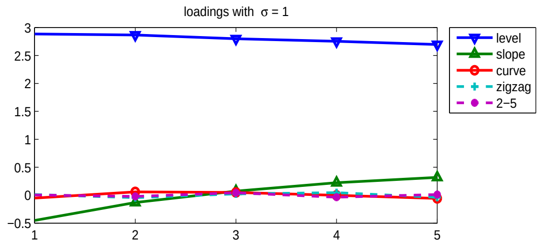
$$\begin{aligned} y_t &= Q\Lambda^{1/2}x_t; cov(x_t x_t') = I \\ x_t &= \Lambda^{-\frac{1}{2}}Q'y \end{aligned}$$

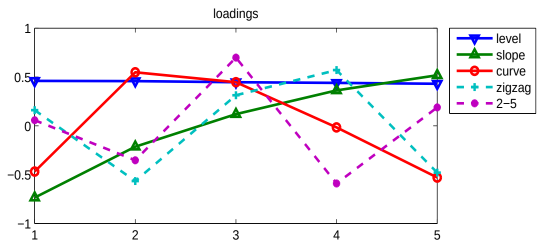
Loadings $Q\Lambda^{1/2}$ are smaller for smaller factors. This is a nice way to show the relative importance as well as the shapes of the factors, and not have your readers spend too much time interpreting the 5th factor.)

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

- Result, applied to FB yield data

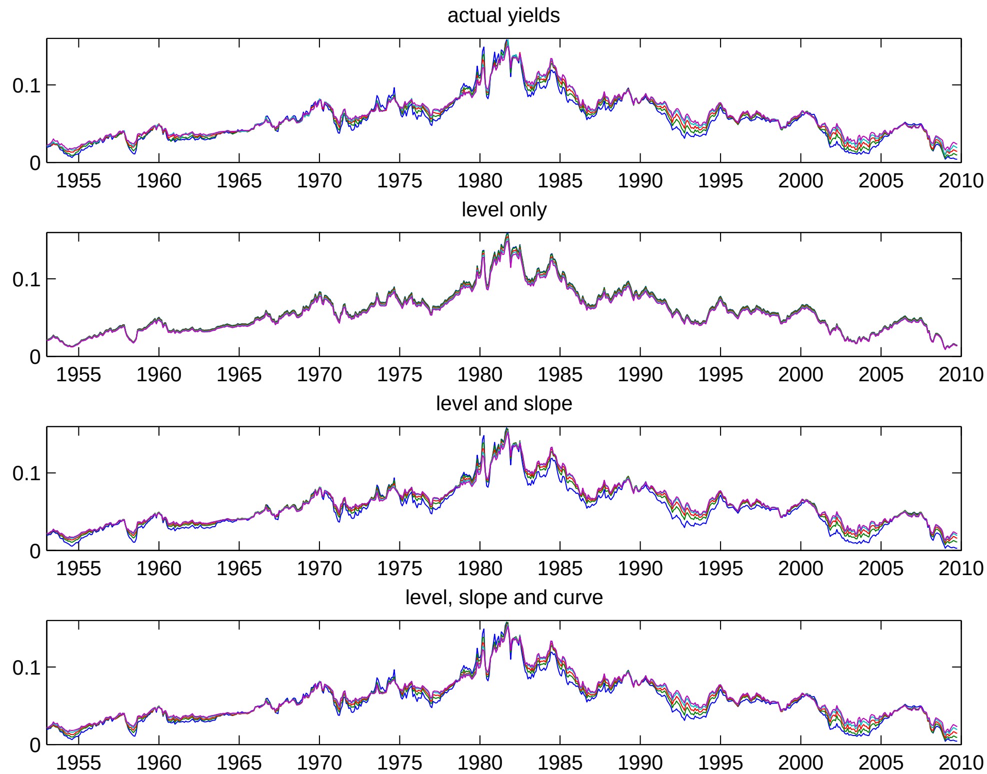
```
Sigma = cov(100*yields);
[Q,L] = eig(Sigma);
disp(diag(L)'.^0.5);
%(my names)2-5          zigzag        curve         slope         level
         0.06         0.07          0.10          0.58          5.80
disp(Q)
         0.06         0.15         -0.47         -0.74          0.46
        -0.35        -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
loads = Q*L^0.5;
plot(loads)
plot(Q)
```

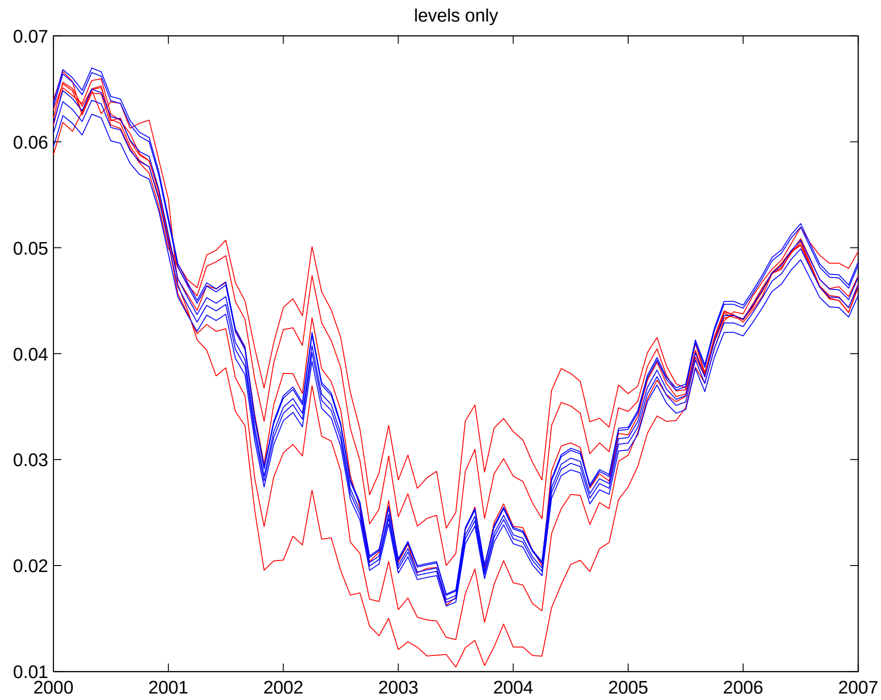


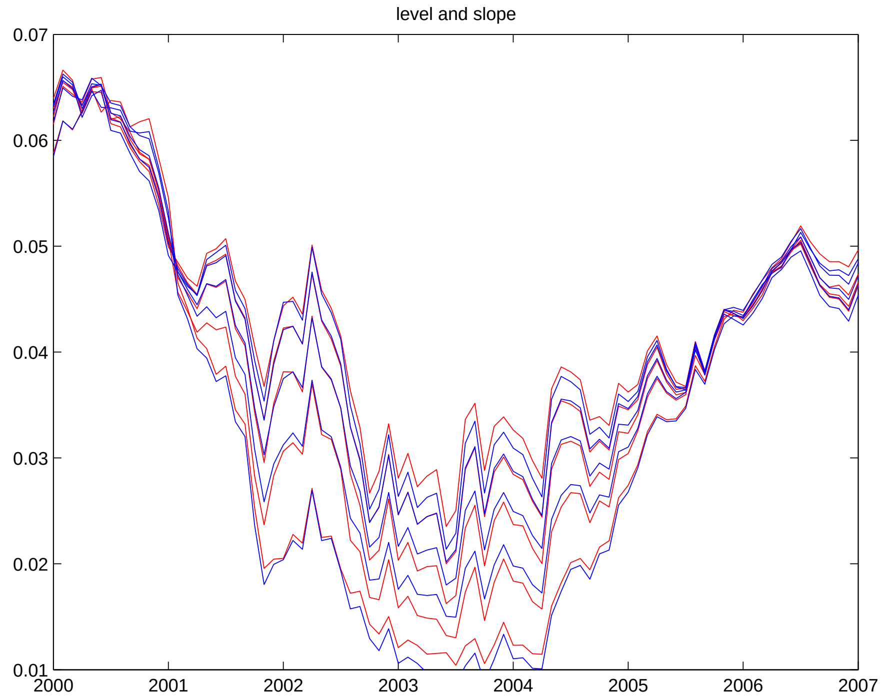


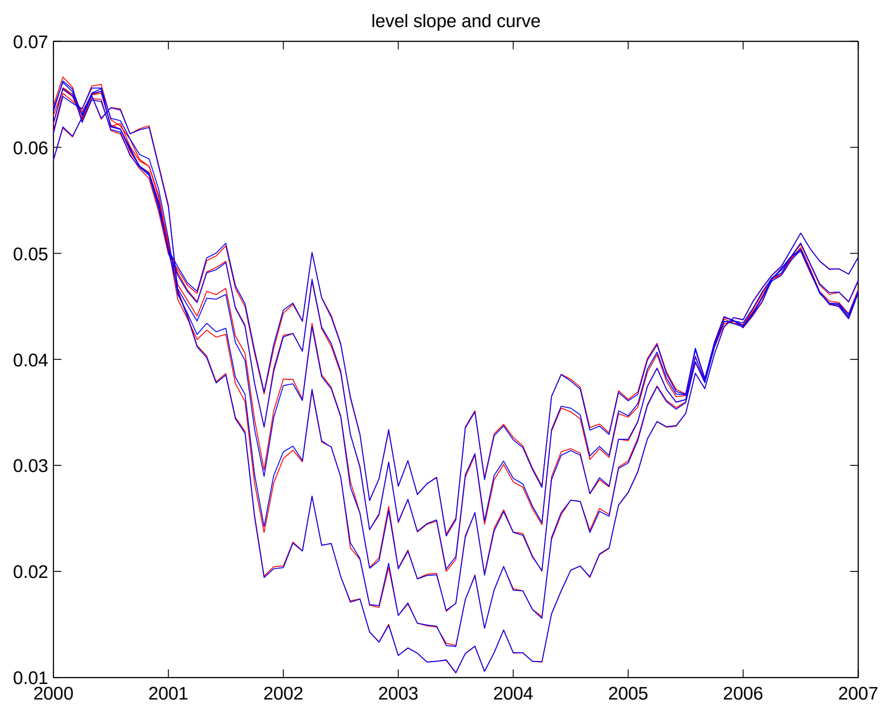
- "Factor models" come from dropping the small eigenvalues. Then a larger number of series are, exactly, driven by a smaller number of factors.
- For example, what if we drop 4 and 5?

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


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 2010


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.

Motivation for term structure models

1. Let's simplify for a bit to an exact factor model,

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

   It suggests a simple description of the yield curve, boil it all down to

$$\begin{aligned} X_t &= \left[\begin{array}{ccc} l_t & s_t & c_t \end{array} \right]' \\ X_{t+1} &= \mu + \Phi X_t + \varepsilon_{t+1} \\ y_t(N \times 1) &= q(N \times 3) X_t(3 \times 1) \end{aligned}$$

2. But

   (a) Can you extend this to other maturities, not included? Since there are 3 factors, spanned, there is only one *arbitrage-free extension.* Can you extend to options – which must also be a function of X?

   (b) *The point of most affine models: Like Black and Scholes, find an arbitrage-free interpolation and extension. For B-S, it was (stock, bond -> option) For us, it is from a few bonds to all maturities, and then to term structure options, whose value only depends on the current yield curve and can be replicated by dynamic trading of bonds.*

   (c) Note: if all you care about is running an arbitrage-free curve through today's bonds (and options) *you don't care about risk premia.* You don't care about the real measure; do everything under risk neutral measure. *But then ignore forecasting.* Real vs. risk neutral measure *only matters if you care about forecasting.*

Term structure models

Note to self: Needs a consistent notation once and for all with my model with piazzesi. Use risk neutral notation for expectations model.

Macro approach

$$P_t^{(N)} = E_t \left(\beta^N \frac{u'(c_{t+N}, m_{t+N})}{u'(c_t)} \right)$$

1. Just test
2. Model

$$X_t = \begin{bmatrix} c_t \\ x_{2t} \\ x_{3t} \end{bmatrix}; X_{t+1} = \Phi X_t + \varepsilon_{t+1};$$

   compute $P_t^{(N)}$?

   (a) Won't be equal to exact $P_t^{(N)}$
   (b) Agents see more than we do, maybe $P_t^{(N)}$ should be in X?

3. Useful for "deep explanation," but there are other uses...This is more than we need for the "arbitrage-free extension" project.

Expectations hypothesis model

This is the simplest "model" to illustrate the logic. The ingredients are a) short rate process b) expectations to get other prices

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

$$f_t^{(2)} = E_t(y_{t+1}^{(1)}) = \delta + \phi(y_t^{(1)} - \delta)$$

$$f_t^{(3)} = E_t(y_{t+2}^{(1)}) = \delta + \phi^2(y_t^{(1)} - \delta)$$

$$f_t^{(N)} = E_t(y_{t+N-1}^{(1)}) = \delta + \phi^{N-1}(y_t^{(1)} - \delta)$$

(And yields,...)

$$\begin{aligned} y_t^{(2)} &= \frac{1}{2}\left[E_t(y_{t+1}^{(1)}) + y_t^{(1)}\right] \\ &= \frac{1}{2}\left[\delta + \phi(y_t^{(1)} - \delta) + y_t^{(1)}\right] \\ y_t^{(2)} - \delta &= \frac{1+\phi}{2}\left(y_t^{(1)} - \delta\right) \\ &\ldots \\ y_t^{(N)} - \delta &= \frac{1+\phi+\phi^2+..}{N}\left(y_t^{(1)} - \delta\right) = \frac{1-\phi^N}{N(1-\phi)}\left(y_t^{(1)} - \delta\right) \end{aligned}$$

Result:

1. Different shapes of the term structure, upward and downward sloping. Up when short rates expected to rise.
2. More complex shapes? Move past AR(1)!

3. A "one factor model" of the yield curve. We obviously could generalize to "multifactor" with a more complex time-series process.

$$
\begin{aligned}
\begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} &= X_t \\
X_t &= \Phi X_{t-1} + \varepsilon_t \\
y_t^{(1)} &= \delta_0 + \delta_1' X_t
\end{aligned}
$$

4. No average slope – $E(y^{(i)})$ all the same. Well, we imposed expectations!

5. "Arbitrage?" We're almost there. If the expectations model were the risk neutral limit, then we could appeal to the risk neutral density theorem and claim these are arbitrage-free prices under the risk neutral density. Alas, they're not..

Discrete-time single-factor Vasicek

Here: A standard "single factor model" – "discrete-time Vasicek." The end result:

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

$$
\begin{aligned}
f_t^{(2)} &= \delta + \phi\left(y_t^{(1)} - \delta\right) - \left[\frac{1}{2} + \lambda\right]\sigma_\varepsilon^2 \\
f_t^{(3)} &= \delta + \phi^2\left(y_t^{(1)} - \delta\right) - \left[\frac{1}{2}(1+\phi)^2 + \lambda(1+\phi)\right]\sigma_\varepsilon^2 \\
&\ldots.
\end{aligned}
$$

$$
\begin{aligned}
y_t^{(2)} &= \delta + \frac{(1+\phi)}{2}\left(y_t^{(1)} - \delta\right) - \frac{1}{2}\left(\frac{1}{2} + \lambda\right)\sigma_\varepsilon^2 \\
y_t^{(3)} &= \delta + \frac{(1+\phi+\phi^2)}{3}\left(y_t^{(1)} - \delta\right) - \frac{1}{3}\left\{\frac{1}{2}\left[1 + (1+\phi)^2\right] + \lambda\left[1 + (1+\phi)\right]\right\}\sigma_\varepsilon^2 \\
&\ldots.
\end{aligned}
$$

Intuition for now: the first equation tells you where interest rates are going over *time*. The second and third sets of equations tell you where each forward rate is at any date, depending only on where the short rate is on that date; a single factor model.

Derivation:

- Suppose m follows the time series model.

$$
\begin{aligned}
x_{t+1} - \delta &= \phi\left(x_t - \delta\right) + \varepsilon_{t+1} \\
m_{t+1} &= \ln M_{t+1} = -\frac{1}{2}\lambda^2\sigma_\varepsilon^2 - x_t - \lambda\varepsilon_{t+1}
\end{aligned}
$$

  This is just a model for m, with a convenient "state variable" x_t. Intuition: x_t shifts the *mean* of $\ln m_{t+1}$ around. If you remember that $R^f = 1/E(m)$, you can see that specifying a model for the mean of m is the key to thinking about interest rates. The $1/2\sigma^2$ term just shifts the mean of $\ln m_t$ down, and offsets a $1/2\sigma^2$ term which will pop up later. To be specific, ε and v are iid Normal with $\sigma_\varepsilon^2, \sigma_v^2, \sigma_{\varepsilon v}$. δ, ρ, λ are free parameters; we'll pick these to make the model fit as well as possible.

- An interpretation:

$$
\begin{aligned}
x_{t+1} - \delta &= \phi (x_t - \delta) + \varepsilon_{t+1} \\
\Delta c_{t+1} &= \frac{1}{\gamma}\left(-\delta + x_t + \frac{1}{2}\lambda^2\sigma_\varepsilon^2 + \lambda\varepsilon_{t+1}\right) \\
&\rightarrow\ -\delta - \gamma\Delta c_{t+1} = -x_t - \frac{1}{2}\lambda^2\sigma_\varepsilon^2 - \lambda\varepsilon_{t+1}
\end{aligned}
$$

"suppose consumption growth followed this ARMA (1,1) process, *but we didn't observe consumption growth.* Like all finance (CAPM etc) ultimately we're abstracting away the connection to macro. Obviously, if we want to "explain" asset prices, we need to put back in tne connection to macro!

- Bond prices.

$$P_t^{(n)} = E_t\left(M_{t+1}M_{t+2}....M_{t+n}\right)$$

This is easier to do recursively,

$$
\begin{aligned}
P_t^{(0)} &= 1 \\
P_t^{(n)} &= E_t\left(m_{t+1}P_{t+1}^{(n-1)}\right).
\end{aligned}
$$

- Here we go.

$$P_t^{(1)} = E_t\left(M_{t+1}\right) = E_t e^{m_{t+1}}$$

$$P_t^{(1)} = e^{-\frac{1}{2}\lambda^2\sigma_\varepsilon^2 - x_t + \frac{1}{2}\lambda^2\sigma_\varepsilon^2} = e^{-x_t}$$

$$
\begin{aligned}
p_t^{(1)} &= \ln E_t\left(m_{t+1}\right) = -x_t \\
y_t^{(1)} &= x_t
\end{aligned}
$$

Now you see why I set up the problem with the $1/2\lambda^2\sigma_\varepsilon^2$ to begin with! *The one year interest rate "reveals the latent state variable x_t"*

- I could have written the model as

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

"a short rate process plus a market price of risk." Then, by taking $-\ln E_t\left(M_{t+1}\right)$ I would have checked that the $y_t^{(1)}$ the model *produces* is the same $y_t^{(1)}$ I *started* with. Take your pick. Which is more confusing: a) starting with an x_t you "can't see" and then showing that it turns out to be $y_t^{(1)}$? b) starting with an assumed $y_t^{(1)}$ process and then showing that it's in fact the one period rate, that the model is "self-consistent" (in the language of CP appendix.)

- On to the next price.

$$
\begin{aligned}
P_t^{(2)} &= E_t\left(m_{t+1}P_{t+1}^{(1)}\right) = E_t\left(e^{m_{t+1}+p_{t+1}^{(1)}}\right) \\
&= E_t\left(e^{-\frac{1}{2}\lambda^2\sigma_\varepsilon^2 - x_t - \lambda\varepsilon_{t+1} - x_{t+1}}\right)
\end{aligned}
$$

$$P_t^{(2)} = E_t\left(e^{-\frac{1}{2}\lambda^2\sigma_\varepsilon^2 - x_t - \lambda\varepsilon_{t+1} - \delta - \phi(x_t-\delta) - \varepsilon_{t+1}}\right)$$

$$P_t^{(2)} = E_t\left(e^{-2\delta - (1+\phi)(x_t-\delta) - \frac{1}{2}\lambda^2\sigma_\varepsilon^2 - (1+\lambda)\varepsilon_{t+1}}\right)$$

$$
\begin{aligned}
p_t^{(2)} &= -2\delta - (1+\phi)(x_t-\delta) - \frac{1}{2}\lambda^2\sigma_\varepsilon^2 + \frac{1}{2}(1+\lambda)^2\sigma_\varepsilon^2 \\
p_t^{(2)} &= -2\delta - (1+\phi)(x_t-\delta) + (\frac{1}{2}+\lambda)\sigma_\varepsilon^2
\end{aligned}
$$

- From prices, we find yields and forwards,

$$y_t^{(2)} = \delta + \frac{(1+\phi)}{2}(x_t-\delta) - \frac{1}{2}\left(\frac{1}{2}+\lambda\right)\sigma_\varepsilon^2$$

$$
\begin{aligned}
f_t^{(2)} &= p_t^{(1)} - p_t^{(2)} \\
&= -\delta - (x_t-\delta) + 2\delta + (1+\phi)(x_t-\delta) - (\frac{1}{2}+\lambda)\sigma_\varepsilon^2 \\
&= \delta + \phi(x_t-\delta) - (\frac{1}{2}+\lambda)\sigma_\varepsilon^2
\end{aligned}
$$

- Now the rest of the maturities. You can "solve the discount rate forward and integrate"

$$p_t^{(3)} = \log E_t\left(M_{t+1}M_{t+2}M_{t+3}\right)$$

$$
\begin{aligned}
p_t^{(3)} &= \log E_t e^{-\frac{1}{2}\lambda^2\sigma_\varepsilon^2 - x_t - \lambda\varepsilon_{t+1} - \frac{1}{2}\lambda^2\sigma_\varepsilon^2 - x_{t+1} - \lambda\varepsilon_{t+2} - \frac{1}{2}\lambda^2\sigma_\varepsilon^2 - x_{t+2} - \lambda\varepsilon_{t+3}} \\
&= \log E_t e^{-3\delta - \frac{3}{2}\lambda^2\sigma_\varepsilon^2 - (1+\phi+\phi^2)(x_t-\delta) - \lambda\varepsilon_{t+1} - \lambda\varepsilon_{t+2} - \lambda\varepsilon_{t+3} - (1+\phi)\varepsilon_{t+1} - \varepsilon_{t+2}}
\end{aligned}
$$

  This will work after much algebra

- Instead, let's do it recursively "derive a differential equation for price as a function of state variables." Guess

$$P_t^{(n)} = A_n - B_n(x_t-\delta)$$

  then

$$P_t^{(n)} = E_t\left(M_{t+1}P_{t+1}^{(n-1)}\right)$$

  (Not in lecture:

$$
\begin{aligned}
A_n - B_n(x_t-\delta) &= \log E_t\left(\exp\left(-\frac{1}{2}\lambda^2\sigma_\varepsilon^2 - x_t - \lambda\varepsilon_{t+1}\right)\exp\left(A_{n-1} - B_{n-1}(x_{t+1}-\delta)\right)\right) \\
&= \log E_t\left(\exp\left(-\frac{1}{2}\lambda^2\sigma_\varepsilon^2 - x_t - \lambda\varepsilon_{t+1} + A_{n-1} - B_{n-1}\phi(x_t-\delta) - B_{n-1}\varepsilon_{t+1}\right)\right) \\
&= \log E_t\left(\exp\left(-\delta + A_{n-1} - (1+B_{n-1}\phi)(x_t-\delta) - \frac{1}{2}\lambda^2\sigma_\varepsilon^2 - \lambda\varepsilon_{t+1} - B_{n-1}\varepsilon_{t+1}\right)\right. \\
A_n - B_n(x_t-\delta) &= -\delta + A_{n-1} - (1+B_{n-1}\phi)(x_t-\delta) + \left(B_{n-1}\lambda + \frac{1}{2}B_{n-1}^2\right)\sigma_\varepsilon^2
\end{aligned}
$$

The constant and the term multiplying x_t must separately be equal. Thus,

$$B_n = 1 + B_{n-1}\phi$$
$$A_n = -\delta + A_{n-1} + \left(B_{n-1}\lambda + \frac{1}{2}B_{n-1}^2\right)\sigma_\varepsilon^2$$

We have "transformed the solution of a stochastic differential equation plus integral to the solution of an ordinary differential equqation.

- That's easy to solve

$$\begin{aligned} B_0 &= 0 \\ B_1 &= 1 \\ B_2 &= 1+\phi \\ B_3 &= 1+\phi+\phi^2 \\ B_n &= \sum_{j=0}^{n-1}\phi^j = \frac{1-\phi^n}{1-\phi} \end{aligned}$$

$$\begin{aligned} A_0 &= 0 \\ A_1 &= -\delta \\ A_2 &= -2\delta + \left(\lambda + \frac{1}{2}\right)\sigma_\varepsilon^2 \\ A_3 &= -3\delta + \left[(1+\phi)\lambda + \frac{1}{2}(1+\phi)^2\right]\sigma_\varepsilon^2 \\ A_4 &= -4\delta + \left[\left(1+\phi+\phi^2\right)\lambda + \frac{1}{2}\left(1+\phi+\phi^2\right)^2\right]\sigma_\varepsilon^2 \end{aligned}$$

You see the pattern from here

- Yields, forwards, returns, etc. follow. $y_t^{(n)} = -1/n \times p_t^{(n)}$. Forwards are even simpler,

$$\begin{aligned} f_t^{(n)} &= p_t^{(n-1)} - p_t^{(n)} \\ &= (A_{n-1} - A_n) - (B_{n-1} - B_n)(x_t - \delta) \\ &= -\delta + \left(B_{n-1}\lambda + \frac{1}{2}B_{n-1}^2\right)\sigma_\varepsilon^2 + \phi^{n-1}(x_t - \delta) \end{aligned}$$

- Result:

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

$$\begin{aligned} f_t^{(2)} &= \delta + \phi\left(y_t^{(1)} - \delta\right) - \left(\frac{1}{2} + \lambda\right)\sigma_\varepsilon^2 \\ f_t^{(3)} &= \delta + \phi^2\left(y_t^{(1)} - \delta\right) - \left[\frac{1}{2}(1+\phi)^2 + \lambda(1+\phi)\right]\sigma_\varepsilon^2 \\ f_t^{(4)} &= \delta + \phi^3\left(y_t^{(1)} - \delta\right) - \left[\frac{1}{2}\left(1+\phi+\phi^2\right)^2 + \lambda(1+\phi+\phi^2)\right]\sigma_\varepsilon^2 \\ &\ldots \end{aligned}$$

$$
\begin{aligned}
y_t^{(2)} &= \delta + \frac{(1+\phi)}{2}\left(y_t^{(1)} - \delta\right) - \frac{1}{2}\left(\frac{1}{2} + \lambda\right)\sigma_\varepsilon^2 \\
y_t^{(3)} &= \delta + \frac{(1+\phi+\phi^2)}{3}\left(y_t^{(1)} - \delta\right) - \frac{1}{3}\left\{\frac{1}{2}\left[1 + (1+\phi)^2\right] + \lambda\left[1 + (1+\phi)\right]\right\}\sigma_\varepsilon^2 \\
&\ldots
\end{aligned}
$$

1. Just like EH but we now see a now a risk premium emerge!
2. Shapes: A steady decline from σ^2 terms, (risk premim) + exponential decay from $y^{(1)} - E(y^{(1)})$. (expectations hypothesis)
3. "Short rate process" plus "one factor model." All yields move in lockstep indexed by $y_t^{(1)}$ (or any other yield). The shape is tied to the level. It looks like "you can price other bonds by arbitrage" but that is only because we restrict our model to have one factor. *"Arbitrage free" pricing in the term structure always comes down to the assumption that there is a perfect factor structure*
4. $y^{(1)}$ is also sufficient to *forecast* all yields.
5. The risk premium comes from $cov(m, y^{(1)}) = \lambda\sigma_\varepsilon^2$ "market price of interest rate risk". If there were a security whose payoff were ε_{t+1} its price would be driven by $cov(m, \varepsilon_{t+1})$.
6. The premium can go either way depending on the sign of λ. My guess: lower $y_{t+1}^{(1)}$means higher m (bad state) means + ($m = .. - \varepsilon$) sign and negative premium. This is a typical result. The real term structure ought to slope down, as long term bonds are safer for long-term investors. However, as long as we separate market prices of risk from consumption and interest data (as we did with the CAPM!) we can incorporate an upward sloping yield curve with $\lambda < 0$
7. The risk premium is constant over time though – as we'll see not in data.
8. "Risk neutrality" $\lambda = 0$ does not mean "expectations" since there is another term. This is Another force for typical downward slope. However, it's quantitatively very small, since $\sigma_\varepsilon \approx 0.01$
9. Another way to see risk premia is to look at returns,

$$
\begin{aligned}
r_{t+1}^{(n)} &= p_{t+1}^{(n-1)} - p_t^{(n)} = (A_{n-1} - A_n) - B_{n-1}(x_{t+1} - \delta) + B_n(x_t - \delta) \\
&= \delta - \left(B_{n-1}\lambda + \frac{1}{2}B_{n-1}^2\right)\sigma_\varepsilon^2 - B_{n-1}(x_{t+1} - \delta) + B_n(x_t - \delta)
\end{aligned}
$$

Expected returns

$$
\begin{aligned}
E_t r_{t+1}^{(n)} &= \delta - \left(B_{n-1}\lambda + \frac{1}{2}B_{n-1}^2\right)\sigma_\varepsilon^2 + (B_n - B_{n-1}\phi)(x_t - \delta) \\
E_t r_{t+1}^{(n)} &= \delta - \left(B_{n-1}\lambda + \frac{1}{2}B_{n-1}^2\right)\sigma_\varepsilon^2 + (x_t - \delta) \\
E_t rx_{t+1}^{(n)} &= -\left(B_{n-1}\lambda + \frac{1}{2}B_{n-1}^2\right)\sigma_\varepsilon^2
\end{aligned}
$$

You see the expected returns differ by maturity, but the risk premium is constant over time – not what Fama-Bliss find.

10. The limiting yield and forward rate are constants. There is no true "level" shift. We'll see this is quite general – "level shifts" imply an arbitrage opportunity at the long end of the yield curve.

11. Yields can be negative – they are normally distributed here. $M > 0$ means $P > 0$ not $P < 1$. The CIR model fixes this up.

12. As a reminder,

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

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

Now, suppose you write a new model with

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

with a new and different mean,

$$\delta^* = \delta - \frac{\sigma_\varepsilon^2}{1-\phi}\lambda$$

and then you set $\lambda = 0$. You'd get

$$\begin{aligned}
f_t^{(2)} &= \left(\delta - \frac{\sigma_\varepsilon^2}{1-\phi}\lambda\right) + \phi\left(y_t^{(1)} - \left(\delta - \frac{\sigma_\varepsilon^2}{1-\phi}\lambda\right)\right) - \left(\frac{1}{2}\right)\sigma_\varepsilon^2 \\
f_t^{(2)} &= \delta + \phi\left(y_t^{(1)} - \delta\right) - \frac{\sigma_\varepsilon^2}{1-\phi}\lambda + \phi\frac{\sigma_\varepsilon^2}{1-\phi}\lambda - \left(\frac{1}{2}\right)\sigma_\varepsilon^2 \\
f_t^{(2)} &= \delta + \phi\left(y_t^{(1)} - \delta\right) - \frac{1-\phi}{1-\phi}\sigma_\varepsilon^2\lambda - \left(\frac{1}{2}\right)\sigma_\varepsilon^2 \\
f_t^{(2)} &= \delta + \phi\left(y_t^{(1)} - \delta\right) - \left(\frac{1}{2} + \lambda\right)\sigma_\varepsilon^2
\end{aligned}$$

the same thing! "Transform to risk-neutral probabilities which alter the drift, then use risk-neutral valuation formula" We can write the whold model this way.

$$\begin{aligned}
f_t^{(2)} &= \delta^* + \phi\left(y_t^{(1)} - \delta^*\right) - (\frac{1}{2})\sigma_\varepsilon^2 \\
f_t^{(3)} &= \delta^* + \phi^2\left(y_t^{(1)} - \delta^*\right) - \left[\frac{1}{2}(1+\phi)^2\right]\sigma_\varepsilon^2 \\
f_t^{(4)} &= \delta^* + \phi^3\left(y_t^{(1)} - \delta^*\right) - \left[\frac{1}{2}\left(1+\phi+\phi^2\right)^2\right]\sigma_\varepsilon^2
\end{aligned}$$

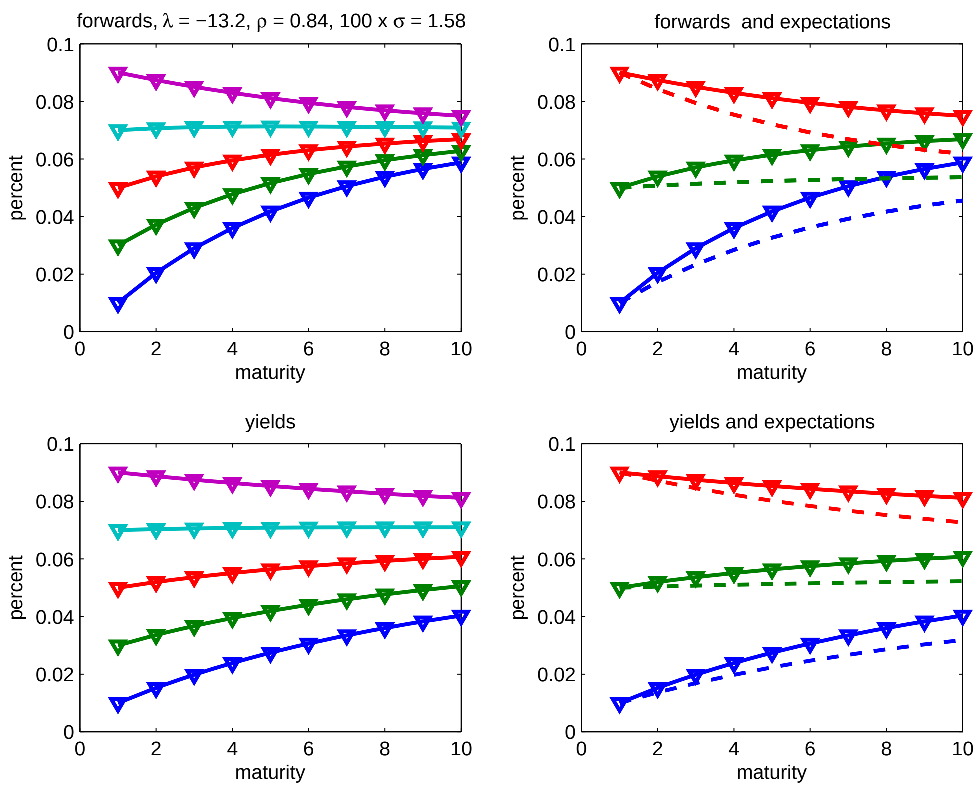
Advantage: prettier formula. Disadvantage: can't describe real dynamics – don't fit the first equation by running a regression of $y_{t+1}^{(1)}$ on $y_t^{(1)}$

- Let's see an example.

1. I chose some parameters to fit the FB zero coupon bond data. I ran a regression of $y_{t+1}^{(1)}$ on $y_t^{(1)}$ to get ρ; I took the variance of errors from that regression to get σ_ε; I took the mean $\delta = E\left(y_t^{(1)}\right)$. Finally, I picked the market price of risk λ to fit the average 5 year forward spread:

$$
\begin{aligned}
f_t^{(5)} &= \delta + \phi^4\left(y_t^{(1)} - \delta\right) - \left[\frac{1}{2}\left(1+\phi+\phi^2+\phi^3\right)^2 + \lambda(1+\phi+\phi^2+\phi^3)\right]\sigma_\varepsilon^2 \\
\lambda &= -\frac{E\left(f_t^{(5)}\right) - \delta}{(1+\phi+\phi^2+\phi^3)\sigma_\varepsilon^2} - \frac{1}{2}\left(1+\phi+\phi^2+\phi^3\right)
\end{aligned}
$$

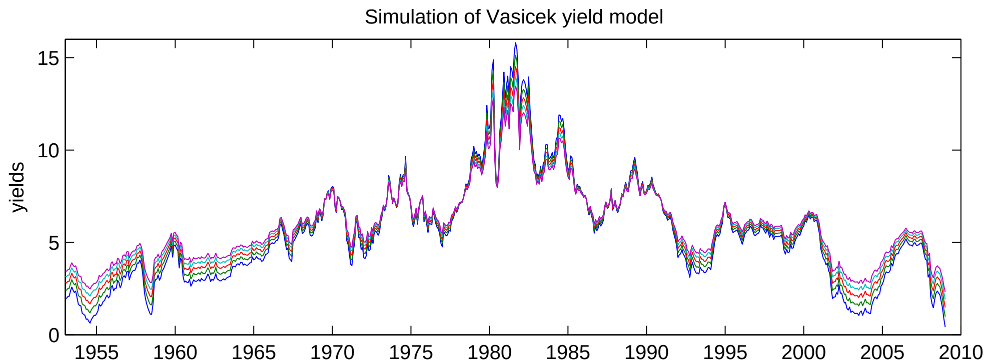
2. I plot $y_t^{(n)}$ for a bunch of $y_t^{(1)}$. The dashed lines in the right hand graph give the expectations hypothesis terms from above, so you can see the distortion from risk aversion λ and the Jensen's inequatlity σ_ε^2 term.

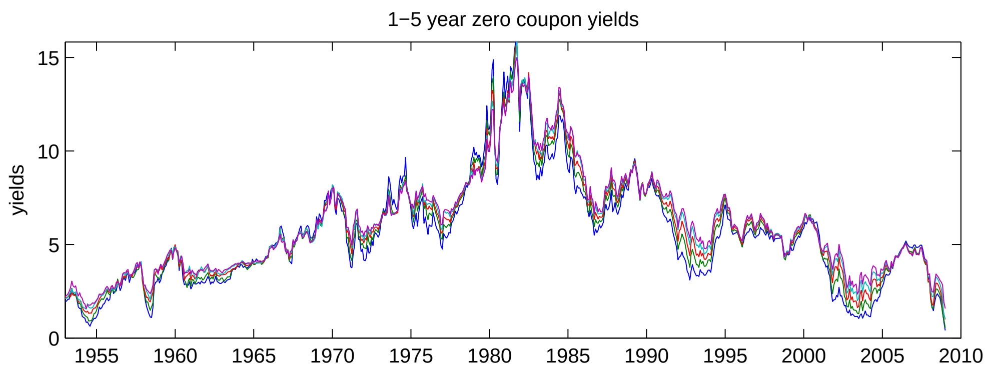


Cool! This captures some basic patterns; yields are upward sloping when lower, downward sloping when higher. The substantial risk premium I estimated to match the average upward slope does introduce a substantial deviation of the model from expectations at the long end.

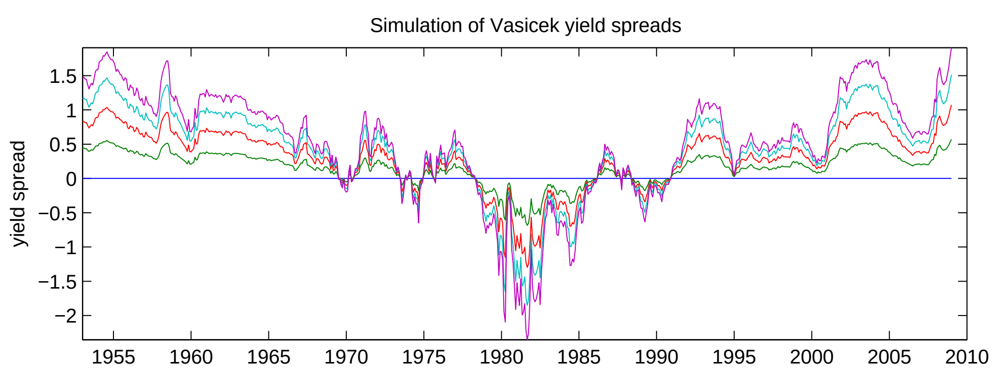
3. Note already: the parameters ϕ, λ can be chosen to match the *cross section* of yields – the shapes of these curves – or the *time series* – the AR(1) coefficient of the short rate and the expected bond returns. These do not necessarily give the same answer, a sign of model misspecification.

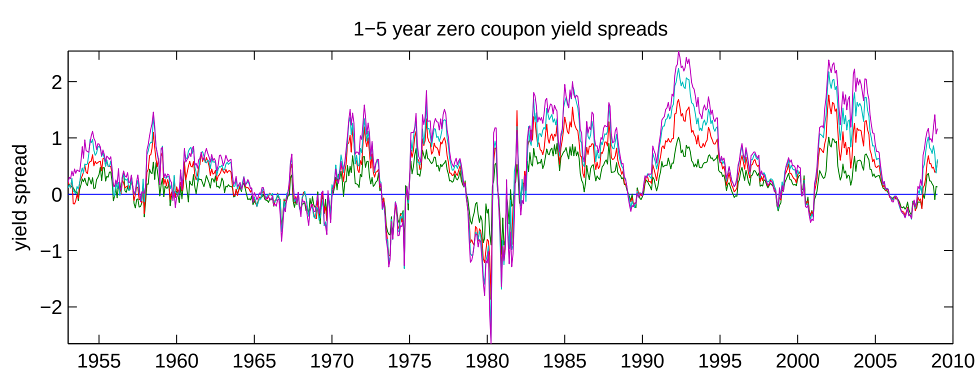
4. Take the history of $y_t^{(1)}$. Find the model-implied $y_t^{(n)}$: compare with data.





5. You can see a decent fit – upwward sloping yields when the interest rate is low. But you can see yields are going up to a constant long-term value, rather than some sort of "local mean. This is clearer if we plot spreads,



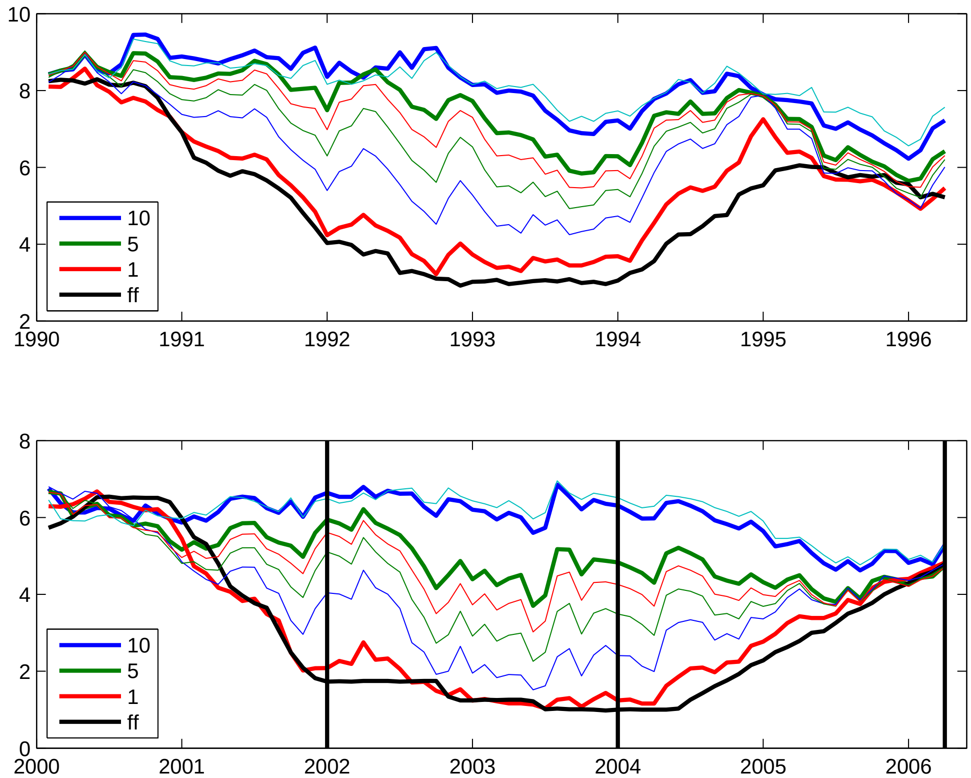


Answer: we need a two-factor model....

Cochrane-Piazzesi 2

1. Our original motivation: How do you forecast rates? How much of $f_t^{(n)}$ is $E_t y_{t+n-1}^{(1)}$?

(We were tasked with, examine the "conundrum," a period in which long term forward rates were anomalous. Task: Understand rates in 2005. Why were long term rates not rising when short term rates rose? Greenspan, "Conundrum" that risk premia seemed to be falling. Us: you must be kidding that you can measure risk premia so well! But let's try. It seems so simple – use all rates to forecast $E_t y_{t+n-1}^{(1)}$, then the residual isk the risk premium)



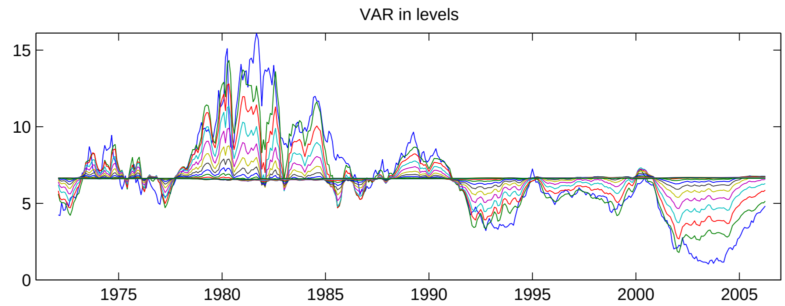
Forward rates in two recessions. The federal funds rate, 1-5, 10 and 15 year forward rates are plotted. Federal funds, 1, 5 and 10 year forwards are emphasized. The vertical lines in the lower panel highlight specific dates that we analyze more closely below.

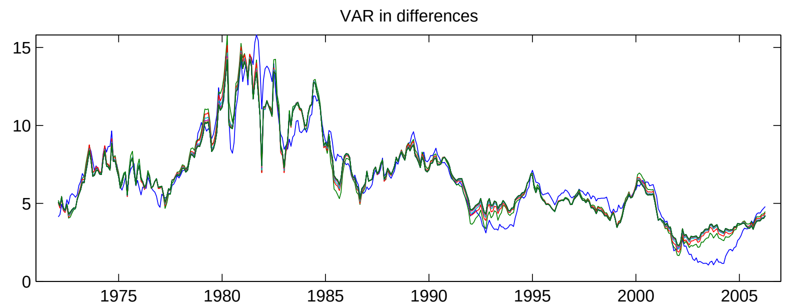
Or is it so easy? VARs leave huge uncertainty about $E_t y_{t+5}^{(1)}$. In addition to wide standard errors, there is a lot of specification uncertainty. Interest rates have a root close to one, and 0.98 vs. 0.99 makes a huge difference. The graph: $E_t y_{t+j}^{(1)}$ based on a VAR using 5 yields at each date. Top:

$$f_{t+1} = \mu + \Phi f_t + \varepsilon_{t+1}.$$

Bottom:

$$\Delta f_{t+1}^{(n)} = \mu_n + \sum_m \phi_{n,m} \left(f_t^{(m)} - y_t^{(1)} \right) + \varepsilon_{t+1}^{(n)}$$





1. *Can we use the structure of a term structure model to learn about the time series from the cross section?*

   (a) Well, $f_t^{(n)} = E^*(y_{t+n}^{(1)})$ but obviously (!) we can't learn about E without saying something about risk premiums. Most basically $p_t u'(c_t) = \sum_{\text{states}} \pi_s \beta u'(c_s) x_s$; probability and marginal utility always enter jointly.

   (b) But we just learned a lot about risk premiums! Can this inform the effort?

2. "Affine model with risk premium." Another objective: learn about longer-run dynamics than the one-year regressions in CP1.

   (a) What's $E_t r_{t+2}^{(n)}$?

   (b) Get back to "implications for prices", feed risk premia through some sort of present value model

3. This is not the usual point of affine model. *Usually the point is just to fit the cross section.* In particular, to price an extra security in terms of given securities. (As Black scholes prices an option in terms of stocks and bonds.) Many authors of affine models lose sight of this.

4. (Digression: affine models are just waking up to the fact that it is very hard to hedge bond options with bonds. This is the heart of the "unspanned stochastic volatility"

message. JC suggestion: well, then you should be hedging one bond option with other bond options, not with treasuries! A suggestion: create something like affine models that span through a basis set of simple bond options. These will reveal volatility much better than bond prices.)

5. (Note: Point to second to last graph of Dec 30 2007. All yield curve models have trouble with data below a year. Plug piazzesi's thesis.)

2. Affine model summary

$$X_t = \left[\begin{array}{cccc} x_t & level_t & slope_t & curve_t \end{array}\right]'.$$
$$X_{t+1} = \mu + \phi X_t + v_{t+1};\; E\left(v_{t+1}v_{t+1}'\right) = V,$$

$$M_{t+1} = \exp\left(-\delta_0 - \delta_1' X_t - \frac{1}{2}\lambda_t' V \lambda_t - \lambda_t' v_{t+1}\right)$$
$$\lambda_t = \lambda_0 + \lambda_1 X_t.$$

(A bit of why; we need $\sigma_t(M)$ to generate $E_t(R^e)/\sigma_t(R^e)$ that varies over time)

$$p_t^{(n)} = \log E_t\left(M_{t+n}\right)$$

Here we go

$$p_t^{(1)} = e^{-\delta_0 - \delta_1' X_t}$$
$$y_t^{(1)} = \delta_0 + \delta_1' X_t$$

This identifies δ_0 and δ_1. (You can also take the short rate as a factor in X, which makes the model simpler. We chose not to)

$$\begin{aligned} p_t^{(2)} &= \log E_t\left(e^{-\delta_0 - \delta_1' X_t - \frac{1}{2}\lambda_t' V \lambda_t - \lambda_t' v_{t+1}} e^{-\delta_0 - \delta_1'(\mu + \phi X_t + v_{t+1})}\right) \\ &= \log E_t\left(e^{-2\delta_0 - \delta_1'\mu - \delta_1'(I+\phi)X_t - \frac{1}{2}\lambda_t' V \lambda_t - (\lambda_t' + \delta_1')v_{t+1}}\right) \\ &= -2\delta_0 - \delta_1'\mu - \delta_1'(I+\phi)X_t + \delta_1' V \delta_1 - 2\delta_1' V \lambda_t \\ &= -2\delta_0 - \delta_1'\mu - \delta_1'(I+\phi)X_t + \frac{1}{2}\delta_1' V \delta_1 - \delta_1' V(\lambda_0 + \lambda_1 X_t) \\ &= (\cdot) - \delta_1'(I+\phi)X_t - \delta_1' V \lambda_1' X_t \end{aligned}$$

$$\begin{aligned} f_t^{(2)} &= p_t^{(1)} - p_t^{(2)} = -\delta_0 - \delta_1' X_t - \left[(\cdot) - \delta_1'(I+\phi)X_t - \delta_1' V \lambda_1 X_t\right] \\ &= (\cdot) + \delta_1'(\phi - V\lambda_1)X_t \end{aligned}$$

....

$$\begin{aligned} f_t^{(n)} &= ..+ \delta_1' \phi^{*(n-1)} X_t \\ \phi^* &\equiv \phi - V\lambda_1 \end{aligned}$$

(Note that since f is a linear function of X and X is an AR(1), this model is homoskedastic for yields and returns. It's interesting to add conditional variance. However, conditional variances in the data are not nearly so simple as the CIR model (variance rises with the level of rates) predicts; variance is also pretty clearly not well spanned by current yields. If you forecast r_{t+1}^2 with past yields and with lagged r_t^2 the latter enter. This is the "unspanned stochastic volatility" observation. In my view, understanding *volatility* is as ripe for plucking as understanding *means* was for Piazzesi and I. Start by forecasting r_{t+1}^2 with all f, What's the factor structure? Then see if r_t^2 enters.... A lot of regressions and plots could get past the mess of the big ML crowd just as our regressions did)

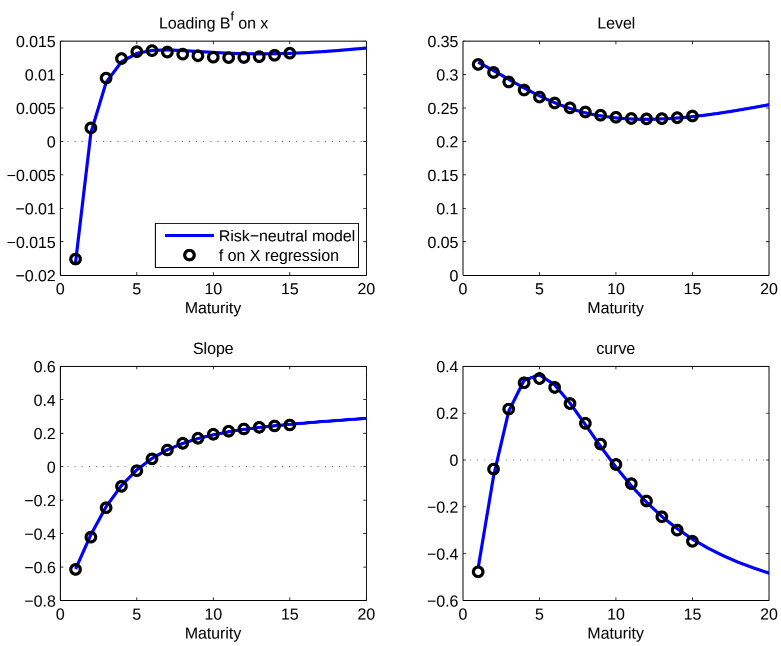
3. We find ϕ^*

From the cross section (nonlinear regression) Note OLS ($f_t^{(n)} = .. + b_n' X_t$) is as good as you can get.

min $\sum_{t=1}^T \sum_{n=1}^N \left\{ f_t^{(n)} - \left[(\cdot) + \phi^{*(n-1)} X_t \right] \right\}^2$ interpret as measurement errors. Note the model predicts *exact* factor structure, so it's easy to "reject."

This uses *no* "time-series information," there is no $(z_{t+1} - aw_t)^2$ (z,w stand for anything) in this.

In our case, the factors X are "observable." However, you can also think of them as "latent." Then, for any guess of ϕ^* you use the first N $f^{(n)}$ to back out the X, then use the remaining $f^{(n)}$ to assess the fit of that ϕ^* (Of course you do this more intelligently, i.e. for any ϕ^* use the first four degrees of freedom of $f^{(n)}$.)



Affine model loadings, B^f in $f^{(n)} = A^f + B^{f\prime} X_t$. The line gives the loadings of the affine model, found by searching over parameters $\delta_0, \delta_1, \mu^*, \phi^*$. The circles give regression coefficients of forward rates on the factors.

4. To ϕ, forecasting? Alas, knowing ϕ^* tells you nothing about ϕ.

Theorem: $\forall \phi^*, \phi, \exists \lambda_1 : \phi^* \equiv \phi - V\lambda_1$

Proof: $\lambda_1 = V^{-1}(\phi - \phi^*)$

Interpretation: The cross section says nothing about the time series without restrictions on market prices of risk. ("A New Perspective on Gaussian DTSMs' Scott Joslin , Kenneth J. Singleton and Haoxiang Zhu")

5. But we know a lot about λ from CP1!

$$\text{Data: } E_t\left(rx_{t+1}^{(n)}\right) = b_n(\gamma' f_t) = b_n x_t$$

$$\text{Model: } E_t(rx_{t+1}) = (\cdot) + cov(rx_{t+1}, v_{t+1}')(\lambda_0 + \lambda_1 X_t)$$

(Why? $M_{t+1} = \exp\left(-\delta_0 - \delta_1' X_t - \frac{1}{2}\lambda_t' V \lambda_t - \lambda_t' v_{t+1}\right)$ $1 = E_t(MR) = E_t\left[e^{-\delta_0 - \delta_1' X_t - \frac{1}{2}\lambda_t' V\lambda_t - \lambda_t' v_{t+1}} e^{r_{t+1}}\right]$ ...)

λ_1 columns: *how does the risk premium vary over time – which state variables tell you that $E_t rx_{t+1}$ is higher at time t than it was at time t-1?*

λ_1Rows: *expected returns depend on covariance of returns with what factor?*

- The lesson of CP1 is that *all variation through time in market prices of risk is carried by x_t*

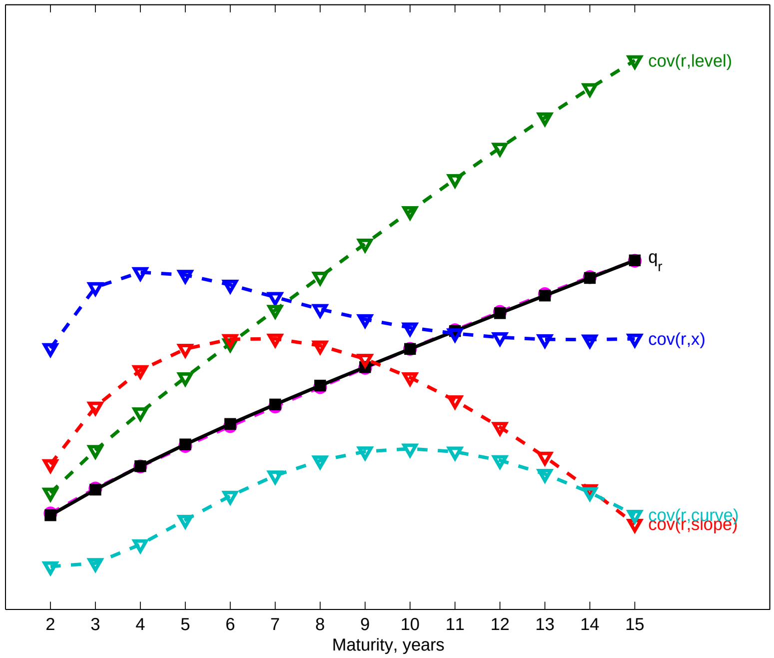
$$\lambda_t = (\lambda_0 + \lambda_1 X_t) = \begin{bmatrix} \lambda_0^{(x)} \\ \lambda_0^{(\text{level})} \\ \lambda_0^{(\text{slope})} \\ \lambda_0^{(\text{curve})} \end{bmatrix} + \begin{bmatrix} \lambda_1^{(x,x)} & 0 & 0 & 0 \\ \lambda_1^{(l,x)} & 0 & 0 & 0 \\ \lambda_1^{(s,x)} & 0 & 0 & 0 \\ \lambda_1^{(c,x)} & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_t \\ \text{level}_t \\ \text{slope}_t \\ \text{curve}_t \end{bmatrix}$$

- So, what covariances to ER line up with?

$$\begin{aligned} E_t\left(rx_{t+1}^{(n)}\right) &= (\cdot) + cov(rx_{t+1}^{(n)}, v_{t+1}')(\lambda_0 + \lambda_1 X_t) \\ b_n x_t &= cov(rx_{t+1}^{(n)}, v_{t+1}^x)\lambda_1^{(x,x)} x_t + cov(rx_{t+1}^{(n)}, v_{t+1}^l)\lambda_1^{(l,x)} x_t + ... \\ b_n &= cov(rx_{t+1}^{(n)}, v_{t+1}^x)\lambda_1^{(x,x)} + cov(rx_{t+1}^{(n)}, v_{t+1}^l)\lambda_1^{(l,x)} + ... \end{aligned}$$

  b_n become the "expected returns" to match. (we're matching not "expected returns" but expected returns conditional on X, i.e. x) The right hand side is a multiple-beta model, with λ as market prices of risk. We want to run a multiple "cross-sectional" regression of b_n on $cov(rx^{(n)}, v^i)$ and see which covariances line up with the expected returns. This is just like Fama - French who look at which betas (h, say) line up with expected returns.

- Visually, you want to know which linear combination of the colored lines add up to the black squares. It's all the green line!



Loading q_r of expected excess returns on the return-forecasting factor x_t, covariance of returns with factor shocks, and fitted values. The fitted value is the OLS cross-sectional regression of q_r on $cov(r, v_{level})$, the dashed line nearly colinear with q_r. The covariance lines are rescaled to fit on the graph.

1. *Market prices of risk correspond entirely to covariance with the level shock.*

2. Details: Black squares: b_n in this notation, q_r in paper notation. These rise nearly literally, which is a finding from CP1. If expected returns on bonds rise (all at the same time, a one factor model), then long maturity expected returns rise more.

3. Colored lines: these are cov(r,v) for various shocks. For example, if all yields rise by the same amount, then $p^{(n)} = ny^{(n)}$ mean prices of long term bonds decline more, and ex-post returns rise more. The green line thus makes sense. The other lines are curved because non-parallel shifts in $y^{(n)}$ imply different patterns of $r^{(n)}$ across maturities.

4. It didn't have to come out this way. Even if there is a single-factor model ($\gamma' f$) it could have come out that all maturities expected returns rise by the same amount (1%) when the factor moves. Then the black squares would have been flat. Or they could have had a curved pattern.

5. *You can estimate λ_1 with a "cross-sectional regression.* In the graph, zeros on all covariances and a single estimated λ on level is plotted as a purple line. Look carefully. (This is the crucial piece of information, and ML will basically focus on this. Note that you have to impose it – we tried running the actual regression and it blows up since there are multiple right hand variables all of which should have zero coefficients.)

6. *Uncertainty about market prices of risk is down to one number!*

$$\lambda_t = \begin{bmatrix} 0 \\ \lambda_{0l} \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ \lambda_{1l} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_t \\ level_t \\ slope_t \\ curve_t \end{bmatrix}$$

6. Summary

a) Find ϕ^* to match the cross section $f_t^{(n)} = \phi^{*(n-1)} X_t$.

b) use a cross sectional regression of b_n on $cov(rx^{(n)}, level)$ to estimate λ_{1l}

c) $\phi \equiv \phi^* + V\lambda_1$

d) $X_{t+1} = \mu + \phi X_t + v_{t+1}$ so we can forecast anything we want.

7. Transition Matrix Estimates

	x	level	slope	curve
Risk-neutral:		ϕ^*		
x	0.35	-0.02	-1.05	8.19
level	0.03	**0.98**	-0.21	-0.22
slope	0.00	-0.02	0.76	0.77
curve	0.00	-0.01	0.02	0.70
Actual:		ϕ		
x	*0.61*	-0.02	-1.05	8.19
level	*-0.09*	**0.98**	-0.21	-0.22
slope	*-0.00*	-0.02	0.76	0.77
curve	*0.00*	-0.01	0.02	0.70

- The risk-neutral ϕ^* from the cross-section = a *lot* of information about the true ϕ!
- All but first column of ϕ are unchanged.

$$\begin{bmatrix} v_{11} & v_{12} & v_{13} & v_{14} \\ v_{21} & v_{22} & v_{23} & v_{24} \\ v_{31} & v_{32} & v_{33} & v_{34} \\ v_{41} & v_{42} & v_{43} & v_{44} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 \\ \lambda_{1l} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} v_{12}\lambda_l & 0 & 0 & 0 \\ v_{22}\lambda_l & 0 & 0 & 0 \\ v_{32}\lambda_l & 0 & 0 & 0 \\ v_{42}\lambda_l & 0 & 0 & 0 \end{bmatrix}$$

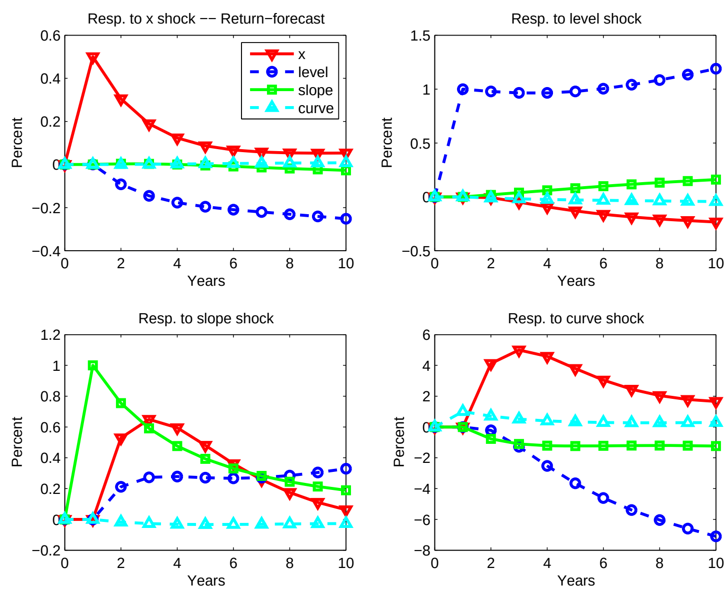
- 0.98 does not change! Near unit-root estimation problems are solved. The root is identified from the cross section. (and λ=0 restrictions)
- True dynamics ϕ : *x is not an AR(1).* Slope$_t$, curve$_t$ $\rightarrow$ x$_{t+1}$. (Top right row.)

1. We can expect *future* risk premium without *current risk premiums.* There is a *term-structure of risk premiums.* If the current yield is above expected future short rates, that could reflect returns that are *expected to be high in the future* even though they are not high now.

$$
\begin{aligned}
y_t^{(n)} &= E_t\left[y_t^{(1)} + y_{t+2}^{(1)} + .. + y_{t+n-1}^{(1)}\right] + rpy_t^{(n)} \\
rpy_t^{(n)} &= \frac{1}{n}\left[E_t\left(rx_{t+1}^{(n)}\right) + E_t\left(rx_{t+2}^{(n-1)}\right) + \ldots + E_t\left(rx_{t+n-1}^{(2)}\right)\right]
\end{aligned}
$$

2. What about $E_t(rx^{(n)})$ for longer horizons? For example, the two year return will combine $E_t rx_{t+1}^{(n)}$ and $E_t E_{t+1} r_{t+2}^{(n-1)}$. What does that look like? A: it is *not* $x_t = \gamma' f_t$. Since x_{t+1} is forecast by l_t, s_t, c_t, then $E_t rx_{t+2}^{(n-1)}$ will depend on l_t, s_t, c_t as well as x_t (Conversely, when we look at shorter horizons, we should not expect x_t to forecast one-month returns.) *x_t is only the return-forecasting factor at an annual horizon.*

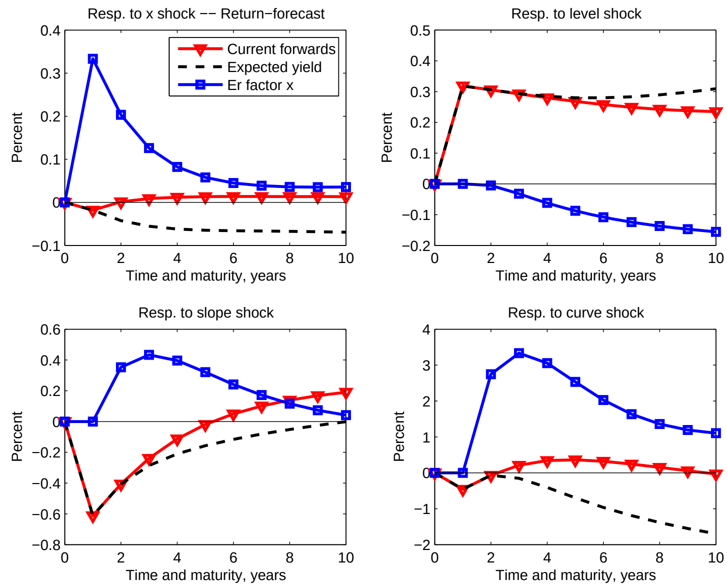
- Plots



1. Top left: x_t does have an AR(1) element. When x_t moves, there isn't much response of the other factors (ok, a bit of level), and a nice decay. If that's all there were, then life would be so simple. But it's not...
2. Bottom left. When slope rises *today*, with no movement in x_t, then x_t rises tomorrow. Thus a high slope at t will imply a high x_{t+1} and a high rx_{t+2}. Again, x is "the" return forecast factor only at one specific horizon.
3. Top right. Level movements seem to have no effect on risk premiums at any horizon. It's interesting that *covariance with level risk generates all risk premiums, but a shift in the level of interest rates does not affect expected returns.* (If you

understand that statement, you really understand the difference between columns and rows of λ, and the difference between $cov(E_t rx_{t+1})$ and $cov(rx_{t+1})$!)

- Plot 2

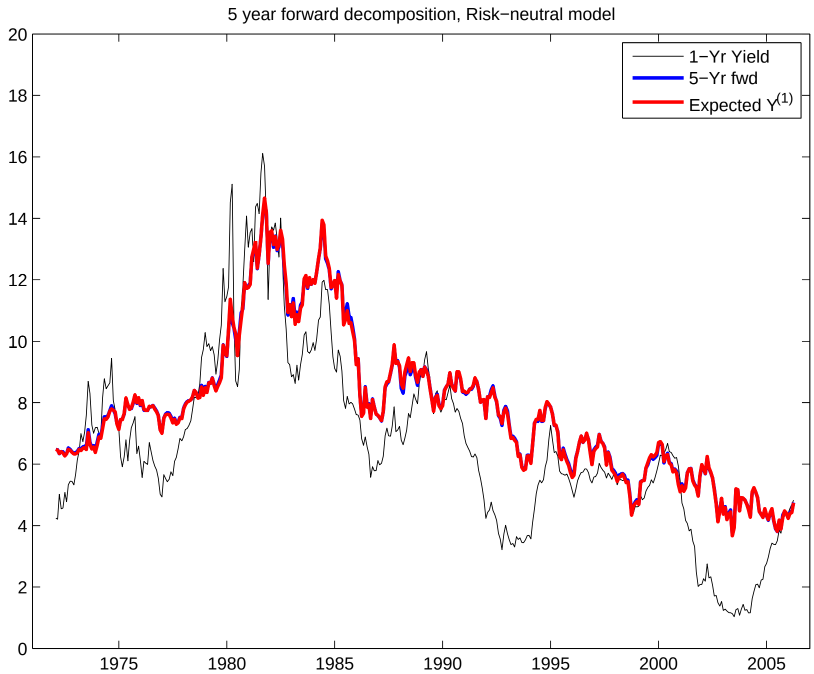


Here, you're looking instead at how a shock at t impacts x_{t+j} (the blue line here is the same as the red line in the last plot – time to change color) along with how that shock changes the forward curve at time t (red) and the path of expected future one year rates $E_t y_{t+n}^{(1)}$ (black dash). The red line plots *current* forwards as a function of *maturity*, and expected x and $y^{(1)}$ as a function of *time.*

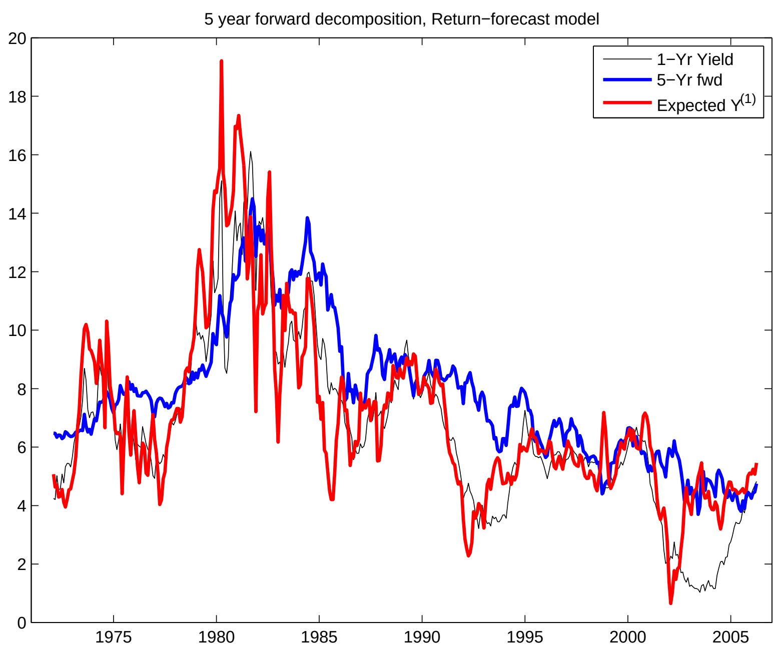
1. Top left: An x shock doesn't do much at all to the forward curve (red). Now you know why it's "nearly unspanned." It does move the risk premium, and hence the expected one year rate line. So a shock to x changes (reflects changes in) expected one year rates, a rise in the risk premium. Since the dynamic pattern is an AR(1), the spread just widens and then sits.
2. A level shock is (red), well, a level shock. It has almost no effect on x, and hence almost no effect on risk premiums. There is a little divergence out in the far number of years.
3. A slope shock moves forwards in a, well, slope-shaped pattern (red). Since it sets off an expectation of *future* risk premiums, there is a divergence between forwards and expected yields in the out years only.

7. Decompositions

Well, what's the answer anyway? When we see $f_t^{(n)}$ move over time, what does that imply about risk premiums and what does it imply about expected future rates?



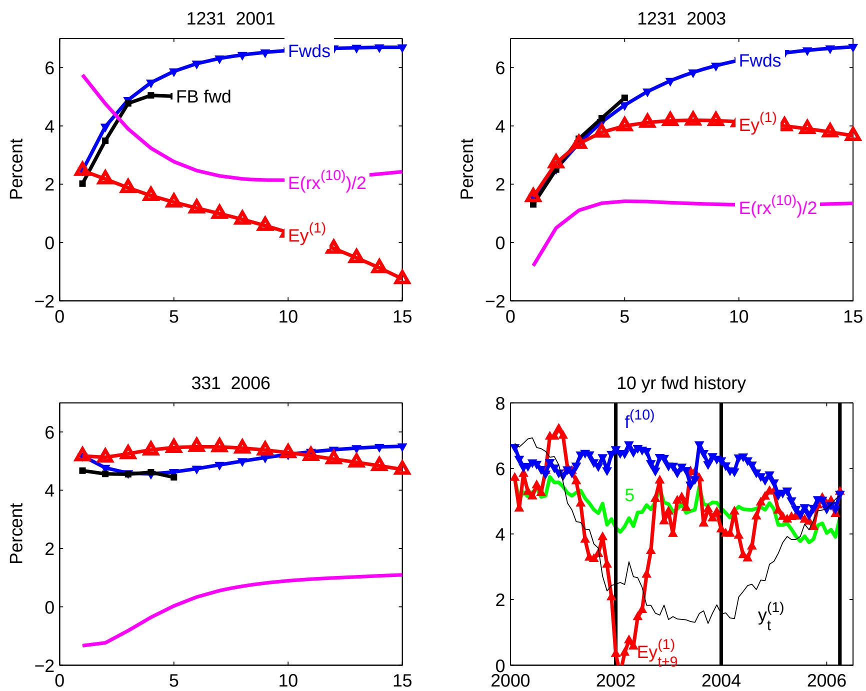
Turn off all frisk premiums The forward rate is exactly the same as the expected 5 year rate. Note the historical pattern, 5 year forwards just glide over the valleys.



When we turn on the risk premium, we get a different picture. In the beginnings of recessions, people seemed to think the interest rate decline might last a long time. Hence, if forward rates are not declining this must reflect a large risk premium. Two effects are offsetting – current short rates decline, the risk premium rises, so long rates do not move much. In the middle of the recession, people get the signal it will end this time – we're not doing Japan 1990 or US 1930-1940 again. Then expected one year rates rise, the risk premium declines.

A last data plea

A last plea to think about data more seriously. There's a lot in this graph, but all I want to say now is to compare the GSW forwards 12 31 2001 and the Fama Bliss forwards. The GSW interpolated data miss the "tent shape" that is so crucial to Fama Bliss. What does the real "forward curve" look like? Time to stop this business of ad-hoc filtering to zeros before starting! This is easy for the estimation stage. An affine model can price coupon bonds, so just pick parameters to make the actual and predicted price of coupon bonds equal. The hard part is, how do you extract *right hand variable* information ($\gamma' f_t$) from the full spectrum of coupon bond prices?



"Term structure model" in general

Ingredients:

1) Write a time series model for the discount factor in discrete or continuous time.

$$
\begin{aligned}
X_{t+1} &= \mu + \phi X_t + \varepsilon_{t+1} \\
M_{t+1} &= F(X_t)
\end{aligned}
$$

X_t may be "observable" or "latent," (which just means we will be able to invert and find them from factors.)

2a) Solve M forward, $M_{t,t+n} = M_{t+1}M_{t+2}...M_{t+n}$. Then

$$P_t^{(n)} = E_t \left[M_{t,t+n} \right]$$

2b) Solve differential / difference equation, i.e. solve $P_t^{(n)}$ "backward" from $P_t^{(0)} = 1$,

$$P_t^{(n)} = E_t \left[M_{t,t+1} P_{t+1}^{(n-1)} \right]$$

Result: $P_t^{(n)} =$ function of state variables that drive M.

This is easiest for logs. Translation from levels to logs: either continuous time or lognormal distributions.

Next steps

- "Unspanned stochastic volatility" and other state variables. (Duffee)
- Liquidity and other spreads (Krishnamurthy, Longstaff, etc.)
- Supply effects (Krishnamurthy and Vissing-Jorgenson; "Understadnding Policy")
- Corporate spreads – basic regressions still open!

FX

- We're going to review the current state of FX work. In my opinion it's a good bit behind the stock and bond work above, which means low-hanging fruit for you to apply the same ideas. (Note Lustig et. al. as a "low hanging fruit" example.")
- The basic idea: Suppose the UK interest rate = 5%, US interest rate = 2%. Should you invest in UK? The naive view: Yes, you'll make 3% more. The traditional view: No, the pound will depreciate 3% (on average) The fact: The pound seems to go *up!* As with D/P the adjustment goes the "wrong way."
- Evidence: Typically country by country time series regressions, following Fama's original in 1981.

$$\$ \text{ Return}^i_{t+1} = a_i + b_i(i^f_t - i^d_t) + \varepsilon_{t+1}$$

  $b \geq 1$. Small R^2.
- This is the basis for the "carry trade," borrow in low interest rate countries and lend in high interest rate countries. (Note the analogy to "ride the yield curve" also called a "carry trade," "borrow in low interest rate *maturities* and lend in high *maturities*. I long for a unifying view!)
- Variations: You can do this for two common choices of *right hand* variables, forward-spot spread or interest differential, and you can do this for two choices of *left hand* variable, exchange rate or excess return. As usual, everything is related by identities. The identies (or arbitrage) say that two ways of getting money to the same place give the same result.

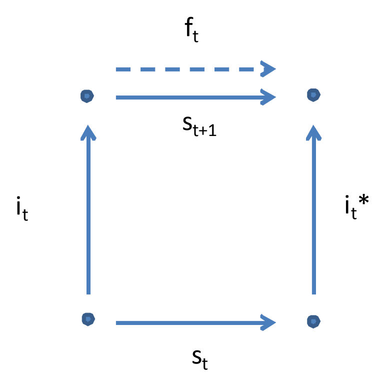


Figure 9:

1. "Covered interest parity" and a right hand variable identity.

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

   i.e. going around the box gets zero

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

   Thus, regressions with $f_t - s_t$ on the right are the same as regressions with $i^*_t - i_t$.

2. Left hand hand variable identity: You can look at exchange rate changes (Δs) or expected returns. By an identity,

$$rx_{t+1} = i_t^* - i_t - \Delta s_{t+1}$$

Thus, if you regress on either right hand variable,

$$b_r + b_s = 0$$

Either exchange rates are predictable or excess returns are predictable.

- Like bonds the first question was "does expectations work?" Is $f_t = E_t(s_{t+1})$? Fama figured out to do this by a) running s_{t+1} on f_t b) much better, running $s_{t+1} - s_t$ on $f_t - s_t$ as a much more powerful test c) once you see that isn't working, it's interesting to note rx_{t+1} on $f_t - s_t$ d) I like expressing it in terms of $i_t^* - i_t$ which is more intuitive to me than forward rates.

- The cp "common factor" investigation has not been done, "covariance with what" is only beginning, the present value relation hasn't been done, etc.

- Fact 1: Expectations is about right in *levels* (433). (Just as the yield curve is pretty flat *on average)*

- Fact 2: We whould see $b_s = +1$ in regressions. In fact: it's negative. Some numbers. From *Asset Pricing*

Table 20.11.

	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

The first row gives the average appreciation of the dollar against the indicated currency, in percent per year. The second row gives the average interest differential—foreign interest rate less domestic interest rate, measured as the forward premium—the 30-day forward rate less the spot exchange rate. The third through fifth rows give the coefficients and R^2 in a regression of exchange rate changes on the interest differential = forward premium,

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

where s = log spot exchange rate, f = forward rate, r^f = foreign interest rate, r^d = domestic interest rate.
Source: Hodrick (1999) and Engel (1996).

An update from Burnside et al.

TABLE 2

UIP Regressions, 1976-2005

	1 Month Regression			3 Month Regression		
	α	β	R^2	α	β	R^2
Belgium†	-0.002	-1.531	0.028	-0.005	-0.625	0.008
	(0.002)	(0.714)		(0.006)	(0.669)	
Canada	-0.003	-3.487	0.045	-0.007	-2.936	0.072
	(0.002)	(0.803)		(0.005)	(0.858)	
France†	0.000	-0.468	0.004	0.001	-0.061	0.000
	(0.002)	(0.589)		(0.005)	(0.504)	
Germany†	-0.005	-0.732	0.005	-0.012	-0.593	0.007
	(0.003)	(0.704)		(0.008)	(0.650)	
Italy†	0.005	-0.660	0.010	0.008	-0.012	0.000
	(0.002)	(0.415)		(0.006)	(0.392)	
Japan*	-0.019	-3.822	0.030	-0.063	-4.482	0.100
	(0.005)	(0.924)		(0.014)	(1.017)	
Netherlands†	-0.009	-2.187	0.029	-0.018	-1.381	0.026
	(0.004)	(1.040)		(0.009)	(0.816)	
Switzerland	-0.008	-1.211	0.012	-0.020	-1.050	0.022
	(0.003)	(0.533)		(0.008)	(0.536)	
USA	-0.003	-1.681	0.017	-0.008	-1.618	0.037
	(0.002)	(0.880)		(0.006)	(0.865)	

* Data for Japan begin 7/78

† Data for Euro legacy currencies ends 12/98

Notes: Regression of [S(t+1)/S(t)-1] on [F(t)/S(t)-1]. Standard errors in parentheses.

Table I
UIP Regressions.

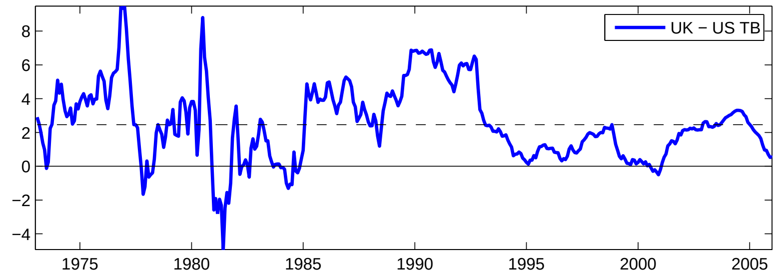
This table reports coefficient estimates from the regression of the time $t+1$ log currency return on the time t log forward premium,

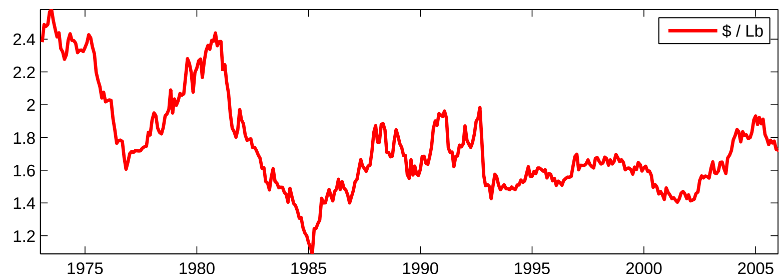
$$s_{i,t+1} - s_{i,t} = a_0 + a_1 \cdot (f_{i,t} - s_{i,t}) + \varepsilon_{i,t+1} \qquad H_0 : a_0 = 0,\, a_1 = 1$$

Currency returns are computed using 21-day rolling windows and span the period from January 1990 to December 2007 (all exchange rates are expressed in terms of dollars per unit of foreign currency). The forward premia are measured using the spread between one-month eurocurrency (LIBOR) rates for loans denominated in U.S. dollars and loans denominated in the foreign currency. The table reports regression coefficients, standard errors (in parentheses), and the χ^2 test statistic for the null hypothesis of UIP (p-values in parentheses). Standard errors in individual regressions are adjusted for serial correlation using a Newey-West covariance matrix with 21 lags (1990-2007: N = 4,673; 1999-2007: N = 2,542). The pooled (panel) regression is run with country-fixed effects; the reported standard errors are robust to within time-period correlation and are adjusted for serial correlation. The pooled regression χ^2 statistic is computed for the null that all country fixed effects are zero and the intercept is equal to one. R^2_{NFE} reports the adjusted R^2 from the panel regression net of the fixed effects (1990-2007: N = 42,057; 1999-2007: N = 22,878). *XS* reports the time series means and standard errors of the regression coefficients from cross-sectional regressions performed for each t. For the cross-sectional regressions R^2 is the mean adjusted R^2 (1990-2007: N = 4,673; 1999-2007: N = 2,324).

	1990-2007				**1999-2007**			
Currency	$\hat{a}_0$	$\hat{a}_1$	R^2_{NFE}	χ^2 test	$\hat{a}_0$	$\hat{a}_1$	R^2_{NFE}	χ^2 test
AUD	-0.0021	-1.6651	0.0087	9.05	-0.0021	-3.6789	0.0240	9.40
	(0.0023)	(1.0487)		(0.01)	(0.0036)	(1.9450)		(0.01)
CAD	0.0006	-0.1754	0.0000	6.87	0.0038	-1.1337	0.0016	5.93
	(0.0010)	(0.5356)		(0.03)	(0.0016)	(2.1465)		(0.05)
CHF	0.0027	-1.2094	0.0060	5.20	0.0100	-4.3559	0.0312	8.64
	(0.0024)	(0.9961)		(0.07)	(0.0041)	(1.8256)		(0.01)
GBP	0.0006	0.0591	-0.0002	1.28	0.0042	-4.3155	0.0401	11.04
	(0.0016)	(0.8981)		(0.53)	(0.0023)	(1.6417)		(0.00)
EUR	0.0021	0.7291	0.0021	3.88	0.0002	-1.6417	0.0053	4.60
	(0.0018)	(1.2675)		(0.14)	(0.0023)	(1.7499)		(0.10)
JPY	0.0059	-1.9686	0.0146	11.83	0.0048	-1.5828	0.0070	5.18
	(0.0025)	(0.8638)		(0.00)	(0.0045)	(1.3785)		(0.08)
NOK	0.0018	0.7030	0.0052	1.88	0.0017	-1.0765	0.0049	5.42
	(0.0017)	(0.6196)		(0.39)	(0.0025)	(1.2000)		(0.07)
NZD	-0.0045	-2.4182	0.0131	15.58	-0.0067	-4.7480	0.0406	16.88
	(0.0034)	(1.1961)		(0.00)	(0.0045)	(1.6800)		(0.00)
SEK	0.0006	0.6554	0.0053	0.53	0.0031	-3.2701	0.0327	10.35
	(0.0017)	(0.5966)		(0.77)	(0.0023)	(1.3594)		(0.01)
Pooled	FE	-0.0986	0.0000	7.1022	FE	-2.7863	0.0190	35.41
		(0.6500)		(0.72)		(1.1179)		(0.00)
XS	0.0007	-0.1622	0.1161	-	0.0017	-0.5154	0.1157	-
	(0.0003)	(0.0853)			(0.0005)	(0.1087)		-

- Can we see it in a graph just like we did for bonds?





1. Higher interest rates *are* associated with stronger exchange. \$/LB goes up when UK rate goes up. There is something to the standard story!
2. If you invest in higher UK rates, you make money until pound weakens, until \$/UK goes down.
3. A higher exchange rate goes on *for many years at a time.* These are 3 month rates. You see the same "sluggish adjustment" as in yields.

- Comments:

1. "Carry trade" by $i^* - i > 0$ and $i^* - i > E(i^* - i)$ are very different! (See hoizontal lines in the graph). The right hand variable is very slow moving.
2. The R^2 is low (monthly data). It's economically large: *All* interest differential (and more?) is expected return, *none* expected depreciation (≤ 1 year). Again, read the regression as "what is the information in the price" not "how do I start my hedge fund?"
3. NB though, many hedge funds do essentially this. As in CP1 they usually trade "enhanced carry," they have some idea of "when to get out." They also form portfolios (something like $\Sigma^{-1}(i - i^*)$)
4. Economics? Low interest rate episodes are recessions, so this has the usual business cycle pattern. When the US risk premium is high, so is the premium for holding currency risk.
5. Wait, that was too quick.
   (a) Why does $cov(u'(c^{us}_{t+1})$,euro$_{t+1}/\$_{t+1}$?) (And the opposite for the Euro investor.) Models have some work to do.

(b) If so, we should not have *separate* regressors $i^* - i$ for each country, there should be a common factor (maybe $\frac{1}{N}\sum_j i^*_j - i_t$) that forecasts *all* currencies. Do like CP1 for currencies? It hasn't been done yet really.

(c) If so, $i^* - i$ should forecast stock and bond returns, as $\gamma' f$ forecasts stock returns. And DP should forecast bond and FX returns. and $\gamma' f$ should forecast FX...Or there should be a reduced factor structure in which a common component of $i^* - i$, $DP, \gamma; f$, forecasts a common component of all returns.

6. a_i is important.

$$\$ \text{ Return}^i_{t+1} = a_i + b_i(i^i_t - i^{us}_t) + \varepsilon^i_{t+1};\ t = 1, 2, ...T$$

There are "country dummies" in the regression. If you leave out a_i you get Turkey or Brazil – perpetually high $i^* - i$ (40%) matched by 40% inflation and devaluation. The fact is "more than usual" interest differential corresponds to a high return. Does this matter? Does slow moving right hand variable mean the a_i estimate biases b_i up? If there is a unit root in inflation, maybe the a_i is meaningless, there is no "usual" differential. Paper topic!

7. The graph suggests "expectations works" in longer run regressions. I'm not aware of papers that do the right hand panel of the Fama-Bliss table well to document this well. (more low hanging fruit.)

8. What about the "present value identity?" Does this link to longer–term regressions? Here is a stab at the question.

$$\begin{aligned}
rx_{t+1} &= i^*_t - i_t - s_{t+1} + s_t \\
s_t &= (i_t - i^*_t) + rx_{t+1} + s_{t+1} \\
s_t &= E_t \sum_{j=0}^{\infty} \left[\left(i_{t+j} - i^*_{t+j} \right) + rx_{t+j+1} \right] + E_t s_{t+j}
\end{aligned}$$

This isn't much help, I'm not getting any discounting and nominal exchange rates can go anywhere. But how about real exchange rates? Let π_t = inflation, so real exchange rate change is

$$s^r_{t+1} - s^r_t = s_{t+1} - s_t - \left(\pi^*_{t+1} - \pi_{t+1} \right).$$

Then,

$$\begin{aligned}
rx_{t+1} &= i^*_t - i_t - \left(s^r_{t+1} - s^r_t + \left(\pi^*_{t+1} - \pi_{t+1} \right) \right) \\
rx_{t+1} &= (i^*_t - i_t) - \left(\pi^*_{t+1} - \pi_{t+1} \right) - \left(s^r_{t+1} - s^r_t \right) \\
rx_{t+1} &= (r^*_t - r_t) - \left(s^r_{t+1} - s^r_t \right)
\end{aligned}$$

where r = ex-post real rates.

$$\begin{aligned}
s^r_t &= (r_t - r^*_t) + rx_{t+1} + s^r_{t+1} \\
s^r_t &= E_t \sum_{j=0}^{\infty} \left[\left(r_{t+j} - r^*_{t+j} \right) + rx_{t+j+1} \right] + E_t s^r_{t+j}
\end{aligned}$$

So, if the expected real exchange rate must approach one in the long run, the current real exchange rate must be matched by real interest differentials or excess returns. I wonder which it is....

9. Big crashes – the "peso problem" was invented precisely for this regression! Many governments do "soft interventions" leading to big left tails and long samples that don't include the left tails.

10. *Strategies that involve small constant gain and occasional big crashes are ubiquitous in hedge funds etc. Dynamic trading can synthesize options. Earth quake insurance. Put options. This is very hard to tell by statistical measures.*

11. The last two comments motivate Jurek, Burnside Eichenbaum Rebelo "crash neutral" currency trades.

Jurek

(April 2009 draft) Good: Updates, it investigates the "peso problem," and the claim that UIP was profitable even if you bought crash insurance against peso problems. It extend pricing questions to currency options. I like to read a recent paper for the latest data, and som reassurance on the "state of the art," so if you do better than this you're doing better than a big literature.

Bad: In many ways it's an example of "how not to write a paper." It's a train of thought and travelogue of experiments. Table IX *is* the paper. If the data kill you in revision, you have to rewrite the paper, not treat it as an "update." Writing papers is the art of throwing things out.

1. Table 1: The standard regressions. Note the much better performance in the later priod. Note the small R2.

2. Table II "carry trade" portfolio returns. At least it's good to look at some portfolios across currencies rather than currency by currency regressions. The basic portfolio one is just long/short depending on the sign of $i - i^*$, ignoring the amount. SPR is proportional to the amount of $i - i^*$. (Why not a real portfolio, $\Sigma^{-1}\mu = \Sigma^{-1}\left(a + b\left(i - i^*\right)\right)$?) Note the high Sharpe ratios. Note that *Portfolio returns can look good with low R^2, like momentum! Portfolios are a way to look at $E(R^e)/\sigma(R^e)$ not $var(bx_t)/var(R^e)$*, and the former is more interesting.

3. "Carry" = average interest differential. This is large – warning! This suggests that there is one data point! *How much of these "profits" are just sitting in one currency through the whole sample = 1 data point? See Figure 3*

4. Figure 1: returns are scaled to same volatility. Note the big crash! Is it over?

5. Crash-neutral construction. P. 12 bottom use out of the money puts, but scale up the portfolio so at the money it has the same sensitivity to exchange rate variations. See Figure 5. Thus, "simultaneously decreasing exposure to depreciations of the high interest rate currency and increasing exposure to its appreications" (p 12)

6. Table VII crash-neutral trades for portfolios through 2007. Note the decline in mean. relative to T II. But there is not so much decline in Sharpe ratios. Jurek: "declines

represent 30-40% of the return to the unhedged strategy." This is where people got the idea that the carry trade worked even if you buy peso-problem protection. (Note a similar story. Before October 1987, out of the money equity put options were cheap. After that date, they rose a lot and have stayed high ever since!)

7. Figure 8 the crash-neutral trades in the crash. Why did even crash protected decline? I thought they were crash protected? The price of put option protection shot up in the last few months.

8. Table 8 The quarterly protection seems to be doing better, echoing my story about Figure 8. But are you allowed to search ex-post over the protection horizon?

9. Conclusion: a muddle. Is anything left of crash-protected returns?

Lustig, Roussanov, and Verdelhan

- Read Up through p. 15 only (April 2009).

- Big picture. Rather than run regressions, sort in to portfolios and look at means; then eigenvalue decompose the portfolio covariance matrix. "Do like Fama-French" (and a bit "like Cochrane Piazzesi") for FX rather than regressions.

- It makes the connection between regression and Fama-French procedures. It starts some extended musing for me on "how should we characterize $E_t R_{t+1}$ and $cov_t(R_{t+1}, f_{t+1})$?

- Table 1. Average returns in portfolios sorted on the basis of $f_t - s_t = i^* - i$ across countries.

  1. Where are the standard errors?
  2. Please put t subscripts ($\Delta s_{t+1}, rx_{t+1}$ but $f - s_t$)!
  3. The Point: high $f - s$ correspond to positive appreciation, not negative, and hence to positive excess returns.
  4. It's nice to put bid-ask spreads in.
  5. High-low portfolios: we really want to know whether $E(R)$ is *different* across portfolios, and this is a simple way to do it. (More thoughts coming on this below.)

- Table 2: Principal components of protfolios.

  1. No surprise "$ moves" is the first component.
  2. The second component is "slope" and third is "curve".

- Table 3: FF style portfolios HML and EW market. This is a cross sectional regression of $ER = \beta\lambda$ with prespecified factors.

  1. Betas panel II: Market betas are all about 1 and hml betas rise. We knew this from the covariance matrix

2. It's not a tautology. *The high interest rate countries will all fall or rise together.* It didn't have to happen
3. Factor risk premiums. Of course, we account for cross sectional variation with a big λ on HML. You see the pattern that ER is higher in the high $f-s$ portfolios, and the betas are high there as well.

- Table 5. This is better. It uses the first two principal components as factors. d is the first component, c is the second.
- *UIP risk premium is earned for covariance with the "slope" factor.*
- Is this different from Cochrane Piazzesi who find covariance with the "level" factor? No, because CP are examining *bonds,* and here they're examining *portfolios.* CP model

$$E_t rx_{t+1}^{(n)} = b_n x_t$$

  Thus, when x_t is positive, CP will put long bonds in portfolio 1 and short bonds in portoflio 5. However, when x_t is negative, CP will put *short* bonds in portoflio 1 and *long* bonds in portfolio 5. Covariance of *bond* returns with a level shock to *yields* is not the same thing as covariance of *portfolio* returns (which change composition) with a level shock to those *portfolio returns.* I suspect if we do the LRV procedure on CP bond data we would get exactly LRV's results. And vice versa? A good problem set/ paper question!

- **Preview on portfolios and time series regressions:**

1. Now, rthink about what they're doing – *relation between portfolio sorts and regressions.*
2. Asset pricing is in the end about $E_t(R^{ei}_{t+1}) = cov_t\left(R^e_{t+1}, f_{t+1}\right)\lambda_t$. $R^e_{t+1} = a + bx_t + \varepsilon_{t+1}$ tells you about $E(R^e_{t+1}|x_t)$, and forming portfolios based on x_t also tells you $E(R_{t+1}|x_t)$. It's really a non-parametric forecasting regression with a rather inefficient kernel!
3. Similarly, the FF procedure amounts to $cov_t\left(R^e_{t+1}, f_{t+1}\right)\lambda_t = cov\left(R^e_{t+1}, f_{t+1}|x_t\right)\lambda$. *FF portfolios are a brilliant way to reduce a time-varying conditional problem to an unconditional problem.*
4. Of course, we've done this many times before, i.e. $0 = E_t(m_{t+1}R^e_{t+1}); 0 = E(m_{t+1}R^e_{t+1} \otimes z_t) = E(m_{t+1}\left[R^e_{t+1} \otimes z_t\right])$
5. Basically, but *the details really matter.* Is time series or cross sectional variation more important? Is the *value* of x_t or the *portfolio rank* more important?

Closing FX/predictability thoughts

- A common pattern across all assets:

1. Dividend yield forecasts stock returns

2. Long yield - short yield forecasts long-short bond returns
3. Foreign - domestic yield forecasts foreign - domestic returns
4. (Cross section – B/M forecasts returns. In this case, both returns and earnings)

- More facts in common with stocks, bonds

1. "Follow yield," "All price variation = Expected returns"
2. "Missing adjustment" (short run, i.e. ≤ 1 year)
3. Expected returns are high in "Bad times", P/D is low, R^f is low relative to R^{f*}, and R^f is low relative to y^{long}.
4. Is there a "single recession related factor?" R^f, term spread, bond forecast factor also forecast stock excess returns. Do they forecast fx? does $i - i^*$ forecast stock returns? Etc.

- More puzzles in international finance

1. News—flow and price correlations

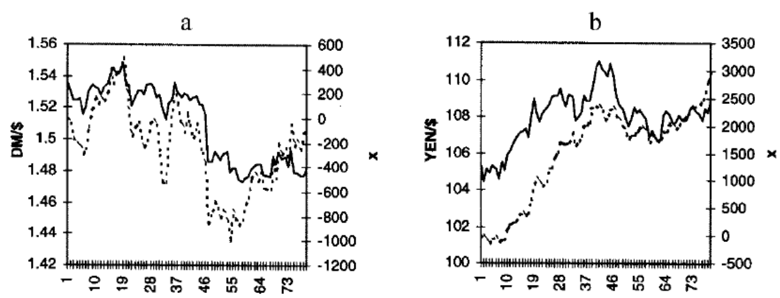


FIG. 1.—Four months of exchange rates (solid) and cumulative order flow (dashed), May 1–August 31, 1996: *a*, deutsche mark/dollar; *b*, yen/dollar.

2. Fact: *Net* order flow *is associated with* price changes. ("order flow" not "trades")
3. Don't jump to: *Any* order *causes* price changes. (Brandt and Kavaiecz coming up)

Brandt, Cochrane, Santaclara

Puzzles

in international finance pretty much define the field.

1. UIP (Just studied): $i^* - i$ seems to imply appreciation, not depreciation, at least for a while, and corresponding profits.

2. The volatility of exchange rates; correlation of real/nominal exchange rates.

$$\sigma\left(\ln\frac{e_{t+1}}{e_t}\right) = 15\%$$

This is not matched 1-1 by inflation, so *real exchange rates vary a lot.*

   (a) (Mussa) Real relative prices across borders change when exchange rates change, suggesting "sticky" nominal prices.

   (b) When countries move from floating to fixed, relative prices across countries (sausage in Munich/Pizza in rome) become more stable, and relative prices of tradeables/nontradeables (Pizza in Rome/Oil in Rome) become more stable. Stock prices are a relative price too – installed vs. uninstalled capital. This is more dramatic.

   (c) Put bluntly, why did this happen? (See pound plot)

Figure 10:

3. Savings = investment and poor risk sharing across countries. One is about allocation across time, the other about allocation across states.

   (a) Permanent income logic means that temporary high income should be exported and then returned later. Also good news about future output (China opening) should lead to a consumption boom and huge imports of capital. Instead, China finances investment from domestic savings and exported the whole time. (This is an open economy facing world interest rates. If the whole world sees a boom, interest rates rise.) ("Feldstein-Horioka puzzle")

   (b) Complete markets, Pareto-Optimum means

$$\max E\left[\lambda_1 \sum \beta^t u(c_{1t}) + \lambda_2 \sum \beta^t u(c_{2t})\right] \text{ s.t. } c_{1t} + c_{2t} = C_t$$

$$\begin{aligned}
\lambda_1 \beta^t u'(c_{1t}) &= \lambda_2 \beta^t u'(c_{2t}) \\
\beta \frac{u'(c_{1t})}{u'(c_{1t-1})} &= \beta \frac{u'(c_{2t})}{u'(c_{2t-1})} \\
\left(\frac{c_{1t}}{c_{1t-1}}\right)^{-\gamma} &= \left(\frac{c_{2t}}{c_{2t-1}}\right)^{-\gamma}
\end{aligned}$$

In fact, $cor(\Delta c_i, \Delta c_j)$ is small (numbers follow). Worse, consumption correlations are less than output correlations.

(c) In both cases, I find it hilarous that as the world starts to look more like our models, people think this is a problem. "Global imbalances" is the buzzword for the idea that we need to *slow down* trade surpluses and deficits. Mortgage backed securities did a great job of sharing risk around the world.

4. "Home bias" in portfolios. US people hold mostly US equities, UK people hold more UK equities and so forth. This is only a puzzle however relative to a world capm model, in which the investor has no job, cares equally about consumption from all countries, etc. There are lots of easy reasons it's optimal for portfolios to focus on your own country, as in my Earth vs. Mars example.

5. Currency crashes, panics, etc. (Much silliniess; "contagion" "capital flight." etc.)

6. International is RIPE for work, as witnessed by Lustig et al and this paper. Simple models are making big progress.

This paper preview:

- We can connect domestic and foreign discount factors by a simple change of units.

$$\begin{aligned} M^f_{t+1} &= M^d_{t+1}\frac{S_{t+1}}{S_t} \\ \frac{utils_{t+1}}{Euro_{t+1}} &= \frac{utils_{t+1}}{\$_{t+1}}\frac{\$_{t+1}/E_{t+1}}{\$_t/E_t} \\ m^f_{t+1} &= m^d_{t+1} + \ln\frac{S_{t+1}}{S_t} \end{aligned}$$

  Equivalently,

$$\begin{aligned} 1 &= E(M^d_{t+1}R_{t+1}) \\ &= E(M^d_{t+1}\frac{S_{t+1}}{S_t}\frac{S_t}{S_{t+1}}R_{t+1}) \\ &= E(M^f_{t+1}R^f_{t+1}) \end{aligned}$$

  where R^f = any return (domestic or foreign) expressed in foreign currency. This is cool! Exchange rates let you see mrs, directly, ex post! Well, they let you see *differences* in mrs, but that's something.

- Important – distinguish the "discount factor for returns expressed in domestic currency" from "discount factor that only prices domestic returns." The latter makes no sense unless markets are segmented somehow. *All* we're doing here is saying that we can find a discount factor that predicts a *given* set of returns converted to Euros from a discount factor that prices *the same* returns expressed in dollars. (Which is, when you see it, rather trivial.) We are *not* constructing a discount factor that prices Euro stocks from a discount factor that (only) prices dollar stocks.

- You can do the same thing with real vs. nominal discount factors. Just mulitply and divide by π.
- Now,

$$\begin{aligned}
\ln \frac{S_{t+1}}{S_t} &= m_{t+1}^f - m_{t+1}^d \\
\sigma^2\left(s_{t+1} - s_t\right) &= \sigma^2\left(m_{t+1}^f\right) + \sigma^2\left(m_{t+1}^d\right) - 2\rho\sigma(m)\sigma(m)
\end{aligned}$$

- What does it take to fit the facts?

$$\sigma\left(\Delta s_{t+1}\right) = 15\%$$

1. Asset pricing, "risk sharing is better than you think." From asset markets, $\frac{E(R^e)}{\sigma(R^e)} <\approx \sigma(m)$ we need at least $\sigma(m) = 50\%$. Since $\sigma(m)$ is much bigger than $\sigma(e)$ we need a lot of positive corrleation.

$$\begin{aligned}
0.15^2 &= 2 \times 0.50^2 - 2 \times \rho \times 0.50^2 \\
0.0225 &= 0.50\,(1 - \rho) \\
0.045 &= (1 - \rho) \\
\rho &= 0.955
\end{aligned}$$

"Risk sharing is better than you think" meaning marginal utility growth is very correlated across countires.

2. Asset pricing, "or exchange rates are too smooth." Imposing $\rho = 0$,

$$\sigma(\Delta s_{t+1}) = \sqrt{2}\sigma(m) = 1.41 \times 0.5 = 0.71$$

We should see 70% variation in exchange rates!

3. Consumption. If we use consumption data, Δc, small risk aversion γ, and $\rho = 0$ as suggested by the data, no matter what we do with ρ on the right hand side $\sigma(m)$ is just not enough to add up to the observed $\sigma(\Delta s)$. $\sigma(m) = \gamma\sigma(\Delta c)$.

$$\begin{aligned}
0.15^2 &= 2 \times \gamma \times \sigma^2(\Delta c) - 2\rho \times \gamma \times \sigma^2(\Delta c) \\
&= 2 \times \gamma \times \sigma^2(\Delta c) \times (1 - \rho) \\
\frac{0.15}{\sqrt{2}} &= 0.106 = \gamma\sigma(\Delta c)(1 - \rho)
\end{aligned}$$

This isn't as bad as the equity premium (that's the whole point), where we needed $\gamma\sigma(\Delta c) = 0.5$. It's still not easy. With $\rho = 0$ and $\sigma(\Delta c) = 0.01$ we need $\gamma = 10$. $\sigma(\Delta c) = 2\%$ gets us down to $\gamma = 5$. And surely there's some positive correlation. In sum, you can see that models (especially models with risk sharing) have trouble producing *enough* exchage rate movement.

The paper details:

- We compute a "risk sharing index"

$$1-\frac{\sigma^2(\ln m^f-\ln m^d)}{\sigma^2(\ln m^f)+\sigma^2(m^d)}=1-\frac{\sigma^2(\ln e_{t+1}/e_t)}{\sigma^2(\ln m^f)+\sigma^2(m^d)}$$

  Why? $\ln m^f=2\ln m^d$ also violates risk sharing, but *correlation* is one.

- Procedure: Just like Hansen-Jagannathan. Find the minimum variance discount factors m to price both domestic and foreign assets, expressed in dollars, and vice versa. We use continuous time so we can do logs vs. levels ("E(log) = log(E) theorem" is true in continuous time, with $1/2\sigma^2$ terms)

- Continuous time

$$\begin{aligned}\Lambda^d &= e\Lambda^f\\ d\ln\Lambda^d &= d\ln e+d\ln\Lambda^f\end{aligned}$$

$$\begin{aligned}\frac{dS}{S} &= (r+\mu)\,dt-\sigma dB\\ \frac{d\Lambda}{\Lambda} &= -rdt-\mu'\Sigma^{-1}\sigma dB\\ \frac{d\ln\Lambda}{\Lambda} &= -\left[r+\frac{1}{2}\mu'\Sigma^{-1}\mu\right]dt-\mu'\Sigma^{-1}\sigma dB\\ \sigma^2\left(\frac{d\ln\Lambda}{\Lambda}\right) &= \mu'\Sigma^{-1}\mu\end{aligned}$$

  so the regular calculation works in logs in continuous time. **Table 2,3** gives the basic calculation.

Strongly recommended reading:

- The introduction on transport costs (Earth vs. Mars) and incomplete markets, and "reconciliation" p. 692 i.e. apples and oranges.

- *Earth vs. Mars*: Just read 673. Better in print. Suppose there are complete financial assets and communication but no goods may flow. If mars gets a good shock, mars stock goes up. The exchange rate must go down. m^f and m^d must be uncorrelated in the end. Knowing this, there is no advantage to Mars stock in the first place, so your portfolio should be completely home biased. *Complete financial assets do not imply "perfect risk sharing" nor constant exchange rates, nor absence of home bias.* Money, capital can't "flow". International is about transport costs, not markets. (Example, crashes "investors pulled capital out." They can't They can sell to locals at a cheap price; trade claim on US govt for claim to factory, but it needs a ship to remove capital.)

- *Incomplete markets.* Now $m^i=m^*+\varepsilon^i$ again.

1. Reminder: Risk sharing in incomplete markets:

$$\begin{aligned} p &= E(m^i x) = E\left[proj\left(m^i|x\right)x\right] = E(m^* x), \\ m^i &= m^* + \varepsilon^i \end{aligned}$$

we should "use asset markets to share as much as possible."

2. It is *not* true that $m^d = m^f + \Delta s$ for arbitrary m pairs. It *is* true that for any m^d that prices assets expressed in dollars, we can construct *an* m^f that prices assets expressed in Euros. Again, this is just a change of units. But that discount factor may not equal foreign consumption growth. It *is* true that $m^{*d} = m^{*f} + \Delta s$. Thus *minimum-variance discount factors in the payoff space do obey the identity.* In this sense, what we are learning is "do transport costs mean that we are not able to use asset markets to share as many risks as possible?"

- The paper also asks if incomplete markets are quantitatively plausible. $\sigma(m)$ rises as m becomes less correlated, so incomplete markets makes the equity premium puzzle that much worse (as in the "correlation puzzle" refinement of Hansen-Jagannathan that lower correlation of m with r implies a higher $\sigma(m)$ bound).

- How much international finance uses vs. does not use complete markets?? Be very careful here. if

$$e = m^d - m^f$$

and then

$$e = \gamma \Delta c^d_{t+1} - \gamma \Delta c^f_{t+1}$$

you *are* assuming complete markets. If

$$e = proj(\Delta c|assets) - proj(\Delta c^f|assets)$$

then you're not.

Misconceptions:

1. *Each investor is allowed to invest in all assets* – the HJ, minimum variance discount factor for all assets as viewed by each investor. Our equation only applies as a change of units. These are NOT the minimum variance discount factor for domestic assets and the minimum discount factor for foreign assets. Why not? You can compute such quantities, but they are not connected by the exchange rate.

   The paper does *nothing* about "what if there are asset market frictions so you can't trade each other's assets?" We allow markets to be incomplete, but once D can buy them, so can F. Would it be interesting to give the countries fundamentally different spaces, or (the same thing) prices that are different by shadow costs or transactions costs as well as by exchange rates? Yes, but we didn't do it.

2. The "discount factor" is *not* the "optimal portfolio" the "market portfolio", or a portfolio anyone actually holds. The *correlation* of discount factors means *nothing* about correlation of portfolios. For example, if our income shocks are the same, we can have very correlated discount factors, very correlated m^* but utterly different portfolios.

3. The correlation of stock markets, which underlie the usual "benefits of international diversification," is not really in the calculation at all. (Technically, it is reflected in the Σ part of $\mu'\Sigma^{-1}\mu$, but higher asset correlation *does not imply* higher discount factor correlation.) Again, we allow trade in both assets by both investors.

4. Converesly, the discount factor is not a portfolio that anyone holds, so highly correlated discount factors do not mean portfolios are highly correlated. Thus, also, home bias does not contradict our calculations.