Bond Risk Premia

John H. Cochrane and Monika Piazzesi
University of Chicago GSB

January 16, 2007

Bottom line

- Forecast 1 year treasury bond returns, over 1 year rate:

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

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

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

- Tent-shaped factor is correlated with slope but is not slope. Improvement comes because it tells you when to bail out – when rates will rise in an upward-slope environment

Background – Expectations and Fama-Bliss.

- 1. Expectations hypothesis. Expected returns are constant over time.

$$rx_{t+1}^{(n)} = a_n + 0 \times x_t + \varepsilon_{t+1}^{(n)}$$

  2. Fama-Bliss.

     (a) Expectations Hypothesis: $\beta = 0$. Instead, $\beta \approx 1$.If the n year forward is 1% higher than the spot, then the n-year bond will earn 1% more on average

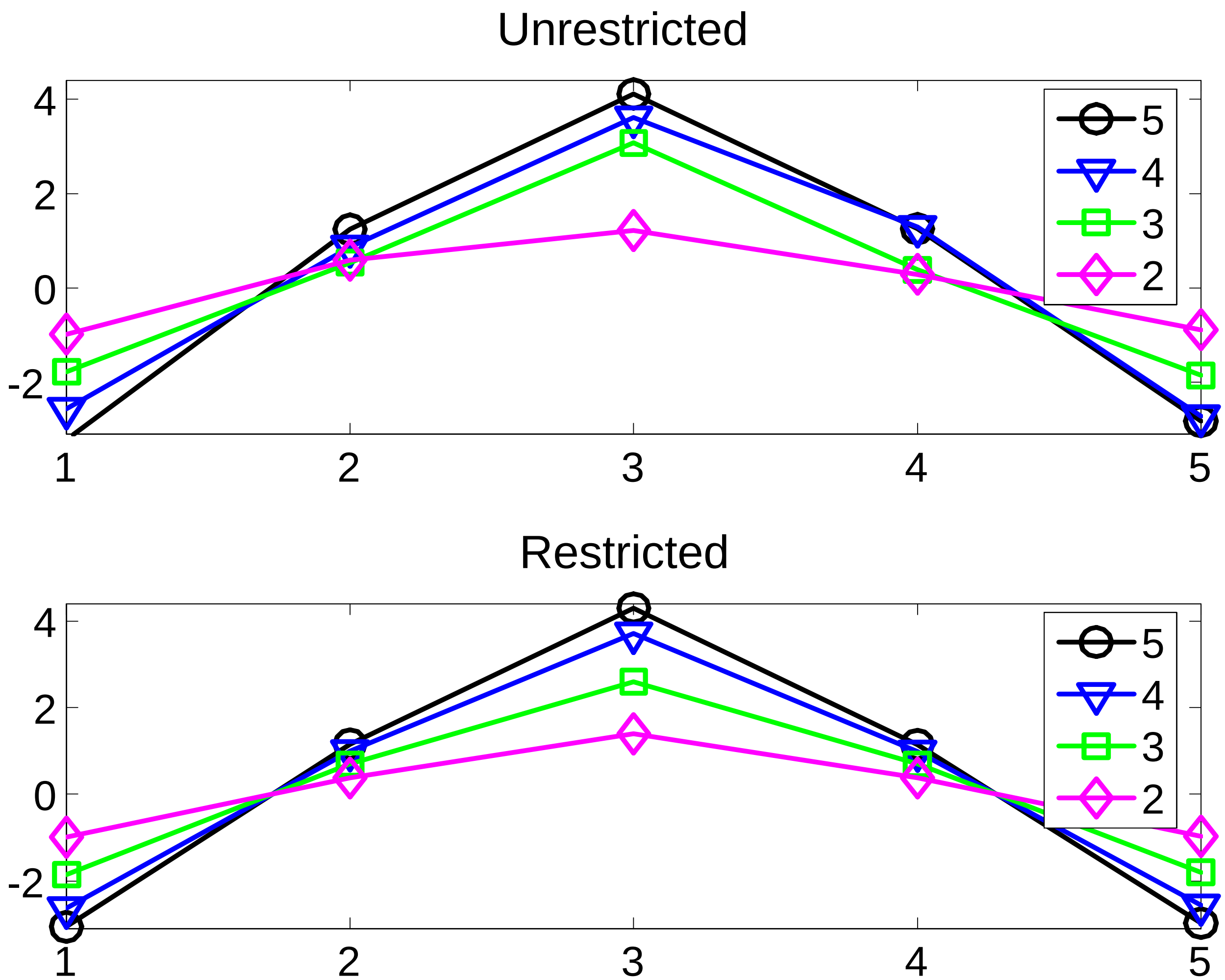
     (b) $R^2 \approx 0.15$; Held up well in 1990s

Table 2. Fama-Bliss excess return regressions

$$rx_{t+1}^{(n)} = \alpha_n + \beta_n \left(f_t^{(n)} - y_t^{(1)} \right) + \varepsilon_{t+1}^{(n)}.$$

Maturity n	β	$s.e.$	R^2	$\chi^2(1)$	p-val
2	0.99	(0.33)	0.16	18.4	$\langle 0.00 \rangle$
3	1.35	(0.41)	0.17	19.2	$\langle 0.00 \rangle$
4	1.61	(0.48)	0.18	16.4	$\langle 0.00 \rangle$
5	1.27	(0.64)	0.09	5.7	$\langle 0.02 \rangle$

Cochrane and Piazzesi



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

- Regressions of bond excess returns on *all* forward rates, not just matched $f - y$
- The *same* linear combination of forward rates forecasts all maturities' returns.

A single factor for expected bond returns

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

- Two step estimation; first γ then b.

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.

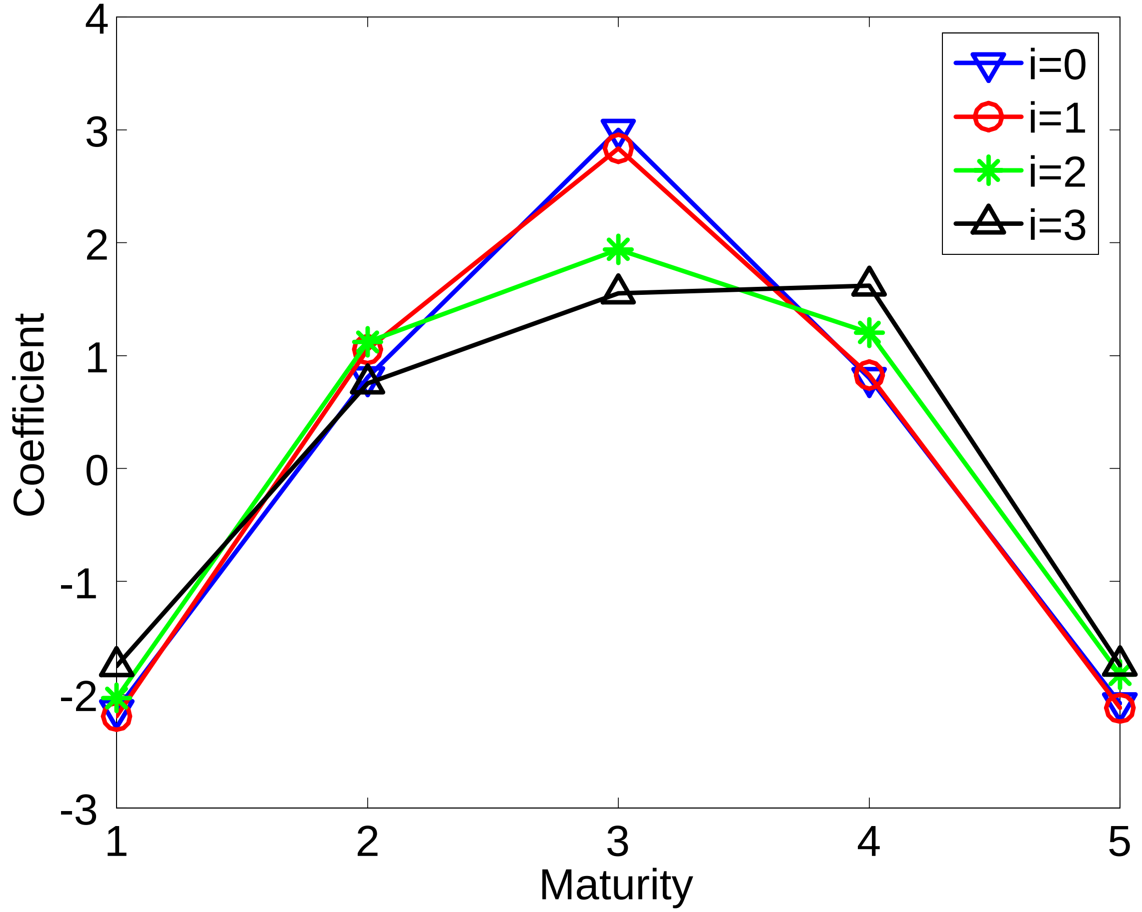
•Restricted model $b_n\gamma$ almost perfectly matches unrestricted coefficients. (well below 1σ)

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

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

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

More lags



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

•More lags are significant, same pattern. Suggests moving averages

$$
\begin{aligned}
rx_{t+1}^{(n)} &= a_n + b_n \gamma' (\alpha_0 f_t + \alpha_1 f_{t-1} + .. + \alpha_k f_{t-k}) + \varepsilon_{t+1}^{(n)} \\
&= a_n + b_n \left[\alpha_0 \left(\gamma' f_t \right) + \alpha_1 \left(\gamma' f_{t-1} \right) + .. + \alpha_k \left(\gamma' f_{t-k} \right) \right] + \varepsilon_{t+1}^{(n)}
\end{aligned}
$$

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

•Interpretation: Yields should be Markov, so a small transitory measurement error. $f_{t-1/12}$ is informative about the true f_t, so it enters with the same pattern.

Stock Return Forecasts

Table 3. Forecasts of excess stock returns (VWNYSE)

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

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

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

Interest Rate Forecasts

Table A4. Forecasting short rate changes $y_{t+1}^{(1)} - y_t^{(1)}$

	$f_t^{(2)} - y_t^{(1)}$	$y_t^{(1)}$	$f_t^{(2)}$	$f_t^{(3)}$	$f_t^{(4)}$	$f_t^{(5)}$	R^2	χ^2/p
FB:	0.01						0.00	0.0
s.e.	(0.26)							$\langle 0.98 \rangle$
CP:		−0.02	0.41	−1.21	−0.29	0.89	0.19	82.7
s.e.		(0.17)	(0.40)	(0.29)	(0.22)	(0.17)		$\langle 0.00 \rangle$

$$E_t\left(rx_{t+1}^{(2)}\right) = -E_t\left(y_{t+1}^{(1)} - y_t^{(1)}\right) + \left(f_t^{(2)} - y_t^{(1)}\right). \quad (1)$$

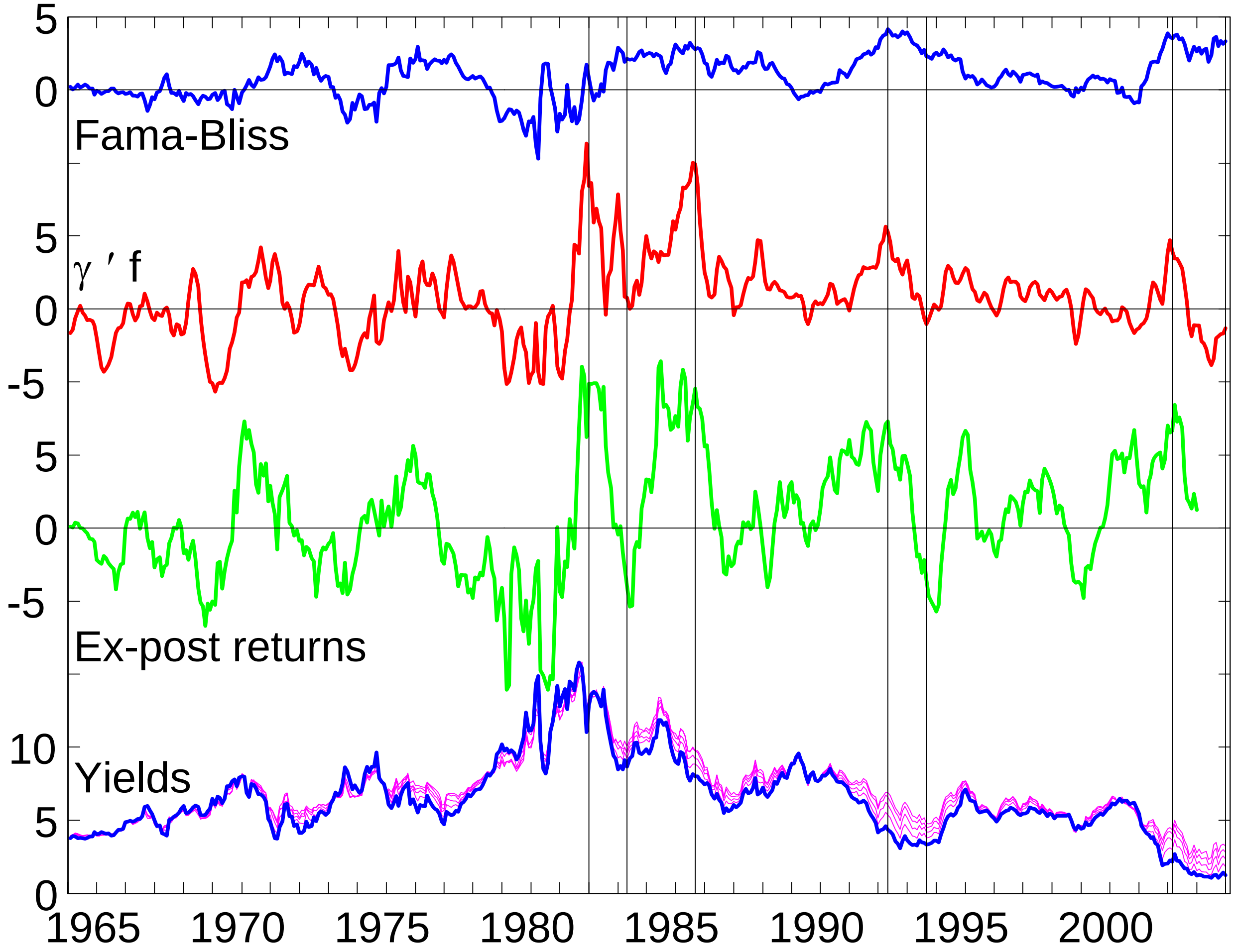
•Expectations: $E_t\left(rx_{t+1}^{(2)}\right) =$ constant

•Fama-Bliss: $E_t\left(y_{t+1}^{(1)} - y_t^{(1)}\right) =$constant

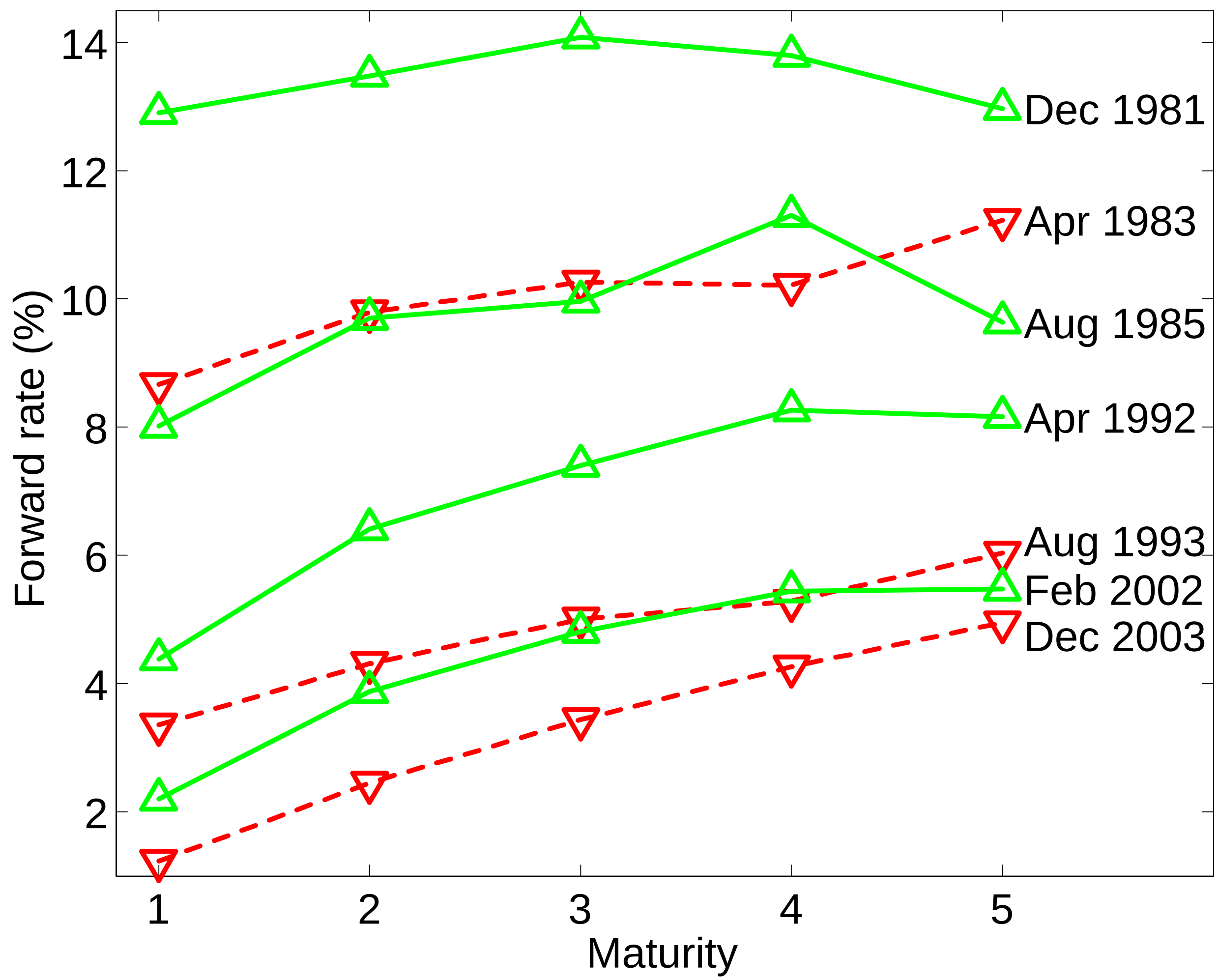
•CP: $E_t\left(rx_{t+1}^{(2)}\right)$ varies more than $\left(f_t^{(2)} - y_t^{(1)}\right)$: must be able to forecast short rates (capital gains on long bonds, not just "ride yields")

•Pattern of coefficients is exactly γ. $\gamma' f$ forecasts short rate changes.

History

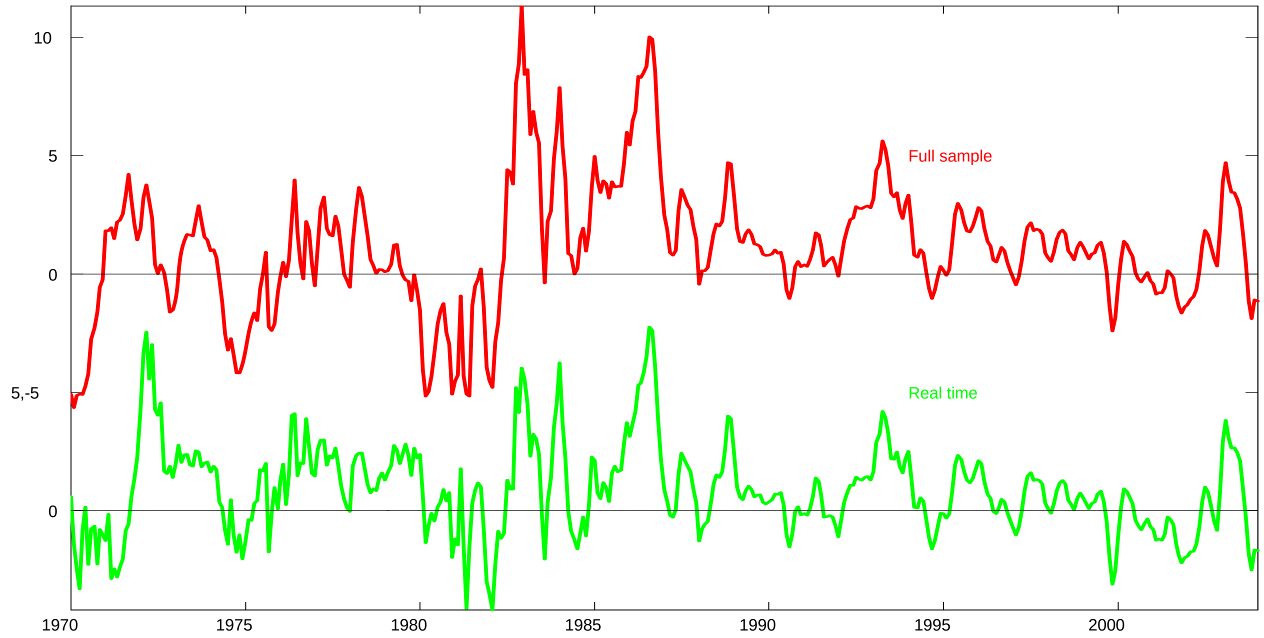


- $\gamma' f$ and slope are correlated. Both show a rising yield curve but no rate rise
- $\gamma' f$ improvement in many episodes. $\gamma' f$ says get out in 1984, 1987, 1994, 2004(?)



Forward rate (%)
14
12
10
8
6
4
2
1
2
3
4
5
Maturity
Dec 1981
Apr 1983
Aug 1985
Apr 1992
Aug 1993
Feb 2002
Dec 2003

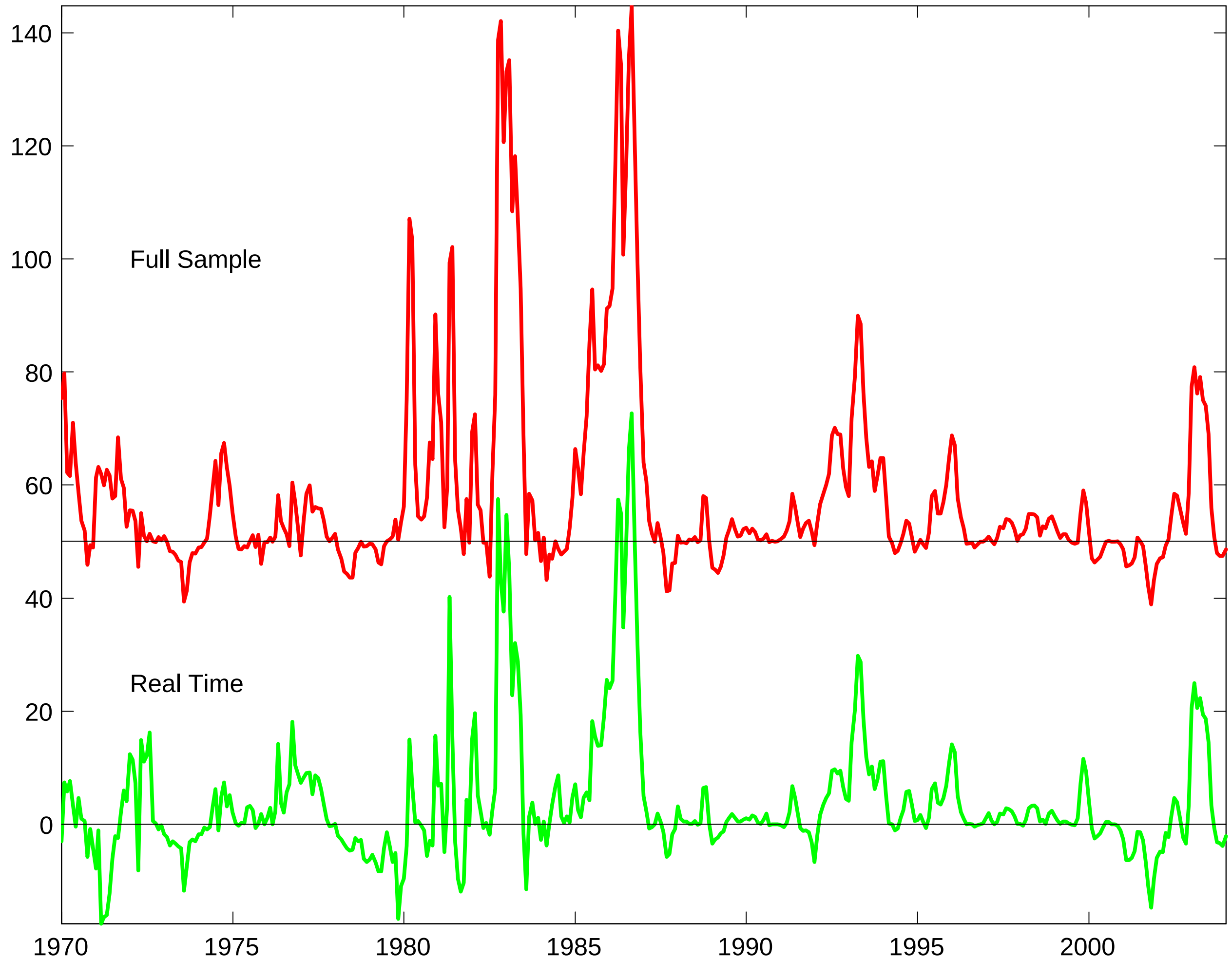
Real time and trading rules



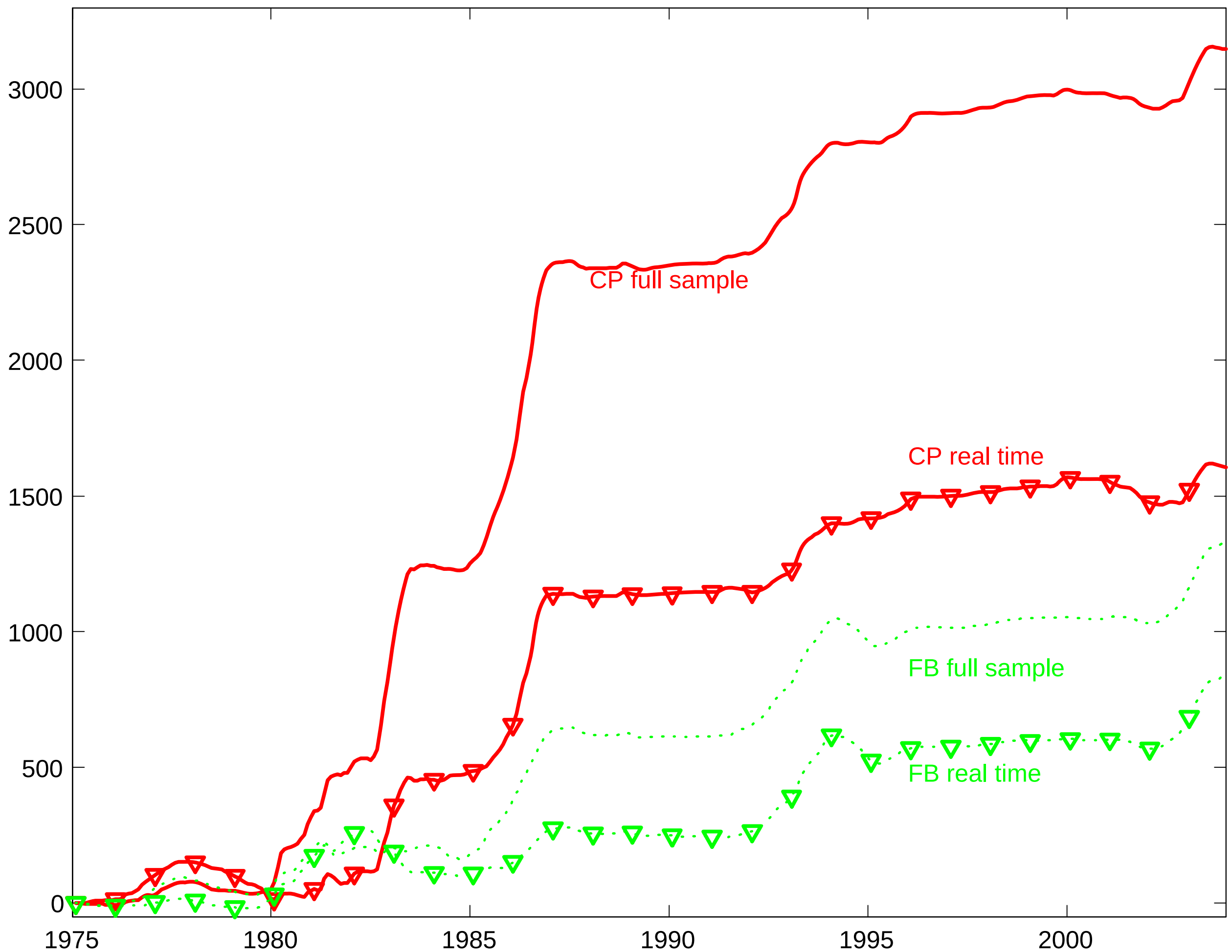
Regression forecasts $\hat{\gamma}^\top f_t$. "Real-time" re-estimates the regression at each t from 1965 to t.

Trading rule

$$\overline{rx}_{t+1} \times E_t(\overline{rx}_{t+1}) = \overline{rx}_{t+1} \times \left[\gamma^\top \left(\alpha_0 f_t + \alpha_1 f_{t-1} + a_2 f_{t-2} \right) \right].$$

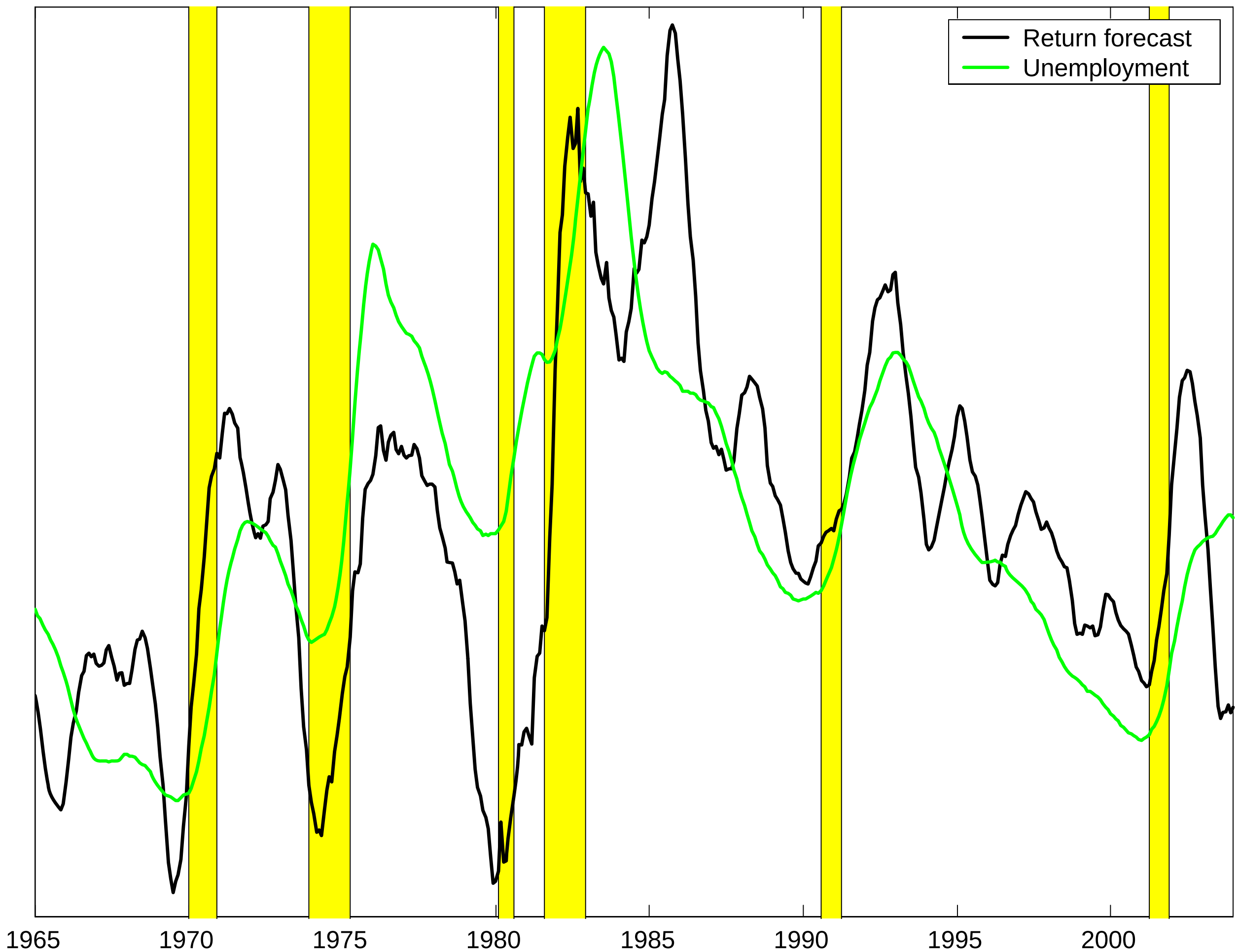


"Trading rule" profits, using full-sample and real-time estimates of the return-forecasting factor.



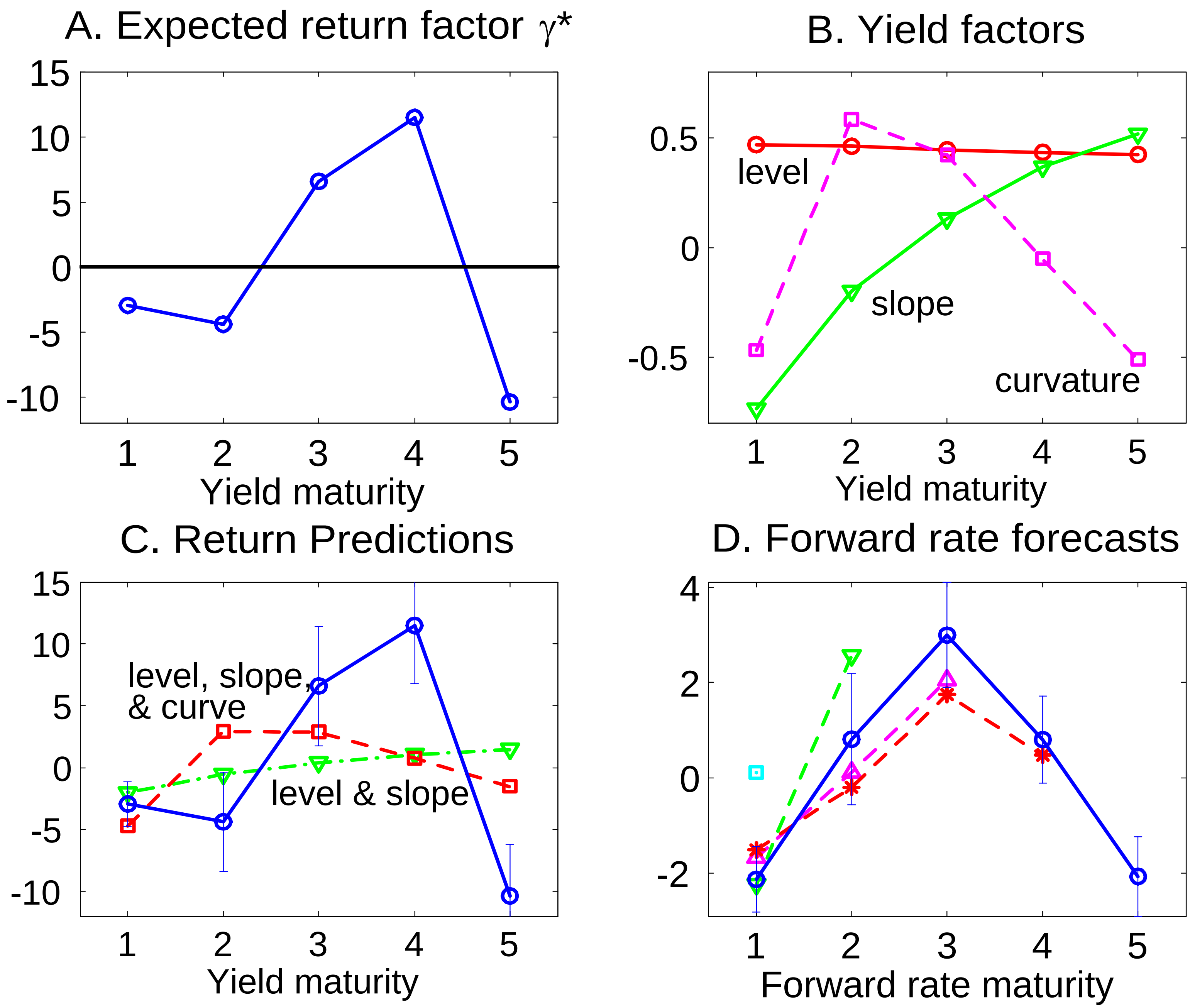
Cumulative trading rule profits; cumulative value of $\overline{rx}_{t+1} \times E_t(\overline{rx}_{t+1})$.

Macro



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

Relation to factor models (why is this news?)



Panel A: Yields are a linear combination of forwards. $\gamma' f = \gamma^{*\prime} y$; which full set is a matter of taste.
γ^* $\approx$ Slope plus 4-5 spread.

Panel B: $\gamma' f$ has *nothing* to do with slope (symmetry: γ' linear $= 0$) and curvature (curved at the long, not short end).

Panel C: You can't approximate $\gamma' f$ well with level, slope, and curvature factors. (See table below)

•Moral 1 Term structure models need L, S, C to get Δy_{t+1} *and* $\gamma' f$ to get $E_t rx_{t+1}$
Adding $\gamma' f$ will not help much to hit yields (pricing errors) but it will help to get *transition dynamics* right (i.e. expected returns, yield *differences*)
Moral 2. You can't *first* reduce to L, S, C, *then* examine $E_t rx_{t+1}$ →Reason #1 this was missed.

Panel D: Stable as we add forward rates.

- Is $\gamma' f$ forecast *significantly* better than forecasts using yield curve factors or simple spreads?

  See Table 4 of paper,

regression of average returns on 3 forward factors

	const	curve	slope	level	R2
b	-0.02			0.17	0.06
t	-0.98			1.23	
b	-0.04		1.39	0.17	0.20
t	-2.44		3.15	1.59	
b	-0.04	2.74	1.39	0.17	0.30
t	-2.65	5.72	3.92	1.97	

regression of average returns on x

	const	x	R2
b	0.00	0.47	0.36
t	0.08	7.99	

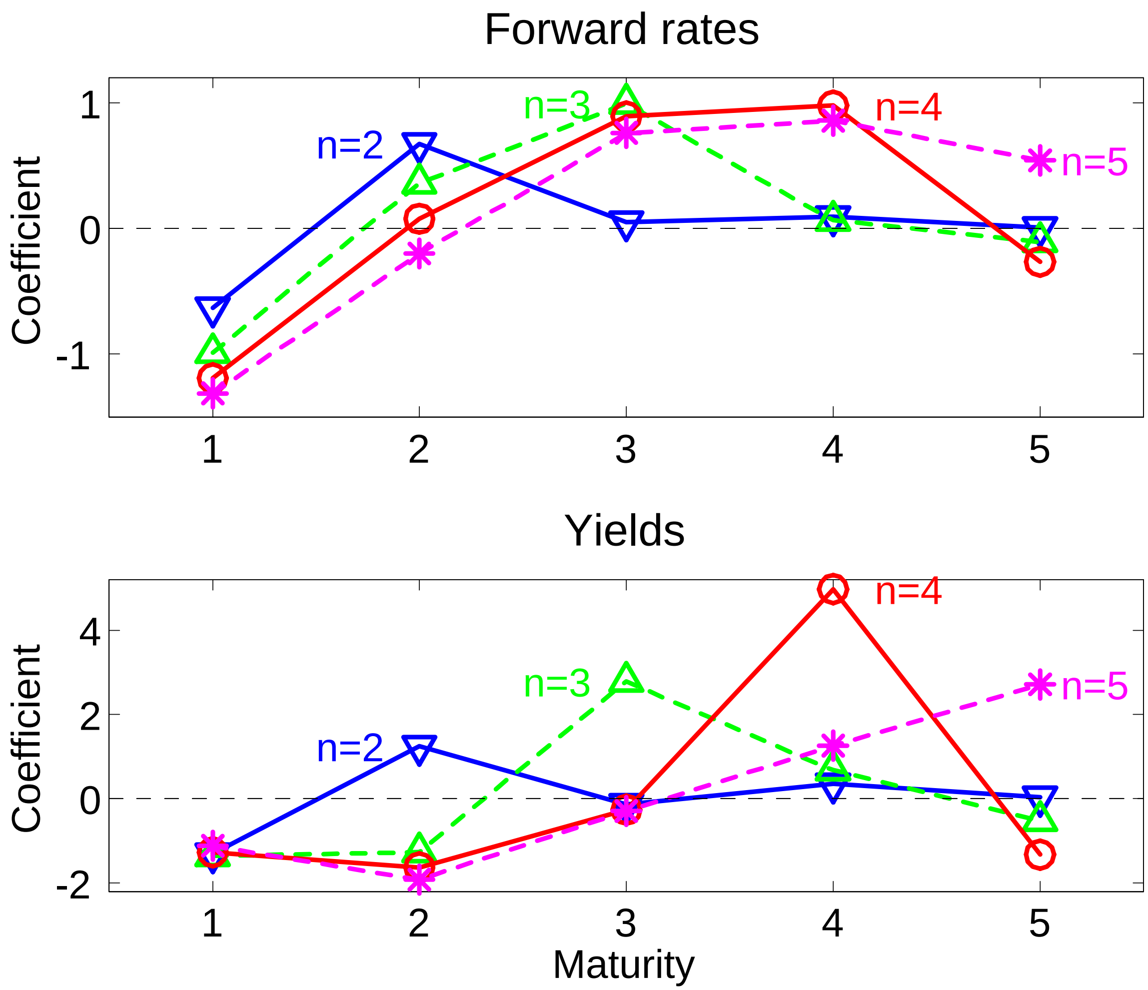
regression of average returns on x and 3 forward factors

	const	x	curve	slope	level	R2
b	0.00	0.47	-0.06	-0.05	0.00	0.36
t	0.02	4.95	-0.10	-0.10	0.04	

- How would you integrate this in to an affine model?

  Paper: shows you how to *construct* market prices of risk so that an affine model *exactly* matches this (any) return regression. (Also see "Decomposing the Yield Curve")

Why is this news? 2. Lags matter, and montly models



Coefficients $rx_{t+1}^{(n)} = a_n + b_n' f_t$ implied by $y_{t+1/12} = \phi y_t + \varepsilon_{t+12}$, then ϕ^{12}.

•*Small* R^2. *No* single-factor. *Pattern* looks like measurement error. Nothing there! Why? Lags matter – yields are not a VAR(1). (VAR(12) works here)

•Moral: Must look at annual returns directly (or fit ARMA(1,1), or deal with measurement error) to see annual horizon return forecasts.

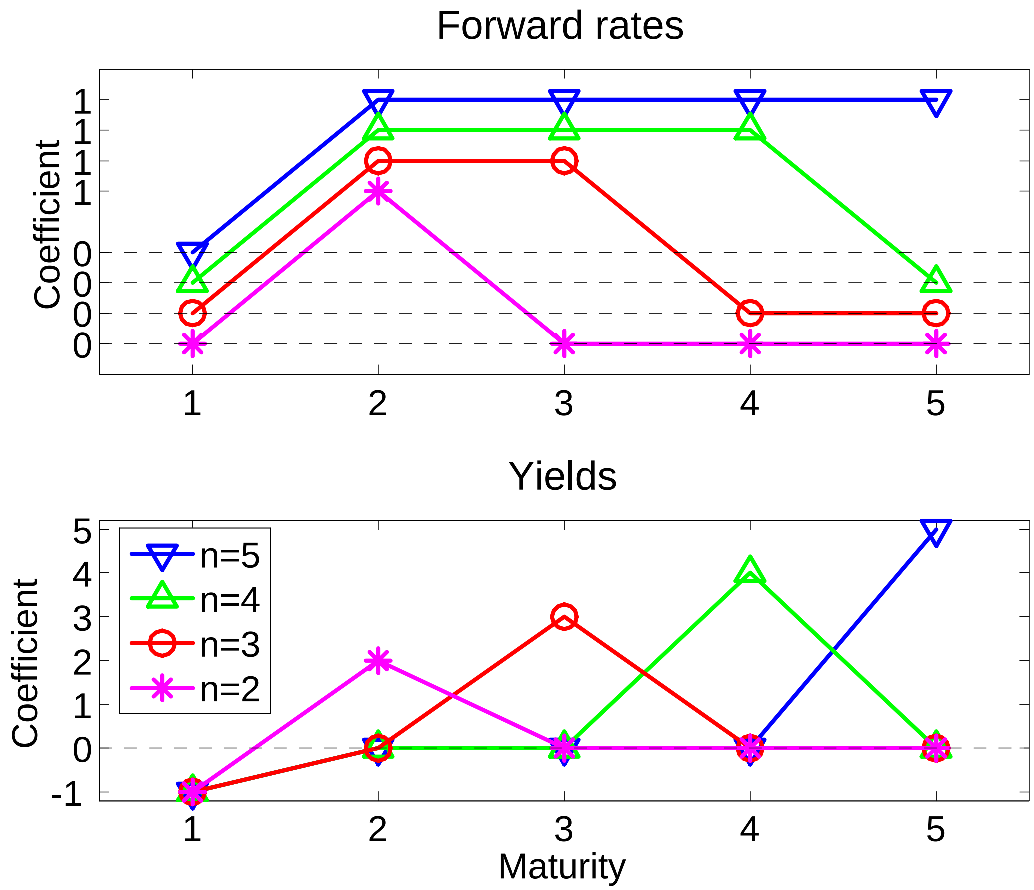
Is this all measurement error?

Danger: if p_t is measured too high, then $r_{t+1} = p_{t+1} - p_t$ will be too low, and a high p_t will seem to forecast a low r_{t+1}. Is this all there is to our results?

1. No. Lags also forecast, with no common price.

2. No. $\gamma' f$ also forecasts stock returns with no common price.

3. Measurement error gives a pattern that the n period yield at t forecasts the n period bond return. It does not give a common factor (m yield helps to forecast n bond return) Measurement error cannot produce our central finding, the single factor.


Forward rates
Coefficient
1
1
1
1
0
0
0
0
1
2
3
4
5
Yields
Coefficient
5
4
3
2
1
0
-1
n=5
n=4
n=3
n=2
1
2
3
4
5
Maturity

Testing the single factor model

- Paper Table 6: the single factor model is dramatically rejected! (*Joint* not individual coefficients)

- $$\begin{bmatrix} rx_{t+1}^{(2)} \\ \vdots \\ rx_{t+1}^{(5)} \end{bmatrix} = (\alpha) + \begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \beta & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix} \begin{bmatrix} y_t^{(1)} \\ f_t^{(2)} \\ \vdots \\ f_t^{(5)} \end{bmatrix} + \begin{bmatrix} \varepsilon_{t+1}^{(2)} \\ \vdots \\ \varepsilon_{t+1}^{(5)} \end{bmatrix} \text{(1)}$$

  vs

  $$\begin{bmatrix} rx_{t+1}^{(2)} \\ \vdots \\ rx_{t+1}^{(5)} \end{bmatrix} = (\alpha) + \begin{bmatrix} b_2 \\ \vdots \\ b_5 \end{bmatrix} \begin{bmatrix} \gamma_1 & \gamma_2 & \cdots & \gamma_5 \end{bmatrix} \begin{bmatrix} y_t^{(1)} \\ f_t^{(2)} \\ \vdots \\ f_t^{(5)} \end{bmatrix} + \begin{bmatrix} \varepsilon_{t+1}^{(2)} \\ \vdots \\ \varepsilon_{t+1}^{(5)} \end{bmatrix} \text{(2)}$$

- $$\begin{aligned} rx_{t+1} &= \beta f_t + \varepsilon_{t+1} \quad \text{(1)} \\ rx_{t+1} &= b\gamma' f_t + \varepsilon_{t+1} \quad \text{(2)} \end{aligned}$$

- GMM

$$E_T(f_t \otimes \varepsilon_{t+1}) = 0 \quad (1)$$

$$\begin{aligned} E_T\left[f_t \otimes (1_4' \varepsilon_{t+1})\right] &= 0 \rightarrow \gamma (2) \\ E_T\left[\left(\gamma' f_t\right) \otimes \varepsilon_{t+1}\right] &= 0 \rightarrow b \end{aligned}$$

- Results: If we do optimal GMM on (2), weird parameters. And *huge* rejections.

- Why??? Restriction:

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

should not be predictable. ($E(\varepsilon \otimes f_t) = 0$) Since

$$\overline{rx}_{t+1} = \gamma' f_t + \bar{\varepsilon}_{t+1}$$

then (Portfolio interpretation)

$$rx_{t+1}^{(n)} - b_n \times \overline{rx}_{t+1} = \tilde{\Gamma}_n^{\top} f_t + w_{t+1}^{(n)}.$$

- Paper, Table 7:

Table 7. Forecasting the failures of the single-factor model

A. Coefficients and t-statistics

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

B. Regression statistics

Left hand var.	R^2	$\chi^2(5)$	$\sigma(\tilde{\gamma}^\top y)$	σ(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

- These *are* predictable!

- With large $R^2$and *statistical* significance!

- But they are *tiny.*

- Pattern: diagonal. $y^{(n)}$ out of line, it reverts back next period. No common factor.

- Tiny measurement errors or tiny (but profitable if you can leverage) “spread trades”

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

- This is why we do a two-step OLS not efficient GMM estimation. Efficient GMM of a single factor model weights by R^2, not size.

More stuff in paper.

Q: What about...

•Subsamples? Yes.

•Other data? McCulloch-Kwan data, not just FB interpolation