Problem Set 5 Answers

1a,b) Green = expectations. Red = Fama Bliss.

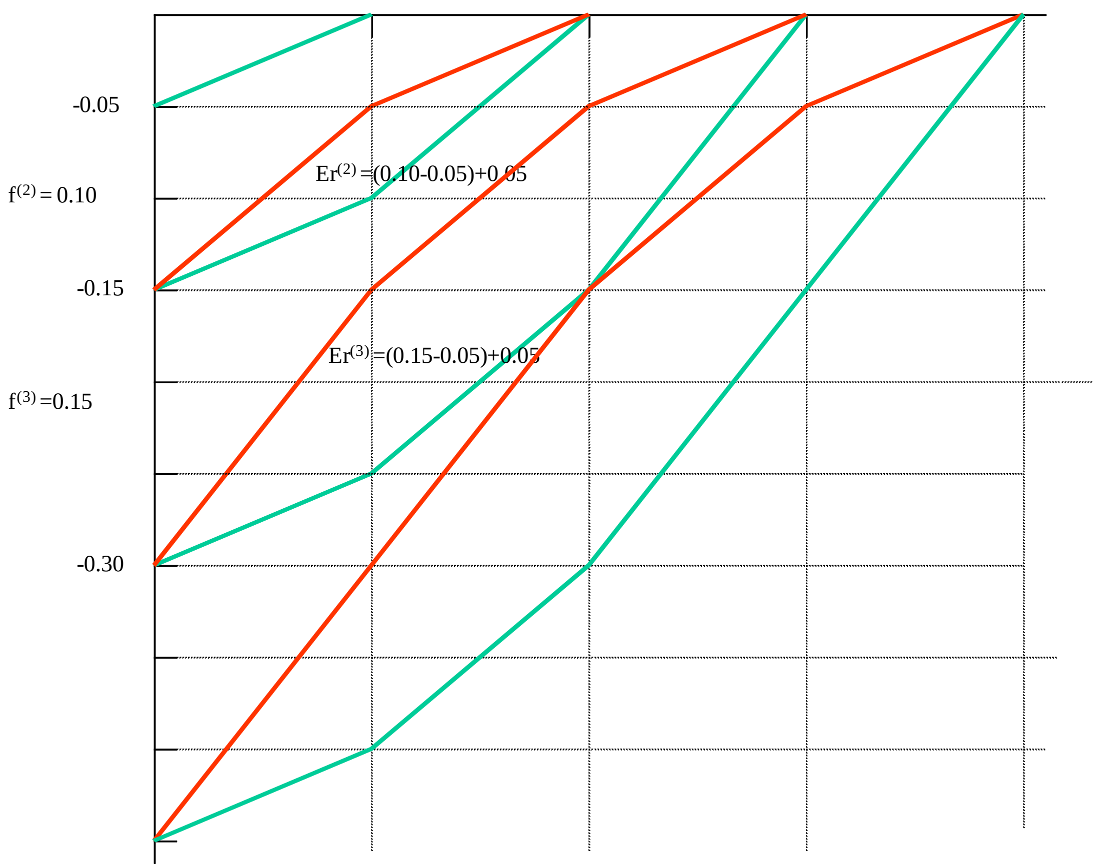


Figure 2:

Notice that iterating forward 0 and 1 one year regressions produces pure random walk behavior in all forward rates. This is inconsistent with the long-term regressions. The 5 year forward-spot spread does forecast the

c) As you can see, a return-forecast coefficient of one implies $f_t^{(n)} = E_t\left(f_{t+1}^{(n)}\right)$. This is easy too prove analytically but is less clear than the graph,

$$
\begin{aligned}
E_t rx_{t+1}^{(n)} &= 1 \times \left(f_t^{(n)} - y_t^{(1)}\right) \\
E_t\left(p_{t+1}^{(n-1)}\right) - p_t^{(n)} - y_t^{(1)} &= 1 \times \left(f_t^{(n)} - y_t^{(1)}\right) \\
E_t\left(p_{t+1}^{(n-1)}\right) - p_t^{(n)} &= 1 \times \left(f_t^{(n)}\right) \\
E_t\left(p_{t+1}^{(n-1)}\right) - E_t\left(p_{t+1}^{(n)}\right) - p_t^{(n)} + p_t^{(n+1)} &= 1 \times \left(f_t^{(n)}\right) - 1 \times \left(f_t^{(n+1)}\right) \\
E_t\left(f_{t+1}^{(n)}\right) - f_t^{(n+1)} &= 1 \times \left(f_t^{(n)}\right) - 1 \times \left(f_t^{(n+1)}\right) \\
E_t\left(f_{t+1}^{(n)}\right) &= 1 \times \left(f_t^{(n)}\right)
\end{aligned}
$$

d)

$$f_{t+1}^{(n)} = E_{t+1} y_{t+1+n-1}^{(1)}$$

by iterated expectations,

$$E_t f_{t+1}^{(n)} = E_t y_{t+n}^{(1)} = f_t^{(n+1)}$$

In sum,

$$E_t f_{t+1}^{(n)} = f_t^{(n+1)}$$

Notice the n. Fama Bliss say$E_t f_{t+1}^{(n)} = f_t^{(n)}$. The expectations hypothesis says $E_t f_{t+1}^{(n)} = f_t^{(n+1)}$ Subtracting,

$$E_t f_{t+1}^{(n)} - f_t^{(n)} = f_t^{(n+1)} - f_t^{(n)}$$

If the forward curve is rising, then long forward rates are expected to rise.

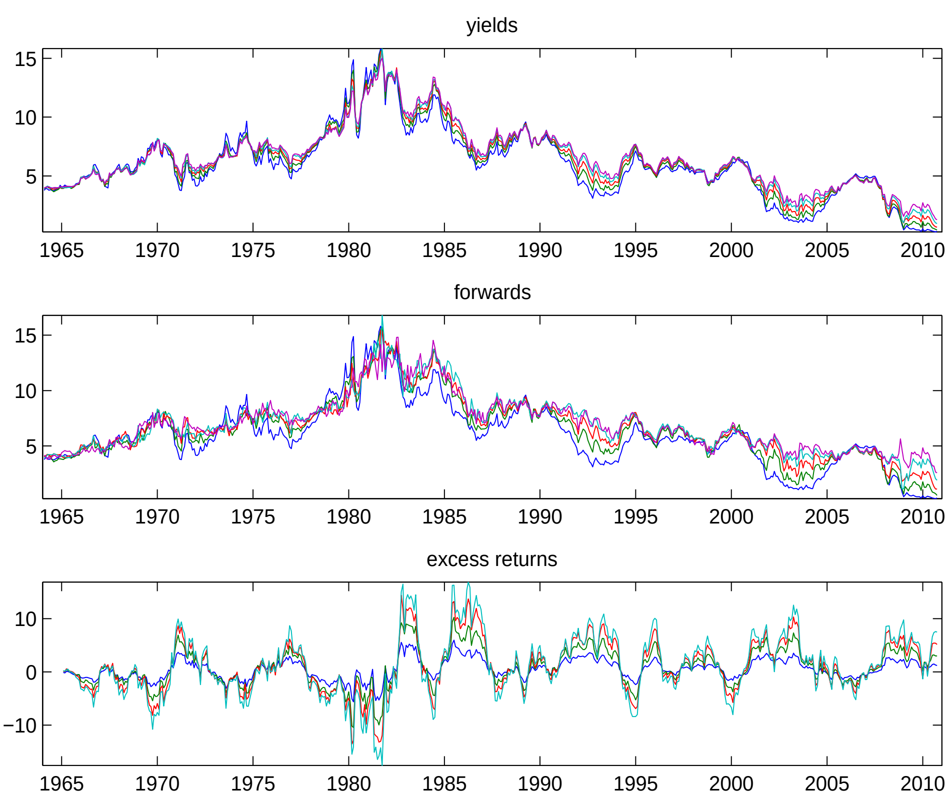
e) As you can see, the graph is not consistent with the long run expectations hypothesis. The interest rate is not a pure random walk but an AR(1) with a large, but less than one, coefficient. This already implies that the two year regression cannot have a coefficient exactly one, but must be a bit below one

$$\begin{aligned} y_{t+1}^{(1)} - y_t^{(1)} &= (1 - b^{(2)})(f_t^{(2)} - y_t^{(1)}) - \varepsilon_{t+1}^{(2)} \rightarrow \\ rx_{t+1}^{(2)} &= b^{(2)}(f_t^{(2)} - y_t^{(1)}) + \varepsilon_{t+1}^{(2)} \end{aligned}$$

We should be able to work out the implied value of $E_t y_{t+n-1}^{(1)} - y_t^{(1)}$ from the full set of excess return forecast coefficients$\{b^{(n)}\}$ in $rx_{t+1}^{(n)} = b^{(n)}(f_t^{(n)} - y_t^{(1)}) + \varepsilon_{t+1}^{(2)}$, and show that the $b^{(n)}$ must be less than one, but I got lost in identities trying to show this compactly.

2.

a. Here we go again...

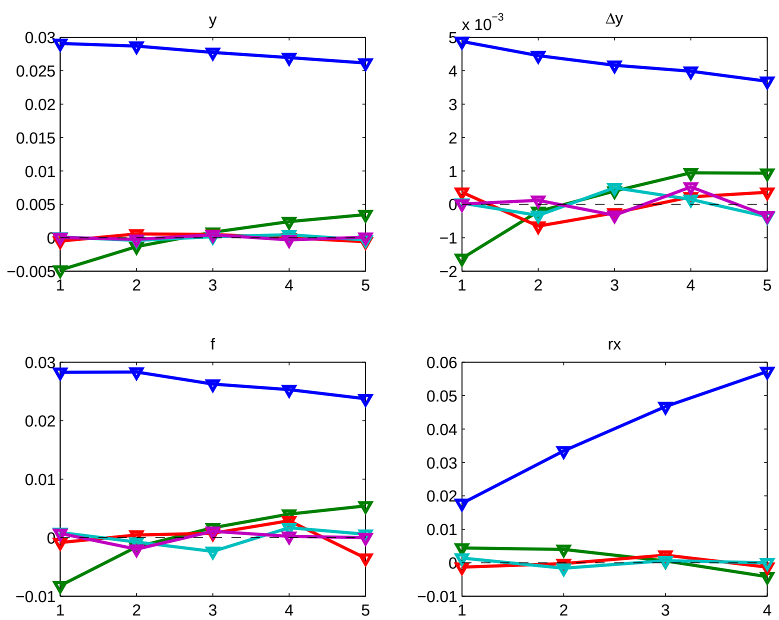


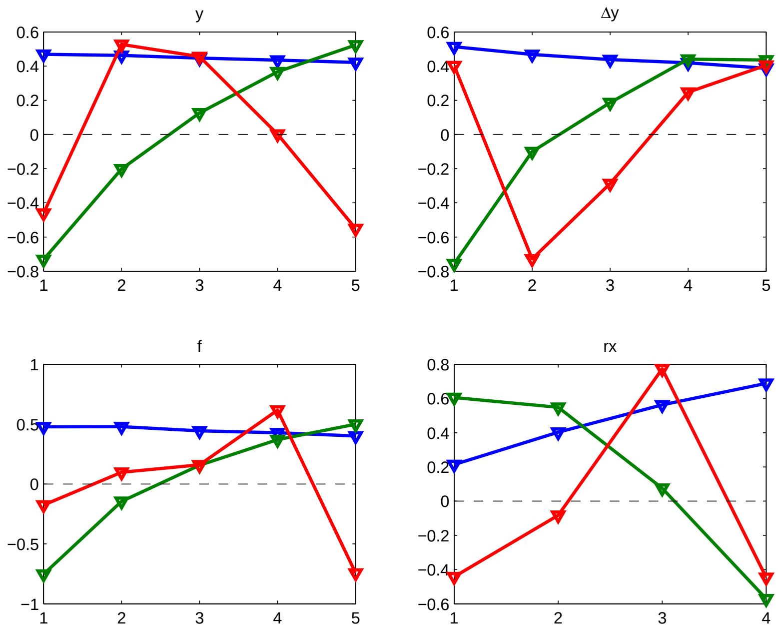
b) i)

```
variance of factors -- eigenvalues
  var y     0.01     0.01     0.03     1.12    98.83
   sd y     0.06     0.07     0.11     0.66     6.20

 var Dy     0.53     0.54     0.83     4.74    93.37
  sd Dy     0.07     0.07     0.09     0.21     0.95

 var f      0.15     0.28     0.61     3.24    95.72
  sd f      0.24     0.32     0.47     1.09     5.91

 var rx     0.07     0.12     0.76    99.04
  sd rx     0.22     0.29     0.73     8.30
```

In each case you see the first factor is really big, while the second and third add small amounts of variance and factors beyond that are tiny.

ii) Here's the picture.



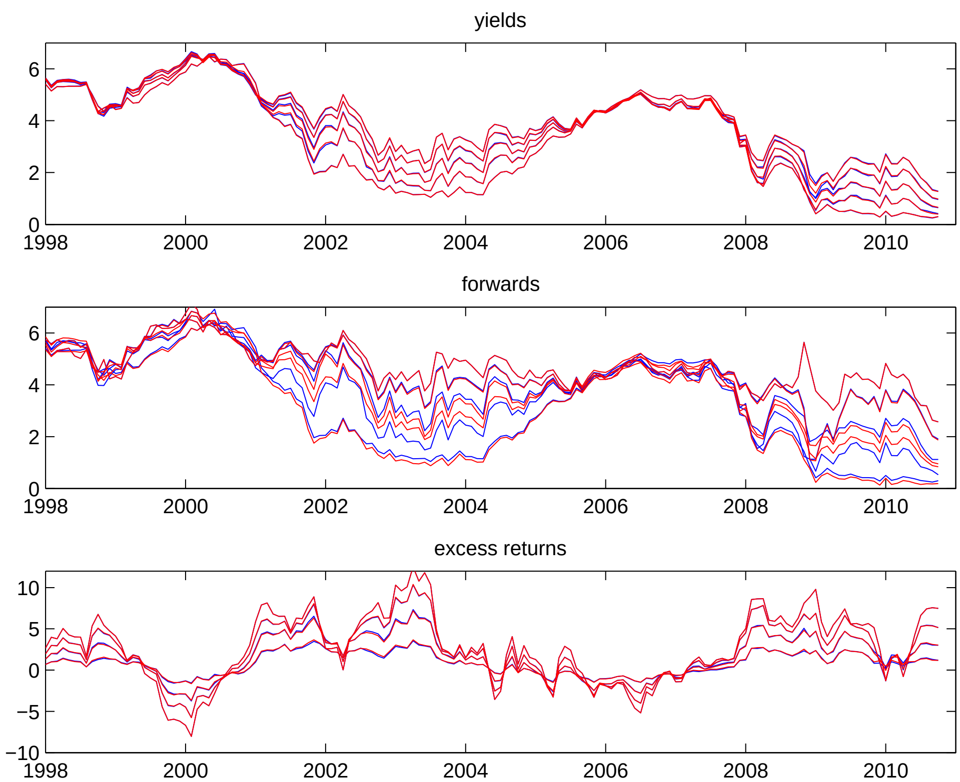


When we scale by $\lambda^{1/2}$ we see that the "level" factor which loads all maturities roughly equally is the largest, with a "slope" factor that is negative for short and positive for long maturities comes next. The others are invisible here, because they contribute such small fractions of variance. The "level" shock to yields affects longer returns more than shorter returns, so, yes, this is the same first factor. It's nice that the level factor moves long term yields and forwards a bit less than short ones since "parallel shifts" are not possible by arbitrage, so what we are seeing can leave the asymptotic yield unchanged.

The next figure plots Q itself. These are dangerous because it draws you to pay attention to the shapes of factors that are so tiny you should be ignoring them. But in figure 3 you can see much more clearly the "level" "slope" and in some cases "curvature" nature of the first three factors. I don't plot the remaining ones for clarity, but the sawtooth shapes mean they send one yield up and the others down – thus they represent "idiosyncratic" rather than "systematic" risks.

It's interesting that forwards and returns give a 4-5 spread as the "curvature" factor. I think that is because we have unsmoothed data. If you include data under one year, you routinely get a "under one year" factor too.

iii)



The yield and return plots look great, as in class. The forward rate plot does not look so great, as we found puzzing loadings. I think this is a sign of measurment error, plus the (good) fact that FB data is not smoothed. $f_t^{(n)} = p_t^{(n-1)} - p_t^{(n)}$ while $y_t^{(n)} = -\frac{1}{n} p_t^{(n)}$. *Any time you difference something across maturities you enhance idiosyncratic error and difference out the signal.*

One point of this excercise is to get you to accept the conventional wisdom that a three-factor model fits pretty well, so finding economically important information in the 4th and 5th factor, as we do is a bit surprising. The fact that the three factor model fits less well for forward rates is a hint though, which I didn't really realize until preparing this problem set.

c) Fama Bliss:

fama-bliss

n	a	b	t	R2
2.00	0.15	0.83	3.13	0.11
3.00	0.04	1.14	3.21	0.13
4.00	-0.17	1.38	3.18	0.15
5.00	0.15	1.05	2.17	0.07

fama-bliss complemetary

n	b	1-b	t	R2
2.00	0.17	0.83	0.63	0.01
3.00	-0.14	1.14	-0.39	0.00
4.00	-0.38	1.38	-0.87	0.01
5.00	-0.05	1.05	-0.11	0.00

fama-bliss long-run interest rate

n	a	b	t	R2

1.00	-0.00	0.17	0.63	0.01
2.00	-0.01	0.53	1.29	0.05
3.00	-0.01	0.84	2.76	0.14
4.00	-0.01	0.92	3.45	0.17

i, ii) Yes, FB is alive and well. Actually it's doing even better – point estimates are over one for 3 of 4 maturities, so this is beginning to look like FX where exchange rates move "the wrong way". Few anomalies are so unchanged after 20 years. R^2 is not biased by overlapping data.

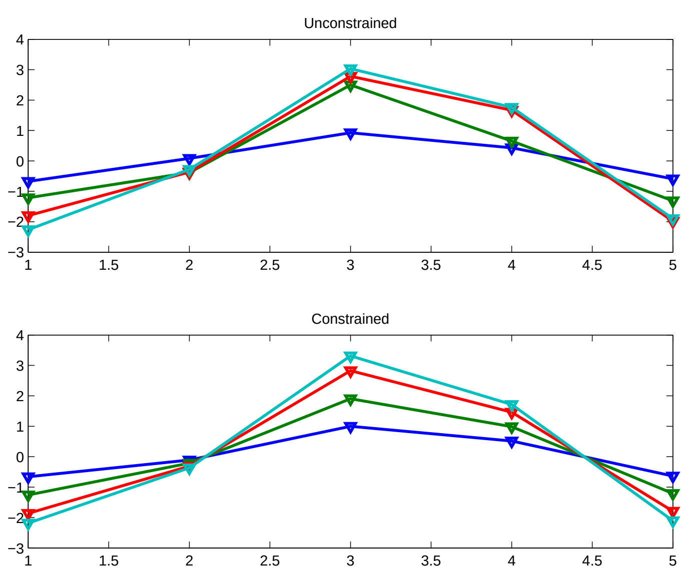
The "complementary" regressions work as they are supposed to. The point here is to hammer home what they are, and that they are not $y_{t+n-1}^{(1)} - y_t^{(1)}$.

The long run interest rate regressions in the bottom panel are also working just as well as before if not better.

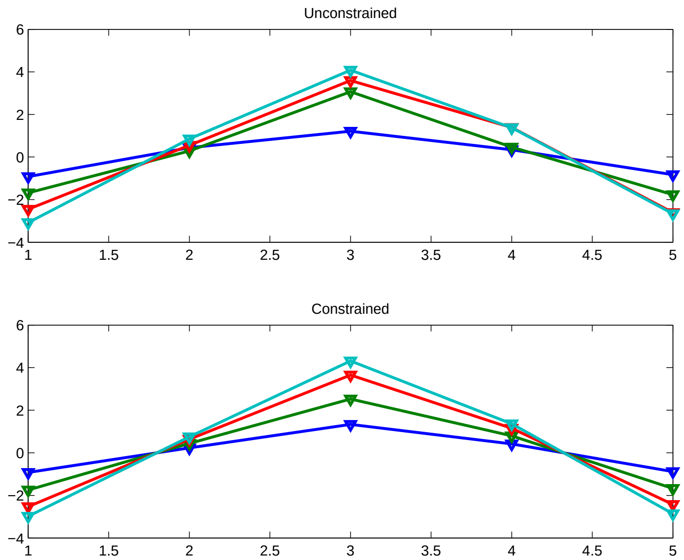
d) i)

	c	y1	f2	f3	f4	f5	R2
rx(2)	-0.01	-0.68	0.08	0.92	0.43	-0.59	0.23
rx(3)	-0.02	-1.21	-0.39	2.50	0.65	-1.31	0.24
rx(4)	-0.02	-1.80	-0.37	2.78	1.67	-1.98	0.27
rx(5)	-0.03	-2.26	-0.28	3.03	1.76	-1.90	0.25

The tent shape is distorted in this longer sample, but the "single factor" representation in which all maturities have the same shape is intact. Here's a plot of the coefficients. (Constrained comes later.) Good thing we didn't make much of the tent!

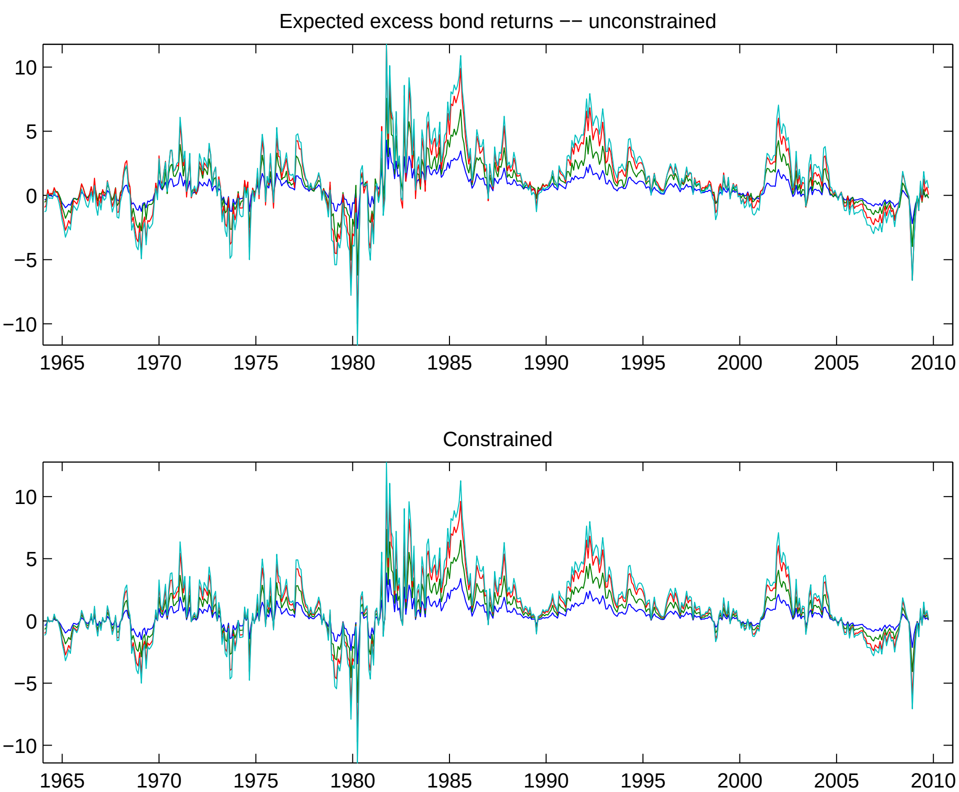


This is a feature induced by the rather weird behavior of Fama Bliss data in the Great Recession. If I stop the data in 2006, I get the usual picture



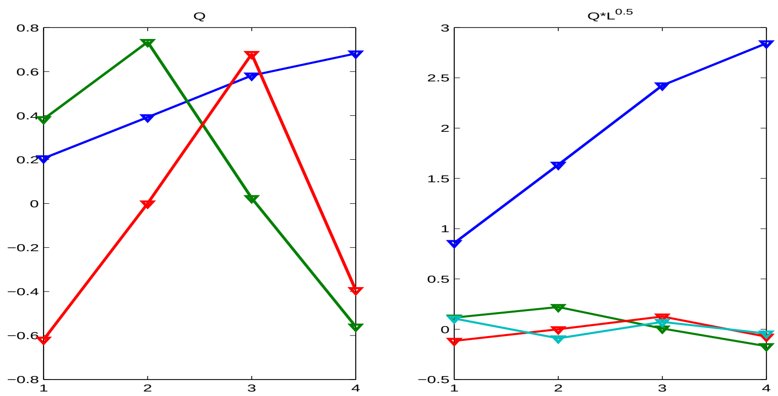
CP to 2006

ii) a) Here's the plot of expected returns over time. The notice they all seem to move together – a one factor model with long maturities moving more than short maturities. The single factor model looks almost the same as the unconstrained model. I'm having you make the plot so you see visually how it is similar to the yield case, and to really drive home the idea that we can look for *factor structure in expected returns* $cov(E_t rx_{t+1}^{(n)})$.



b)

```
fraction of variance, eigenvalues of expected returns
    0.15      0.20      0.52     99.13
```



No tent shape. This Q tells you how to construct the factor by combining *expected returns.* The tent shape tells you how to construct the factor from *forwards.*

e) Here are my regressions to forecast the average (across maturity) return.

regression of average returns on 3 yield factors, full sample

	const	curve	slope	level	R2
b	-0.00			0.07	0.01
t	-0.06			0.56	
b	-0.02		2.40	0.06	0.18
t	-1.50		3.37	0.65	
b	-0.02	3.35	2.39	0.07	0.19
t	-1.06	1.25	3.37	0.69	

Slope is the best, as per Campbell Shiller and Fama Bliss. Curve does not help. The single and multiple coefficients are almost the same as the right hand variables are orthogonal (except losing 12 data points to lags), but by runnning it in steps you see the R^2

regression of average returns on x

	const	x	R2
b	0.00	0.47	0.25
t	0.06	4.77	

X is better than alll three yield factors.

regression of average returns on x and 3 forward factors

	const	x	curve	slope	level	R2
b	-0.00	0.47	0.12	-0.02	0.01	0.25
t	-0.01	5.26	0.05	-0.03	0.08	\bigskip

And X drives them out of multiple regressions. Here you see the point, factors that explain the *variance of yields* don't necessarily capture all the information in yields about expected returns, beause that information is contained in differences of differences of yields.

Whoa, 25% R^2? This is because of the poor performance in the recession.

```
regression of average returns on 3 yield factors, sample to 2006
     const    curve    slope    level       R2
b    -0.01                       0.12     0.03
t    -0.63                       0.93
b    -0.04              2.75     0.13     0.25
t    -2.71              4.11     1.30
b    -0.03     6.15     2.73     0.14     0.27
t    -1.89     2.76     4.06     1.58
regression of average returns on x
     const        x       R2
b    -0.01     0.56     0.34
t    -1.18     7.76
regression of average returns on x and 3 yield factors
     const        x    curve    slope    level       R2
b    -0.01     0.54     2.94    -0.09     0.09     0.36
t    -0.69     7.33     1.43    -0.16     1.07        \bigskip
```

What about forward factors? Actually the three factors come pretty close, and R^2 of 0.23 is pretty close to 0.25. x still beats them all, and drives them out in multiple regressions, but the factor model comes a lot closer.

```
regression of average returns on 3 forward factors, full sample
     const    curve    slope    level       R2
b    -0.01                       0.10     0.02
t    -0.36                       0.81
b    -0.03              1.24     0.09     0.14
t    -1.51              2.86     0.87
b    -0.02     2.36     1.23     0.10     0.23
t    -1.26     4.40     3.29     1.02
regression of average returns on x
     const        x       R2
b     0.00     0.47     0.25
t     0.06     4.77
regression of average returns on x and 3 forward factors
     const        x    curve    slope    level       R2
b     0.00     0.48    -0.09    -0.06     0.00     0.25
t     0.01     2.29    -0.10    -0.09     0.04
```

And this isn't specific to the full sample either

```
regression of average returns on 3 forward factors, sample to 2006
     const    curve    slope    level       R2
```

b	-0.02			0.17	0.06
t	-0.99			1.24	
b	-0.04		1.37	0.15	0.20
t	-2.44		3.12	1.40	
b	-0.04	2.83	1.31	0.18	0.31
t	-2.64	5.90	3.81	2.13	

regression of average returns on x

	const	x	R2
b	-0.01	0.56	0.34
t	-1.18	7.76	

regression of average returns on x and 3 forward factors

	const	x	curve	slope	level	R2
b	-0.01	0.69	-0.64	-0.53	0.05	0.35
t	-0.55	4.93	-0.85	-1.06	0.60	\bigskip

What's going on? As before, the forward rates are *differences* of the yields so their curvature factor is a lot different from the yield curvature factor, and the eigenvalue technique looks deeper and almost spans x.

In more general factor models and "data" , (like the GSW data analyzed in CPII) forward rates are much smoothed compared to Fama-Bliss, so a curvature factor that loads on the 4-5 spread does not emerge.

3.

$$\begin{aligned} x_{t+1}-\delta &= \rho(x_t-\delta)+\varepsilon_{t+1} \\ \log m_{t+1} &= -x_t-\frac{1}{2}\left(\lambda_1 x_t\right)^2\sigma_\varepsilon^2-\left(\lambda_1 x_t\right)\varepsilon_{t+1} \end{aligned}$$

$$\begin{aligned} p_t^{(1)} &= \log E_t e^{-x_t-\frac{1}{2}(\lambda_1 x_t)^2\sigma_\varepsilon^2-(\lambda_1 x_t)\varepsilon_{t+1}} = -x_t \\ p_t^{(2)} &= \log E_t e^{\log m_{t+1}+p_{t+1}^{(1)}} = \log E_t e^{-x_t-\frac{1}{2}(\lambda_1 x_t)^2\sigma_\varepsilon^2-(\lambda_1 x_t)\varepsilon_{t+1}-x_{t+1}} \\ &= \log E_t e^{-x_t-\frac{1}{2}(\lambda_1 x_t)^2\sigma_\varepsilon^2-(\lambda_1 x_t)\varepsilon_{t+1}-\delta-\rho(x_t-\delta)-\varepsilon_{t+1}} \\ &= \log E_t e^{-2\delta-(1+\rho)(x_t-\delta)-\frac{1}{2}(\lambda_1 x_t)^2\sigma_\varepsilon^2-(\lambda_1 x_t)\varepsilon_{t+1}-\varepsilon_{t+1}} \\ &= \log e^{-2\delta-(1+\rho)(x_t-\delta)+\frac{1}{2}\sigma_\varepsilon^2+(\lambda_1 x_t)\sigma_\varepsilon^2} \\ p_t^{(2)} &= -2\delta-(1+\rho)\left(x_t-\delta\right)+\frac{1}{2}\sigma_\varepsilon^2+\left(\lambda_1 x_t\right)\sigma_\varepsilon^2 \end{aligned}$$

$$\begin{aligned} f_t^{(2)} &= p_t^{(1)}-p_t^{(2)} \\ &= -\delta-\left(x_t-\delta\right)+2\delta+(1+\rho)\left(x_t-\delta\right)-\frac{1}{2}\sigma_\varepsilon^2-\left(\lambda_1 x_t\right)\sigma_\varepsilon^2 \\ &= \delta+\rho\left(x_t-\delta\right)-\frac{1}{2}\sigma_\varepsilon^2-\left(\lambda_1 x_t\right)\sigma_\varepsilon^2 \\ &= \left[\delta\left(1-\rho\right)-\frac{1}{2}\sigma_\varepsilon^2\right]+\left(\rho-\lambda_1\sigma_\varepsilon^2\right)x_t \end{aligned}$$

You can see there is a time-varying distortion relative to the expectations hypothesis.

b)

$$
\begin{aligned}
rx_{t+1}^{(2)} &= p_{t+1}^{(1)} - p_t^{(2)} - y_t^{(1)} = -x_{t+1} + 2\delta + (1+\rho)(x_t - \delta) - \frac{1}{2}\sigma_\varepsilon^2 - (\lambda_1 x_t)\sigma_\varepsilon^2 - x_t \\
&= -\delta - \rho(x_t - \delta) + \varepsilon_{t+1} + 2\delta + (1+\rho)(x_t - \delta) - \frac{1}{2}\sigma_\varepsilon^2 - (\lambda_1 x_t)\sigma_\varepsilon^2 - x_t \\
&= -\frac{1}{2}\sigma_\varepsilon^2 - (\lambda_1 x_t)\sigma_\varepsilon^2 + \varepsilon_{t+1} \\
&= -\left(\frac{1}{2}\right)\sigma_\varepsilon^2 - \left(\lambda_1 \sigma_\varepsilon^2\right) x_t + \varepsilon_{t+1} \\
E_t rx_{t+1}^{(2)} &= -\left(\frac{1}{2}\right)\sigma_\varepsilon^2 - \left(\lambda_1 \sigma_\varepsilon^2\right) x_t
\end{aligned}
$$

Hence, it varies through time.

c) What does it take to get this one FB coefficient right?

$$
\begin{aligned}
f_t^{(2)} - y_t^{(1)} &= \left[\delta(1-\rho) - \left(\frac{1}{2}\right)\sigma_\varepsilon^2\right] + \left(\rho - \lambda_1\sigma_\varepsilon^2 - 1\right) x_t \\
E_t rx_{t+1}^{(2)} &= -\left(\frac{1}{2}\right)\sigma_\varepsilon^2 - \left(\lambda_1 \sigma_\varepsilon^2\right) x_t
\end{aligned}
$$

using the formula in the question, b/d is

$$
E_t rx_{t+1}^{(2)} = \alpha + \frac{\lambda_1\sigma_\varepsilon^2}{1 + \lambda_1\sigma_\varepsilon^2 - \rho}(f_t^{(2)} - y_t^{(1)})
$$

Interestingly, we can't get all the way to one. It approaches one as λ_1 rises. You do need a huge λ_1,however. Higher ρ helps. In the limit $\rho = 1$, we get huge

This is still a one factor model, so any interest rate variable is as good as any other for forecasting, $y_t^{(1)}$ would be just as good as $f_t^{(2)} - y_t^{(1)}$, which is certainly not true in the data. I would definitely want a two (at least) factor model! Obviously, we can't even try for CP regressions.

However, you have just solved a one-factor Vasicek model with time-varying risk premia that replicates (mostly) the Fama-Bliss regressions. Congratulations!