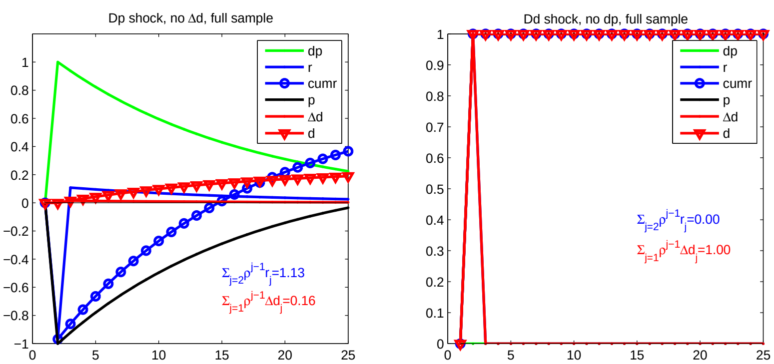
Business 35905 **John H. Cochrane**

Problem Set 3 Answers

1.

a. Full sample

	dp_t	R2
dp	0.937	0.875
t	(22.117)	
r	0.108	0.052
t	(2.183)	
dd	0.015	0.002
t	(0.380)	
ddi	0.015	0.002
t	(0.380)	



This all looks as expected. The slight positive b_d coefficient means there is a small bit of dividend growth news in the first shock, but this is all insignificant. My graphs include $\sum_{j=2}\rho^{j-1}r_j$ and $\sum_{j=1}\rho^{j-1}\Delta d_{t+j}$, the reason for the asymmetric treatment is below.

b i) By the identity $r_t = -\rho dp_t + dp_{t-1} + \Delta d_t$, the extra right hand variables are perfectly correlated. You can include two of the three of r_t, dp_{t-1} and Δd_t but not all three. How the results look will depend on which you choose

ii) Here are the three choices. The R2 values should tip you off if the impulse-responses don't-*these are all the same regression*. They produce the same forecasts and same impulse-response functions.

	dp_t	dp_t-1	dd_t	R2
dp	1.090	-0.160	-0.268	0.884
t	(9.769)	(-1.646)	(-2.082)	
r	0.088	0.015	0.232	0.073
t	(0.580)	(0.099)	(1.285)	
ddi	0.142	-0.140	-0.027	0.022
t	(1.417)	(-1.379)	(-0.203)	

	dp_t	r_t	dd_t	R2
dp	0.935	-0.160	-0.108	0.884
t	(23.867)	(-1.646)	(-0.845)	
r	0.102	0.015	0.218	0.073
t	(2.054)	(0.099)	(1.313)	
ddi	0.007	-0.140	0.113	0.022
t	(0.173)	(-1.379)	(0.923)	

	dp_t	dp_t-1	r_t	R2
dp	0.830	0.108	-0.268	0.884
t	(6.321)	(0.845)	(-2.082)	
r	0.312	-0.218	0.232	0.073
t	(1.912)	(-1.313)	(1.285)	
ddi	0.116	-0.113	-0.027	0.022
t	(1.000)	(-0.923)	(-0.203)	

I had you do it three ways to undermine your confidence in looking at individual coefficients and spotting individual t statistics. The multiple-regression coefficients depend a lot on what else is in the regression! When we replace, say r by Δd in a regression that has dp, the coefficient on that variable stays the same, since its orthogonal information with respect to dp is the same. The dp coefficient changes.

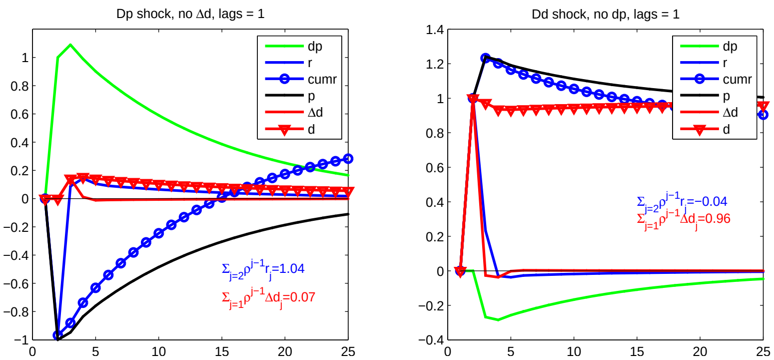
For example dd_t (top row) has a –0.268 coefficient and 2.082 t statistic with dp_t-1, but only a -0.108 coefficient and 0.845 t stat with r_t. Note also that the coefficient is the same in the bottom row with r_t in place of dd_t. *Conditioned on dividend yields*, returns give the same information as dividend growth.

In the return-forecasting rows, not much looks significant here. However, "economic significance" is a tipoff. In the last regression,

$$
\begin{aligned}
r_{t+1} &= 0.312dp_t - 0.218dp_{t-1} + 0.232r_{t-1} \\
&= 0.1dp_t + 0.2(dp_t - dp_{t-1}) + 0.23r_{t-1}
\end{aligned}
$$

This looks like a fairly strong "recent change" effect in dividend yields, and a very strong return autocorrelation coefficient. The dp regression shows the same behavior. The t stats are not huge, but a rise in R^2 from 5 to 7% is important in this business. Not enough to hang your hat on in this sample, but suggestive. (In other data sets, this "recent change" in dividend yield or b/m ratio is much more important.)

The impulse response functions are the same for all three ways of running the regression. This is one strong argument for evaluating these sorts of questions with the impulse-response function rather than the regression coefficients. The VAR literature does this routinely; we do not try detailed t stat fishing, but instead run fairly overparameterized regressions and hope that statistically small coefficients don't muck up the responses too much. We'll see that philosophy is a bit dangerous too, especially once you start raising things to many powers.



These responses look to me basically like the response with no lags. Dp gets nice "hump-shaped" responses because of the "recent change" effect.

Looking at coefficients, 2/3 of which have $dd_{t+1} = 0.11dp_t$, the same as the previous r coefficient, you might have thought "wow, dividend growth is now predicitable!" But dd is all responding to recent change (if recent change helps forecast returns, it must help forecast dividend growth), so the $-0.11dp_{t-1}$ coefficient quickly undoes it. This is another case of "short run forecastability" that does not translate into "long run forecastability."

As I sum of the responses, we add some wiggles but they are not economically (or statistically) significant.

c) More lags. By using one more lag of dp than the others, it doesn't matter whether you use returns or dividends, up to the approximation.

$$r_t = -\rho dp_t + dp_{t-1} + \Delta d_t$$

so the space $(r_t, r_{t-1}, dp_t, dp_{t-1}, dp_{t-2})$ is the same as the space $(\Delta d_t, \Delta d_{t-1}, dp_t, dp_{t-1}, dp_{t-2})$. I ran the regression using returns and implied dividends.

Here are my regressions. I chose to express the right hand side with use one extra dp lag and returns, though of course you can reexpress it different exactly equivalent ways which I did not explore.

	dp_t	dp_t-0	r_t-0	R2
dp	0.830	0.108	-0.268	0.884
t	(6.321)	(0.845)	(-2.082)	
r	0.312	-0.218	0.232	0.073
t	(1.912)	(-1.313)	(1.285)	
ddi	0.116	-0.113	-0.027	0.022
t	(1.000)	(-0.923)	(-0.203)	

	dp_t	dp_t-0	r_t-0	dp_t-1	r_t-1	R2
dp	0.907	0.121	-0.244	-0.084	0.202	0.890
t	(6.275)	(0.614)	(-1.793)	(-0.535)	(1.132)	
r	0.250	-0.235	0.207	0.074	-0.160	0.081
t	(1.406)	(-0.966)	(1.116)	(0.308)	(-0.769)	
ddi	0.128	-0.117	-0.029	-0.007	0.036	0.025
t	(1.037)	(-0.645)	(-0.216)	(-0.041)	(0.270)	

var coeffs, dp-1, dp-2, r-1, dp-3, r-2, etc. Last col is R2

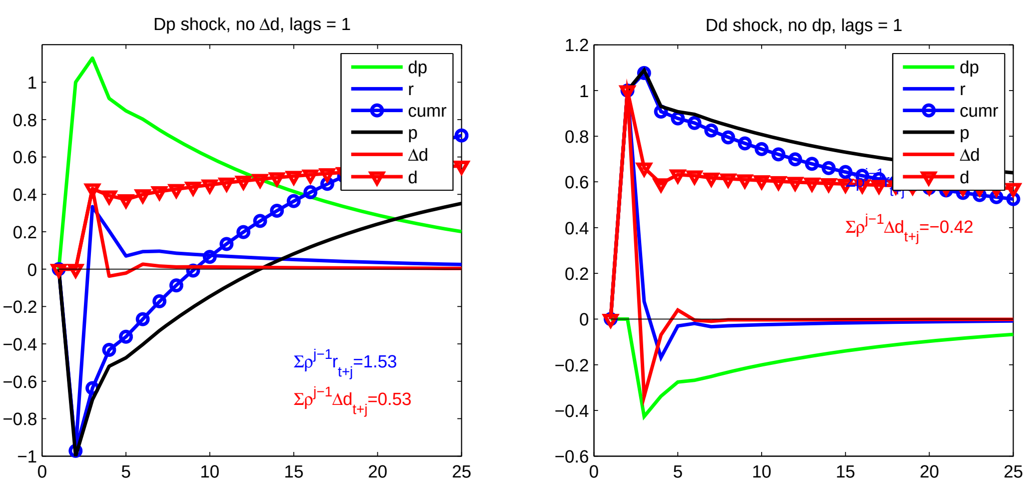
	dp_t	dp_t-0	r_t-0	dp_t-1	r_t-1	dp_t-2	r_t-2	R2
dp	0.914	0.125	-0.237	-0.303	0.225	0.223	-0.104	0.893
t	(5.976)	(0.557)	(-1.793)	(-1.466)	(1.247)	(1.432)	(-0.734)	
r	0.251	-0.259	0.218	0.299	-0.168	-0.212	0.054	0.100
t	(1.327)	(-0.990)	(1.180)	(0.955)	(-0.814)	(-1.108)	(0.343)	
ddi	0.135	-0.139	-0.011	0.005	0.050	0.004	-0.047	0.024
t	(1.043)	(-0.762)	(-0.084)	(0.023)	(0.369)	(0.027)	(-0.422)	

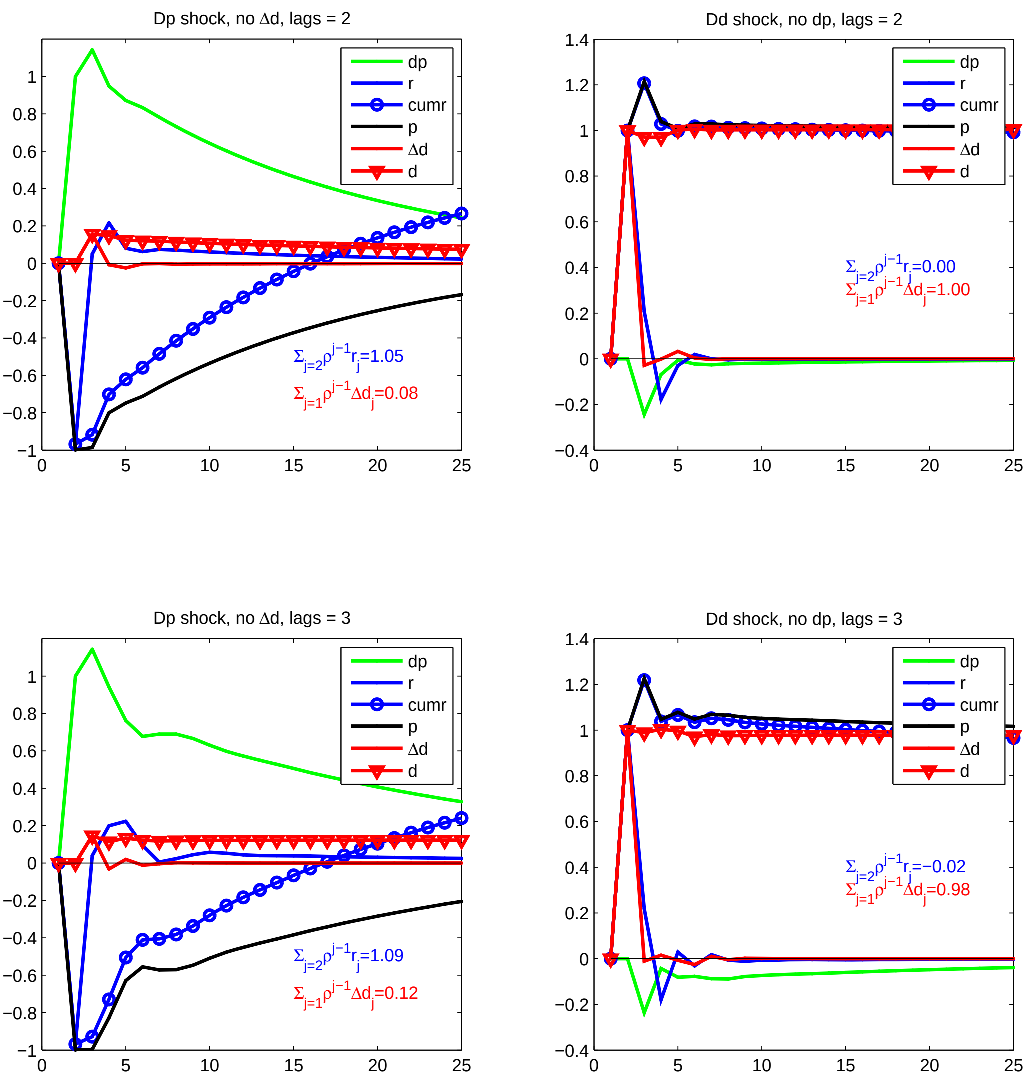
var coeffs, dp-1, dp-2, r-1, dp-3, r-2, etc. Last col is R2

	dp_t	dp_t-0	r_t-0	dp_t-1	r_t-1	dp_t-2	r_t-2	dp_t-3	r_t-3	R2
dp	0.949	0.086	-0.209	-0.308	0.217	0.298	-0.124	-0.075	0.036	0.894
t	(6.112)	(0.383)	(-1.508)	(-1.427)	(1.203)	(1.237)	(-0.879)	(-0.429)	(0.273)	
r	0.171	-0.143	0.141	0.238	-0.144	-0.264	0.083	0.087	-0.140	0.096
t	(1.013)	(-0.626)	(0.778)	(0.770)	(-0.733)	(-0.872)	(0.572)	(0.382)	(-0.909)	
ddi	0.090	-0.060	-0.062	-0.060	0.066	0.025	-0.037	0.015	-0.105	0.044
t	(0.754)	(-0.389)	(-0.465)	(-0.267)	(0.493)	(0.114)	(-0.329)	(0.092)	(-0.954)	

The dp forecasting numbers are interesting, but note the emerging pattern of offsetting + and - coefficients, which is often a sign of overfitting. The same thing happens in the r equations. Big numbers but suspicious offsetting positive and negative signs. The R^2 increase is attractive. But I would have to see some interpretatble pattern in the coefficients before I believe it. That may just mean I haven't looked hard enough.

Here are my impulse responses. As you can see the somewhat offsetting numbers don't add up to much change in the actual response functions. There seems to be a bit of mean-reversion in the dividend growth response but that's about it. Again, this means that the extra variables really aren't doing much to change our picture of where *dividend yield volatility* comes from. They may however change our picture of *one period expected returns*, if we could believe some of the R^2 rise in these vastly overfit regressions.





d) in retrospect, I am getting to like more the representation with a single lag of dp and lots of returns and dividend growth. I am influenced here by Ralph Koijen's ideas about filtering lagged returns and dividend growth to get at additional expected return variation not revealed by dp. But then it's harder to see the "change in dividend yield" effect, if it's there. I guess as always, the best approach is to understand all the ways of looking at the regressions and how to relate them.

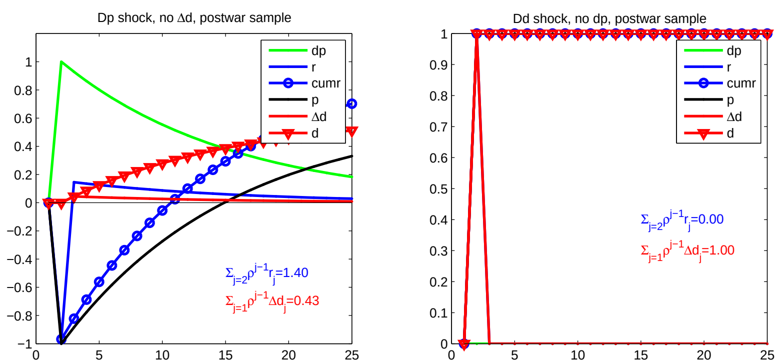
2

a)

	dp_t	R2
dp	0.929	0.891
t	(20.573)	
r	0.146	0.108
t	(2.469)	

```
dd    0.049    0.023
 t  ( 0.923)
ddi   0.045    0.023
 t  ( 0.843)\
```

There is a substantial difference – the dividend growth coefficient is a lot bigger! It's half as big as the return coefficient, but in the "wrong" direction. Of course, it's still not close to siginificant. But VARs use point estimates, so we're going to have pretty big changes in the point estimates. Here's the response function. As you might have suspected, the expected dividend plot is now throughout at about half the level of the expected return plot and cumulative expected dividends rise.



b) Now we have some serious coefficients to think about.

```
      dp_t   dp_t-0   r_t-0      R2
 dp    0.672    0.277   -0.427  0.904
  t  ( 4.695)( 1.825)(-2.898)
  r    0.405   -0.262    0.065   0.148
  t  ( 3.294)(-1.911)( 0.405)
ddi    0.055    0.006   -0.348   0.206
  t  ( 0.385)( 0.040)(-2.414)
```

- Both lagged dp and lagged r enter the dp equation. Lagged dp isn't really that big a deal since it has a positive coefficient. But the 2.98 t stat on lagged r is impressive.

- In the return forecast, we again can usefully orthogonalize to a "recent change" effect, which now has a 1.9 t statistic as well as a good economic intuition. The increase in R^2 from 10 to 15% is enough for many papers to get in the JF.

$$\begin{aligned} r_{t+1} &= 0.41dp_t - 0.26dp_{t-1} + 0.07r_{t-1} \\ &= 0.15dp_t - 0.26\left(dp_t - dp_{t-1}\right) + 0.07r_{t-1} \end{aligned}$$

- Dividends now are *really* forecastable. The R^2 is now 0.2, larger than returns! The main mechanism is feedback from lagged returns – a big lagged return holding dp constant means a lower future dividend growth. That seems weird of course – shouldn't returns rise on news of good dividends? Aha, but we're conditioning on dp. Maybe looking at it another way helps.

```
      dp_t  dp_t-0  dd_t-0      R2
 dp   1.085  -0.149  -0.427  0.904
  t ( 8.034)(-1.190)(-2.898)
  r   0.343  -0.197   0.065   0.148
  t ( 1.986)(-1.133)( 0.405)
ddi   0.393  -0.342  -0.348   0.206
  t ( 3.765)(-3.138)(-2.414)

       dp_t   r_t-0  dd_t-0      R2
 dp   0.940  -0.149  -0.277  0.904
  t (22.906)(-1.190)(-1.825)
  r   0.152  -0.197   0.262   0.148
  t ( 2.455)(-1.133)( 1.911)
ddi   0.062  -0.342  -0.006   0.206
  t ( 1.218)(-3.138)(-0.040)

       dp_t  dp_t-0   r_t-0      R2
 dp   0.672   0.277  -0.427  0.904
  t ( 4.695)( 1.825)(-2.898)
  r   0.405  -0.262   0.065   0.148
  t ( 3.294)(-1.911)( 0.405)
ddi   0.055   0.006  -0.348   0.206
  t ( 0.385)( 0.040)(-2.414)
```

Notice the great 2.9 t stat in dp forecast disappears in the middle representation. Of course r and dd are highly corrleated

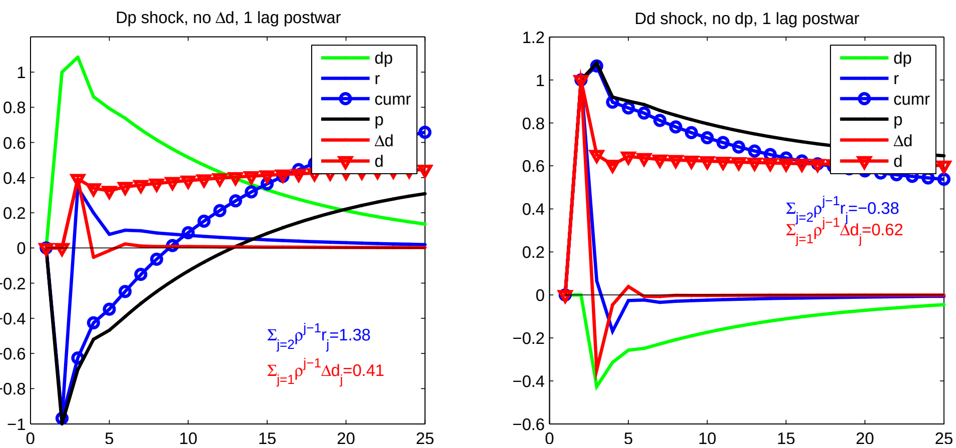
The "source" of better return forecasts looks different across the three representations. With dp, dd, or dp, r it looks like a "recent change in dp" effect, $r_{t+1} = 0.35dp_t - 0.2dp_{t-1} = 0.15dp_t + 0.2(dp_t - dp_{t-1})$. That one makes sense to me because the sign is right – a short-term second factor in expected returns would do this.

In the r, dd regression it looks like a weird sort of momentum effect – high dividend growth Δd_t implies higher returns r_{t+1}.

In the dividend growth regressions, we see the "recent change in dp effect" that (by identity) forecasts dividend growth and returns, – plus a serious -0.348 serial correlation. However, in the other regressions it either looks like negative cross-correlation from lagged return. Of course all these are equivalent, and multiple regressions always depend on what else is held constant.

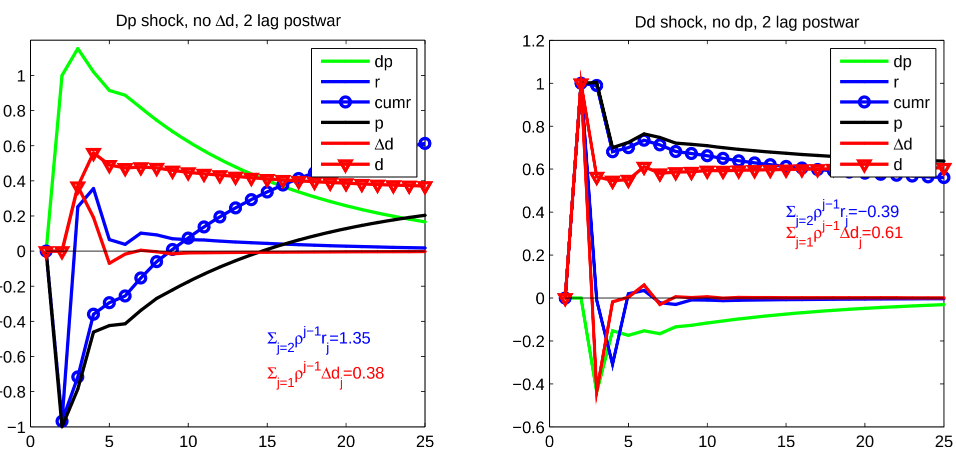
The regression with r and dd on the right hand side is interesting. I expected "filtering", positive autocorrelation. Instead I got negative autocorrelation for both returns and dividend growth.

Let's look at the response function:



Now, there is a serious difference. Our dp shock means a big negative current return, so forecasts one period of dividend growth, sending dividends permanently higher as well as the usual sort of pattern for expected returns. We also see an expected return effect to the dividend yield shock.

As I look to two lags, I see the same basic pattern. So these first lags pretty much captured the important difference for the postwar data



3)

a) Impulse response functions are just revisions in expectations:

$$(E_{t+1}-E_t) \quad : \quad dp_t = \sum_{j=1}^{\infty}\rho^{j-1}r_{t+j} - \sum_{j=1}^{\infty}\rho^{j-1}\Delta d_{t+j}$$

$$0 = \sum_{j=1}^{\infty}\rho^{j-1}e_{z_{t+1}\to r_{t+j}} - \sum_{j=1}^{\infty}\rho^{j-1}e_{z_t\to\Delta d_{t+j}}$$

The latter is Campbell's return identity,

$$e_{z_{t+1}\to r_{t+1}} = -\sum_{j=2}^{\infty}\rho^{j-1}e_{z_{t+1}\to r_{t+j}} + \sum_{j=1}^{\infty}\rho^{j-1}e_{z_t\to\Delta d_{t+j}}$$

Rotating the responses is easy. The "expected return" response is

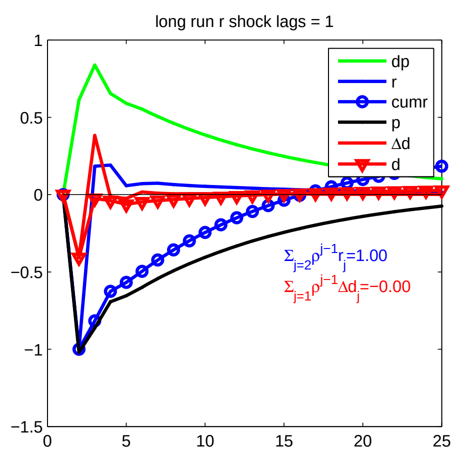
$$ae_{dp_{t+1}\to r_{t+j}} + be_{\Delta d_t\to \Delta d_{t+j}}.$$

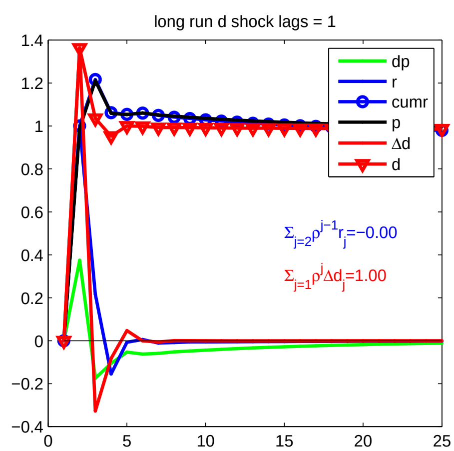
We pick a and b so that

$$\begin{aligned} a\left(\sum_{j=2}^{\infty}\rho^{j-1}e_{dp_{t+1}\to r_{t+j}}\right)+b\left(\sum_{j=1}^{\infty}\rho^{j-1}e_{\Delta d_t\to \Delta d_{t+j}}\right) &= 1 \\ a\left(\sum_{j=2}^{\infty}\rho^{j-1}e_{dp_{t+1}\to r_{t+j}}\right)+b\left(\sum_{j=1}^{\infty}\rho^{j-1}e_{\Delta d_t\to \Delta d_{t+j}}\right) &= 0 \end{aligned}$$

For the other shock, we pick 0 and 1. This is just $A^{-1}[10]'$.

Here's the responses:





You can see that the "long run" right hand side of these responses looks very much as they did in the simple system, with $b_d = 0$. The "short run" is different. For the expected return shock, now dividend growth also is shocked downward, along with a smaller movement in dividend yield. The dividend growth rebounds so that the overall sum of dividend growth is zero. The expected cashflow shock (right) now has a very small return bounce.

```
  time 0 responses --     dp         r         dd
"expected return"       0.6123   -1.0000   -0.4072
"expected cashflow"     0.3748    1.0000    1.3628
```