Exchange rate risk and the forward premium anomaly
Revised: April 15, 2008

These notes follow Fama (JME, 1984) in outlining the statistical properties of exchange rate risk premia implied by the data. It then turns to the implications of random walk exchange rates for models of monetary policy as recently emphasized by Alvarez, Atkeson and Kehoe (AER, 2007).

1 Exchange rate risk premia

Recall from Note 1 that two basic predictions of international monetary economics are *covered interest parity* (CIP), meaning

$$f_t - e_t = i_t - i_t^*$$

and *uncovered interest parity* (UIP), meaning

$$\mathbb{E}_t\{\Delta e_{t+1}\} = i_t - i_t^*$$

where f_t is the (log) one-period forward nominal exchange rate, e_t is the (log) one-period spot nominal exchange rate, i_t is the one-period nominal interest rate on home currency dollar bonds, and i_t^* is the one-period nominal interest rate on foreign currency euro bonds. The covered interest parity relation is exact up to the approximation $\log(1 + x) \approx x$ for small x (so it is exact in continuous time) and holds well in the data up to a small adjustment for bid/ask spreads arising from the OTC nature of foreign exchange trading.

By contrast, the uncovered interest parity relation which comes from log-linearizing the bond pricing equations for each countries' one-period nominal bonds, does not hold in the data. To see this, notice that CIP and UIP together imply

$$f_t = \mathbb{E}_t\{e_{t+1}\}$$

so that the forward is an unbiased forecast of the future spot exchange rate. However if we run a regression of the form

$$e_{t+1} - e_t = \alpha_0 + \alpha_1(f_t - e_t) + \text{noise}$$

a typical estimate is $\hat{\alpha}_1 = -1$ or thereabouts, not the $\hat{\alpha}_1 = +1$ that we expect from the uncovered interest parity hypothesis. Roughly speaking, theory says an investor should demand a premium $i_t - i_t^* > 0$ for holding euro bonds if the euro is expected to *depreciate* against the dollar, $\mathbb{E}_t\{\Delta e_{t+1}\} > 0$, but the data say the reverse, relatively high interest rates go hand-in-hand with exchange rate *appreciations*, $\mathbb{E}_t\{\Delta e_{t+1}\} < 0$.

The term $f_t - e_t$ is the (log) forward/spot premium and so this fact is often known as the 'forward premium anomaly'.

Let us write the forward premium as

$$f_t - e_t = f_t - \mathbb{E}_t\{e_{t+1}\} + \mathbb{E}_t\{e_{t+1}\} - e_t$$

If we define an exchange rate *risk premium* p_t by

$$p_t := f_t - \mathbb{E}_t\{e_{t+1}\} \tag{1}$$

then we have

$$f_t - e_t = p_t + \mathbb{E}_t\{\Delta e_{t+1}\} = i_t - i_t^*$$

(taking CIP as given).

Notice that the use of the term 'risk premium' is a little bit misleading: at the moment we're really just giving a label to our ignorance, just as we call the residual in a production function accounting exercise 'total factor productivity'.

We can now do a little bit of work to figure out the statistical properties of the risk premium that would give rise to $\hat{\alpha}_1 < 0$. This discussion essentially follows Fama (JME, 1984). See also Backus, Foresi and Telmer (JF, 2001).

Basic econometrics. The probability limit of the slope coefficient in the Fama regression is

$$\alpha_1 = \frac{\mathrm{Cov}\{f_t - e_t, \Delta e_{t+1}\}}{\mathrm{Var}\{f_t - e_t\}}$$

Rational expectations. Rational expectations requires that $\mathbb{E}_t\{e_{t+1}\} - e_{t+1}$ must be uncorrelated with any information that is observable at date t (why?). Because of this, we have

$$\mathrm{Cov}\{f_t - e_t, \Delta e_{t+1}\} = \mathrm{Cov}\{f_t - e_t, \mathbb{E}_t\{\Delta e_{t+1}\}\}$$

Hence

$$\alpha_1 = \frac{\mathrm{Cov}\{f_t - e_t, \mathbb{E}_t\{\Delta e_{t+1}\}\}}{\mathrm{Var}\{f_t - e_t\}} = \frac{\mathrm{Cov}\{p_t + \mathbb{E}_t\{\Delta e_{t+1}\}, \mathbb{E}_t\{\Delta e_{t+1}\}\}}{\mathrm{Var}\{p_t + \mathbb{E}_t\{\Delta e_{t+1}\}\}}$$

Properties of covariances. If we have two random variables X and Y, then

$$\mathrm{Cov}\{X, X+Y\} = \mathrm{Var}\{X\} + \mathrm{Cov}\{X, Y\}$$

So

$$\alpha_1 = \frac{\mathrm{Cov}\{p_t + \mathbb{E}_t\{\Delta e_{t+1}\}, \mathbb{E}_t\{\Delta e_{t+1}\}\}}{\mathrm{Var}\{p_t + \mathbb{E}_t\{\Delta e_{t+1}\}\}} = \frac{\mathrm{Var}\{\mathbb{E}_t\{\Delta e_{t+1}\}\} + \mathrm{Cov}\{p_t, \mathbb{E}_t\{\Delta e_{t+1}\}\}}{\mathrm{Var}\{p_t + \mathbb{E}_t\{\Delta e_{t+1}\}\}}$$

Several implications follow from this calculation. First, if the risk premium were constant, $p_t = p$ all t, then $\text{Var}\{p_t\} = 0$ and we would have slope coefficient

$$\alpha_1 = \frac{\text{Var}\{\mathbb{E}_t\{\Delta e_{t+1}\}\}}{\text{Var}\{\mathbb{E}_t\{\Delta e_{t+1}\}\}} = 1$$

no matter what the statistical properties of $\mathbb{E}_t\{\Delta e_{t+1}\}$. So in order for the slope coefficient to be other than one, we definitely have to have a *time-varying risk premium*. Also, since $\text{Var}\{\mathbb{E}_t\{\Delta e_{t+1}\}\} \geq 0$, in order for us to have $\alpha_1 < 0$ we must have a risk premium with the property

$$\text{Cov}\{\mathbb{E}_t\{\Delta e_{t+1}\}, p_t\} < 0$$

That is, there must be a systematic tendency for expected *appreciations*, $\mathbb{E}_t\{\Delta e_{t+1}\} < 0$, to go hand-in-hand with *increases in the risk premium*. Moreover, the negative covariance has to have magnitude such that

$$|\text{Cov}\{\mathbb{E}_t\{\Delta e_{t+1}\}, p_t\}| > \text{Var}\{\mathbb{E}_t\{\Delta e_{t+1}\}\}$$

As we will see below, this requirement will turn out not to be particularly stringent, so the literature does not focus on it much.

Another implication for the statistical properties of the risk premium comes from expanding the denominator of α_1 so that we have

$$\alpha_1 \quad = \quad \frac{\text{Var}\{\mathbb{E}_t\{\Delta e_{t+1}\}\} + \text{Cov}\{p_t, \mathbb{E}_t\{\Delta e_{t+1}\}\}}{\text{Var}\{\mathbb{E}_t\{\Delta e_{t+1}\}\} + 2\text{Cov}\{p_t, \mathbb{E}_t\{\Delta e_{t+1}\}\} + \text{Var}\{p_t\}}$$

Now recall that a typical estimate for α_1 is negative. But whenever $\alpha_1 < \frac{1}{2}$, we must have

$$\text{Var}\{\mathbb{E}_t\{\Delta e_{t+1}\}\} + \text{Cov}\{p_t, \mathbb{E}_t\{\Delta e_{t+1}\}\} < \frac{1}{2}\left[\text{Var}\{\mathbb{E}_t\{\Delta e_{t+1}\}\} + 2\text{Cov}\{p_t, \mathbb{E}_t\{\Delta e_{t+1}\}\} + \text{Var}\{p_t\}\right]$$

or

$$\text{Var}\{\mathbb{E}_t\{\Delta e_{t+1}\}\} < \frac{1}{2}\left[\text{Var}\{\mathbb{E}_t\{\Delta e_{t+1}\}\} + \text{Var}\{p_t\}\right]$$

In short

$$\text{Var}\{p_t\} > \text{Var}\{\mathbb{E}_t\{\Delta e_{t+1}\}\}$$

Hence the risk premium must be relatively volatile in comparison with expected depreciations.

To sum up: in order to get a negative slope coefficient in the forward premium regression under rational expectations, we must have a risk premium that is relatively volatile (compared to the expected depreciation) and that co-varies negatively with the expected depreciation.

The requirement that the risk premium co-vary negatively with expected depreciations is particularly counterintuitive if one thought of certain currencies as being safe-havens during times of distress in global financial markets. That is, if there are '*flights to quality*,' maybe one ought to expect currencies to appreciate when their risk premium goes down! This would imply that $\text{Cov}\{p_t, \mathbb{E}_t\{\Delta e_{t+1}\}\} > 0$.

Random walks. It is also worth bearing in mind that nominal exchange rate movements are well approximated by random walks (see Meese and Rogoff, JIE 1983) so $\mathbb{E}_t\{e_{t+1}\} \approx a + e_t$ (some constant a) and $\mathrm{Var}\{\mathbb{E}_t\{\Delta e_{t+1}\}\} \approx 0$. So the forward premium puzzle may not really be that the volatility of the risk premium is so large but instead that the volatility of expected depreciations is *so small*.

Notice that if exchange rates were a random walk so that $\mathrm{Var}\{\mathbb{E}_t\{\Delta e_{t+1}\}\} = 0$ **nearly all of the observed variations in interest differentials are accounted for by the movements in the 'risk premium'**. That is, if $\mathrm{Var}\{\mathbb{E}_t\{\Delta e_{t+1}\}\} = 0$, we have

$$i_t - i_t^* = f_t - e_t = p_t + \mathbb{E}_t\{\Delta e_{t+1}\} \approx p_t + a$$

and so

$$\mathrm{Var}\{i_t - i_t^*\} \approx \mathrm{Var}\{p_t\}$$

We couldn't get a negative coefficient α_1 in the Fama regression if exchange rates were a pure random walk and they are not. But they are well approximated as such.

2 Implications for monetary policy?

We now take this statistical characterization of the properties of risk premia implied by exchange rate data and, following Alvarez, Atkeson and Kehoe (AER, 2007), ask 'what are the implications for models of monetary policy?'.

Notation. Let M_{t+1} denote the nominal stochastic discount factor (SDF) for home currency assets so that, in a slight abuse of previous notation, $\exp(-i_t) = \mathbb{E}_t\{M_{t+1}\}$ where i_t denotes the *continuously compounded* nominal interest rate on one period zero coupon bonds denominated in home currency.

Nominal SDFs and exchange rates. Recall from Note 1 that no-arbitrage in the (complete) international asset market implies

$$M_{t+1}^* = M_{t+1}\frac{\mathcal{E}_{t+1}}{\mathcal{E}_t}$$

where $\mathcal{E}_t$ is the spot nominal exchange rate. This follows from the uniqueness of the pricing kernel (and the SDF) with complete markets. Taking logs and expectations then

$$\mathbb{E}_t\{\log M_{t+1}^*\} = \mathbb{E}_t\{\log M_{t+1}\} + \mathbb{E}_t\{\Delta e_{t+1}\}$$

Nominal SDFs and interest rates. Again from Note 1, the interest rate on dollar bonds is

$$i_t = -\log \mathbb{E}_t\{M_{t+1}\}$$

and similarly the interest rate on euro bonds is

$$i_t^* = -\log \mathbb{E}_t\{M_{t+1}^*\}$$

Risk premia and the volatility of SDFs. Since the risk premium can be written

$$p_t = i_t - i_t^* - \mathbb{E}_t\{\Delta e_{t+1}\}$$

we can now write it as the difference between the 'log of the expectation' and the 'expectation of the log,' specifically

$$p_t = \left(\log \mathbb{E}_t\{M_{t+1}^*\} - \mathbb{E}_t\{\log M_{t+1}^*\}\right) - \left(\log \mathbb{E}_t\{M_{t+1}\} - \mathbb{E}_t\{\log M_{t+1}\}\right)$$

This is a generalized measure of the volatility of the SDFs.

An aside: specifically, the measure of volatility of a random variable X given by $\log \mathbb{E}(X) - \mathbb{E}(\log X)$ is known as 'Theils second entropy measure'. Alvarez and Jermann (Econometrica, 2005) make extensive use of this measure of volatility to characterize the persistence of the pricing kernel implied by long term interest rates.

As a specific example, suppose the nominal SDF is conditionally lognormal so that $\log M_{t+1}$ is normal with conditional mean $\mathbb{E}_t\{\log M_{t+1}\}$ and conditional variance $\text{Var}_t\{\log M_{t+1}\}$ (and similarly for the foreign SDF). Then the conditional mean of M_{t+1} is

$$\mathbb{E}_t\{M_{t+1}\} = \exp\left(\mathbb{E}_t\{\log M_{t+1}\} + \frac{1}{2}\text{Var}_t\{\log M_{t+1}\}\right)$$

so

$$\log \mathbb{E}_t\{M_{t+1}\} - \mathbb{E}_t\{\log M_{t+1}\} = \frac{1}{2}\text{Var}_t\{\log M_{t+1}\}$$

and similarly for the foreign SDF. So in this lognormal example the difference between the log of the expectation and the expectation of the log is proportional to the conditional variance and the risk premium is the difference of these terms across the two countries

$$p_t = \frac{1}{2}\left[\text{Var}_t\{\log M_{t+1}^*\} - \text{Var}_t\{\log M_{t+1}\}\right]$$

Clearly if conditional variances are constants, then so is the risk premium.

Cumulants. Another aside: as discussed in Backus, Foresi and Telmer (JF, 2001) with more general assumptions about the distribution of the SDFs, the risk premium contains all the *higher order cumulants.*

The (conditional) cumulant generating function of a random variable X can be written as the log of the moment generating function when the latter exists

$$g_t(z) = \log \mathbb{E}_t\{\exp(zX)\} = \sum_{j=1}^{\infty} \kappa_{j,t} \frac{z^j}{j!} = \kappa_{1,t} z + \kappa_{2,t} \frac{z^2}{2!} + \kappa_{3,t} \frac{z^3}{3!} \dots$$

The coefficients $\kappa_{j,t}$ of this expansion are the conditional cumulants. The jth cumulant can be recovered by taking the jth derivative of $g_t(z)$ with respect to z and setting the derivative to zero. So for example

$$g_t'(z) = \kappa_{1,t} + \kappa_{2,t} z + \kappa_{3,t} \frac{z^2}{2} + \dots \Rightarrow g_t'(0) = \kappa_{1,t}$$

The cumulants are closely related to moments. In fact staring at the expansion for a short while reveals $\kappa_{1,t}$ is the conditional mean, $\kappa_{2,t} = \mu_{2,t}$, $\kappa_{3,t} = \mu_{3,t}$ and $\kappa_{4,t} = \mu_{4,t} - 3\mu_{2,t}^2$ (where $\mu_{j,t}$ is the jth conditional central moment).

Now let us take as our random variable $X = \log M_{t+1}$ so that the log conditional mean of the SDF can be written

$$g_t(1) = \log \mathbb{E}_t\{M_{t+1}\} = \sum_{j=1}^{\infty} \kappa_{j,t} \frac{1}{j!}$$

Since the log forward premium is $f_t - e_t = i_t - i_t^* = \log \mathbb{E}_t\{M_{t+1}^*\} - \log \mathbb{E}_t\{M_{t+1}\}$ we have

$$f_t - e_t = \sum_{j=1}^{\infty} (\kappa_{j,t}^* - \kappa_{j,t}) \frac{1}{j!}$$

Since the expected depreciation rate and the expectations of the logs are linked by $\mathbb{E}_t\{\log M_{t+1}^*\} = \mathbb{E}_t\{\log M_{t+1}\} + \mathbb{E}_t\{\Delta e_{t+1}\}$ we can conclude that the risk premium is the sum of the differences in higher order cumulants $j \geq 2$, specifically

$$p_t = \sum_{j=2}^{\infty} (\kappa_{j,t}^* - \kappa_{j,t}) \frac{1}{j!}$$

Implications for monetary policy. As discussed above, if nominal exchange rates are approximately a random walk, $\mathbb{E}_t\{\Delta e_{t+1}\} \approx a$ and

$$i_t - i_t^* \approx p_t + a$$

as discussed above. Therefore changes in the stance of monetary policy go hand-in-hand with changes in the risk premium p_t. Alvarez, Atkeson and Kehoe (AER, 2007) argue that this has important implications for models of monetary policy. They characterize the typical model as assuming constant conditional variances, say, that the nominal SDF is lognormal with time-varying mean but constant conditional variance, so that

$$i_t = -\log \mathbb{E}_t\{M_{t+1}\} = -\mathbb{E}_t\{\log M_{t+1}\} + \text{constant}$$

Debates about the neutrality of money in such models then reduce to debates about how a change in i_t is decomposed into changes in real interest rates and expected inflation. But if exchange rates are a random walk, then we know that across countries the interest differential $i_t - i_t^*$ is approximately the risk premium p_t plus a constant. So if we take the foreign nominal interest rate as given (say, a constant), then

$$i_t = p_t + \text{different constant}$$

Alvarez, Atkeson and Kehoe therefore characterize the typical debate in monetary economics as meaningless because it is arguing about terms that are 'essentially constant' (p 342). Interest changes should show up in changes in risk premia but this channel of monetary policy is ruled out by assumption in a model with constant conditional variances (or in a log-linearized model).

Causality? Without specifying how monetary policy is introduced into a model, Alvarez, Atkeson and Kehoe cannot say anything about the direction of causality. Is it that interest rate changes cause changes in the risk premium? Or does the monetary policy react to exogenous changes in the risk premium? Based on data from several recent episodes in US and UK monetary policy, they argue that the causality is from i_t to p_t, i.e., that monetary policy decisions are (part of) the source of time-varying risk.

For more on this and related topics see their paper in this year's NBER macro annual.

Chris Edmond

April 2008