RETURN FORECASTS AND TIME-VARYING RISK PREMIUMS

. . .

John H. Cochrane

The most obvious early efficiency tests examined whether returns on a given security are predictable over time:

$$R^e_{t+1} = a + bx_t + \varepsilon_{t+1}$$

where R^e_{t+1} denotes the excess return on a security, and x_t is any variable that investors might use at time t to forecast returns.

Tests using daily, weekly, or monthly return data found coefficients b that are small, with tiny R^2 values. Thus even some statistically "significant" results were judged economically insignificant. A coin with 51%/49% probabilities is darn close to fair, surely, it would seem, within transactions costs.

In the mid-1970s, Gene started looking at long-run return forecasts and, perhaps more importantly, at forecasts using prices as right-hand variables. Lo and behold, you *can* forecast returns using prices, and these forecasts are particularly striking at longer horizons. The b coefficients are economically large, and the R^2 values rise to impressive values.

Figure 1 illustrates this phenomenon.

The dashed line is the dividend/price ratio of the CRSP value-weighted portfolio. Think of it as prices upside down. The dividend/price ratio goes down in the big price booms, such as the 1960s and 1990s, and goes up in the big busts, such as the 1970s. It also wiggles with business cycles. Among other things, this graph points out the astounding volatility of stock valuations, which Bob Shiller shared the Nobel Prize in part for pointing out.

The solid line is the average return for the seven *following* years. So, times of high prices relative to dividends are reliably followed by seven years of low returns. Times of low prices relative to dividends are reliably followed by seven years of high returns.

The graph is my way of illustrating Fama and French's "Dividend Yields and Expected Stock Returns" regressions,

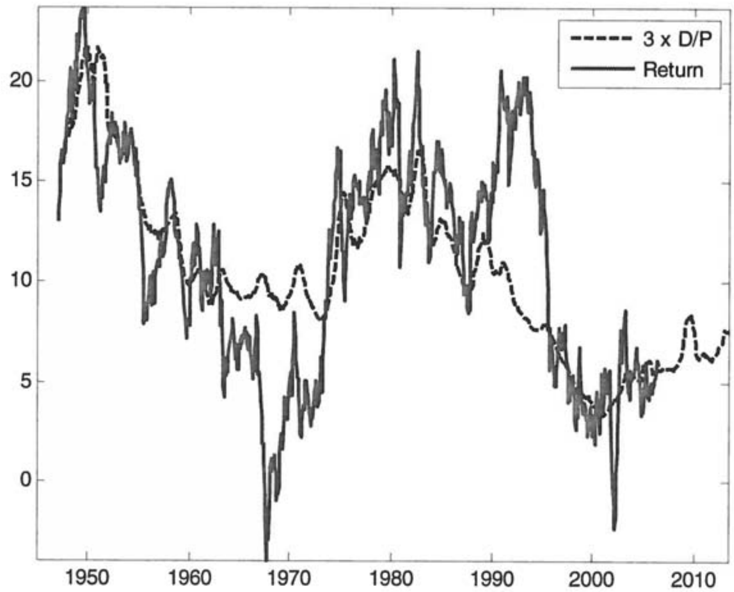

Figure 1. Dividend yield of value-weighted NYSE stock portfolio and following seven-year return

$$R^e_{t,t+7} = a + b\left(\frac{D_t}{P_t}\right) + \varepsilon_{t+7}. \tag{1}$$

These regressions have large R^2 values, visible in the correlation between lines in my graph. By this and other measures, long-run returns can be predicted with economic and statistical significance.

What do we make of this finding? Such a regression means that expected returns vary through time,

$$E\left(R^e_{t,t+7}\middle|\frac{D_t}{P_t}\right) = a + b\left(\frac{D_t}{P_t}\right) \tag{2}$$

It all seems simple in retrospect, but once again the idea that one could just run such simple regressions to measure time-variation in expected returns was not at all obvious at the time. Gene's "life in finance" explains some of the muddy contemporary thinking on the topic. Contemporary thinking revolved around looking for "proxies" for expected return variation. Using ex post returns, as in the left-hand side of this regression, looked fishy from that perspective. Ex post returns contain expected returns plus a large unpredictable component, so they are poor proxies for expected returns. But the forecasting regression doesn't use a proxy concept, and the unpredictable part of returns is uncorrelated with dividend yields, so the regression is valid.

Some of this confusion, I think, comes down to information sets, which are still a source of much confusion. We can think about expected returns given investors' information, given all public information, or given the information in one particular study. (And, to do it properly, one should distinguish conditional expectation from linear projection, which I will gloss over in the interest of simplicity.) Each set is smaller than the previous one, and they are not the same.

When we run a regression like (1) and interpret the "expected returns" as in (2), we measure the expectation given the dividend yield only. We do not measure investors' "true" expectations, or even the expectations we could measure with multiple regressions.

But the measurement in (2) still tells us a lot. By the law of iterated expectations, it gives the expected value of agent's expectations, $E[E(R|A, B)|B] = E(R|B)$. Furthermore, the variation over time of expected returns documented by (1) and (2) is a lower bound for the variation in expected returns conditioned on larger information sets: $\text{var}[E(R|A,B)] \geq \text{var}[E(R|B)]$. Adding more variables to a regression always increases the R^2.

So, regressions such as (1) do not isolate "the" expected return. But they do inform us about investors' expected returns and provide a lower bound for the variation over time of investors' expected returns.

To this day, many studies "proxy" for investors' information, deriving implications that hold only when investors use exactly the same information that a study uses to forecast returns, and no more. The Fama forecasting regressions, by contrast, "condition down." We start with a theoretical statement that is true based on investors' information, then take conditional expectations on both sides to derive expressions that remain valid based on a subset of that information, such as dividend/price ratios. Investors will always have much more information than we can include in a study, so avoiding proxies is a great advance. In this way, Fama's return-forecasting regressions are the precursor to the first-order condition instrumental variable estimation technique introduced in Hansen (1982) and Hansen and Singleton (1982). They mirror the inversion mentioned in Ray Ball's essay, and the grouping procedure we mentioned in discussing mutual funds. Markets reflect all sorts of information that we will never see, and tests of market efficiency must respect that fact.

After a long controversy, I think it is fair to say that long-horizon regressions are most important for showing the *economic* rather than *statistical* significance of forecasting regressions. The number of nonoverlapping observations declines as the horizon lengthens, so larger standard errors make up for larger coefficients, and there is not really a huge statistical advantage either way.

But that observation does not make the finding any less important. Really, much of the second revolution in finance—predictability, value, and momentum—comes from looking at the same phenomena in different ways that reveal their previously overlooked economic significance, rather than finding new techniques that wring more statistical significance out of the data.

When we unite return regressions with the evolution of the right-hand variable,

$$R^e_{t+1} = a + bx_t + \varepsilon_{t+1}$$

$$x_{t+1} = a_x + \phi x_t + \varepsilon^x_{t+1}.$$

a highly persistent forecasting variable (i.e., ϕ near one) means that regression coefficients b rise with horizon, and R^2 values rise with horizon. So, consider a regression forecast that one might have dismissed in 1970 as having a small b with low R^2, statistically "significant" but on that basis "economically insignificant." Yet, if the forecasting variable is persistent, that regression is exactly equivalent to a long-run forecast with a large b and a large R^2. There is no separate fact at long horizons. The long-horizon regressions are just a consequence of short-horizon regressions and a persistent forecasting variable. But the long-horizon regressions let us see that the fact is economically significant after all.

A better way to demonstrate economic significance, I think with a lot of hindsight, is to compare the variation in expected returns to the level of expected returns, rather than to divide the variation of expected returns by the variance of actual returns in R^2. We will never perfectly forecast returns, so the latter is an unhelpful comparison.

In a regression $r_{t+1} = a + bx_t + \varepsilon_{t+1}$, variation in expected returns is $\sigma(E_t(r_{t+1})) = b\sigma(x_t)$. If stock returns have a 5% unconditional mean, and the conditional mean varies by $\sigma(E_t(r_{t+1})) = 4\%$, that's an economically *huge* variation in expected returns. Expected returns are 5% on average, yet sometimes 0% and sometimes 10%! Such numbers are typical of dividend-yield forecasts such as equation (1). But since stock returns vary by $\sigma(r_{t+1}) = 20\%$, the R^2 of this forecast is only $(4/20)^2 = 0.04$, which seems small.

Unlike R^2, this comparison of the volatility of expected returns with the level of expected returns is not strongly affected by horizon, since the standard deviation of expected returns and the level of expected returns both grow roughly linearly with horizon.

The Fama forecastability papers have something else in common: the right-hand variable contains a price, and there is usually only one right-hand variable. If we were just looking for large variation in expected returns, then any variable could go on the right, and the more the merrier, up to the danger of fishing. Why prices, and why just prices?

In one sense prices are a natural variable. If expected returns are abnormally high on a given date, prices will be low since dividends are discounted at a higher rate. Thus, low prices partially reveal to us the fact that investors have information of high expected returns. If the investors are rational, average returns will be higher following a "low" price. Using prices is a clever way to aggregate and help us see some of the widely dispersed information that investors see.

But the use of prices is deeper. These regressions are not really about documenting variation in expected returns, which we could do by putting anything on the right-hand side. *These regressions tell us why prices move.* They are one more brilliant Fama use of OLS regressions for unusual purposes.

In economics, when you run $y = a + bx + \varepsilon$, it's conventional to think of x as "causing" y. In return-forecasting regressions, causality goes the other way. A rise in risk premium or other event causes expected returns to rise. Higher expected returns means prices fall as they discount the same dividends more strongly. We then see, on average, higher actual returns following the lower price, which is the fact that drives the regression. It is like a regression of Saturday's weather on Friday's forecast. The weather causes the forecast, not the other way around.

Fama's return-forecasting regressions are here not so much to tell us about expected returns but to tell us about *how prices are formed*. They are about what explains variation in the *right*-hand variable, not the *left*-hand variable. We put the "cause" on the "wrong" side because forecast errors are orthogonal to forecasters, and orthogonality, not causality, is the ultimate arbiter of which side each variable should lie on.

The message of the regressions is that variation in expected returns—

variation in discount rates—is an important determinant of the variation in prices.

This interpretation became clearest as return regressions were integrated with volatility tests, the latter made famous by Robert Shiller (1981). In modern expression, volatility tests amount to the observation that regressions of long-run dividend growth on the dividend/price ratio have a coefficient close to zero. They "should," in a constant-expected-return, efficient-market world, have a strong negative coefficient: If prices are lower than current dividends, it should mean that investors think dividends will decline in the future.

An identity links these observations. A high price/dividend ratio must mean higher subsequent dividend growth, lower subsequent returns, or a perpetually rising price/dividend ratio (a "rational bubble," "violation of the transversality condition"). These three items must add up as a matter of arithmetic, by the definition of return. Therefore, the regression coefficient of long-run return, long-run dividend growth, and terminal dividend yield on initial dividend yields must add up to one. The question is, which one is it? The answer is that all variation in price/dividend ratios corresponds to expected returns—and none of it to expected dividend growth (confirming Shiller) or perpetually rising prices ("rational bubbles"). In this sense, the subsequent digestion of Fama and French's regressions has completely united volatility tests and return predictability regressions.

Does return predictability imply that markets are inefficient? No, or at least not necessarily. In 1970, Gene's joint-hypothesis theorem emphasized that you *can* get better returns by shouldering more risk, and the reward for bearing risk can vary over time and across assets. We don't see important forecasts at daily, weekly, or seasonal frequencies, where it would be pretty hard to concoct a theory of time-varying risk premiums. The significant return forecasts seem to line up with business cycles and larger movements in economic activity, just where one might expect variation in economic risk premiums. Of course, that's not proof—one needs to write down and check economic models of time-varying risk premiums. But it certainly is suggestive.

For example, in December 2008, prices fell and, by the regression, expected stock returns rose. In the risk-premium view, typical investors answered, "Yes, I see it's a buying opportunity. But stocks are still risky, and the economy is falling to pieces. I just can't take risks right now. I'm selling." Many university endowments did just that.

There is another possibility: perhaps people were irrationally optimistic in the booms, and irrationally pessimistic in the busts.

And a third, more recent, challenge: perhaps the institutional mechanics of financial intermediation cause variation in the risk premium. When leveraged hedge funds lose money, they sell. If not enough buyers are around, prices fall.

These views agree on the facts so far. So how do we tell them apart? Answer: we need "models of market equilibrium." We are not here to tell stories. We need economic models, psychological models, or institutional models that tie price and expected-return fluctuations to data, in a nontautological way. Gene proved in the 1970 joint hypothesis theorem that there is no test based only on asset prices that can distinguish these explanations. And constructing such models is exactly what a generation of researchers including myself do, which is a measure of the large influence of Fama's forecasting regressions.

The forecastability regressions radically changed our worldview about variation in prices. In the 1970s, we might have thought that variation in market-wide price/dividend or price/earnings ratios came from changing expectations of dividend growth, earnings growth, etc. A high price relative to current dividends means that people expect higher dividends in the future. The return-forecast regressions, together with the "complementary" dividend-growth regressions, mean that variations in the risk premium, rather than variation in expected cash flows, account *entirely* for the volatility of market-wide stock valuations. A high price relative to current dividends entirely means that returns will be lower in the future. (This is a simple, agreed-on fact, not an explanation. The disagreement concerns whether those lower returns represent rationally avoided risk or irrational expectations. I carefully use the word "account," not "cause," here.)

Our worldview changed from "variation in price/dividend ratios corresponds 100% to variation in cash flow expectations with constant expected returns" to "variation in price/dividend ratios corresponds 0% to variation in cash flow expectations and 100% to expected return variation." You can't ask for greater "economic" significance of the point estimates—though you can still argue with statistical significance, and a large literature does.

The fact that the risk premium accounts for *all* variation in valuations changes everything we do in finance and related fields from accounting to macroeconomics. And the fact that variation in risk premiums is so correlated with business cycles tells macroeconomics that recessions have a lot to do with the ability to bear risk, a feature largely missing in current macroeconomic modeling.

A few caveats, because these results are frequently misinterpreted. First, one must be clear about information sets. Dividend growth is unpredictable at annual horizons from dividend/price ratios. That does not mean that dividend growth is unpredictable from other variables. In fact, dividend growth does seem to be predictable about one year ahead, using variables other than dividend/price ratios as predictors. So don't jump from "dividend yields don't predict dividend growth" to "dividend growth is unpredictable!"

Second, one might think that when additional variables show higher dividend growth predictability, they must imply less expected return variability. But adding variables to a regression always increases R^2. When we add other variables, expected returns must vary *even more* than dividend yield regressions indicate. The resolution of this puzzle is that additional dividend growth and additional return forecastability must be perfectly correlated. Any variable that forecasts higher dividend growth, holding dividend yields constant, must *also* forecast higher returns. For example, at the bottom of a recession, current dividends are depressed. Expected dividend growth is high, but expected returns are unusually high as well. The cash flow and discount rates offset each other, leaving no additional effect on prices. Thus, the fact that dividend growth is forecastable means that expected return variation accounts for more than 100% of price/dividend ratio variation!

Third, the fact that variation in *valuations*—price/dividend or price/earnings ratios—are fully accounted for by expected return variation does not mean that variation in *prices* or *returns* have the same sources. Contemporaneous dividend growth shocks account for about half the variance in price changes and returns. Prices and dividends decline together, leaving price/dividend ratios unaffected.

Fourth, these facts hold for time-series variation in the market as a whole. Cross-sectional variation in dividend/price, price/earnings, and book-to-market ratios seems to come about half from variation in expected cash flows and half from variation in expected returns. However, these decompositions are still somewhat overlooked, and a much better job of quantifying them is possible.

We start this section of the volume, however, with the much earlier "Short-Term Interest Rates as Predictors of Inflation." This was the first paper in the series, and it inaugurated the technique of regressing ex post values on prices to see how prices are determined.

This paper investigates the Fisher relationship, which states that the nominal interest rate should equal the real interest rate plus the expected rate of inflation,

$$i_t = r_t + E(\pi_{t+1}|I_t) \tag{3}$$

where I_t denotes the investor's information set. As Gene explains, previous efforts to examine this relationship looked at proxies for expected inflation.

Gene ran it backwards, with the price determined by the market i_t on the right-hand side:

$$\pi_{t+1} = r + bi_t + cx_t + \varepsilon_{t+1} \tag{4}$$

If the real rate is constant—the needed "model of market equilibrium"—then the coefficient b on the interest rate in (4) should be one, and, more importantly, the coefficient on c multiplying any other variable x_t in investors' information sets should be zero. The nominal interest rate should be a sufficient statistic, capturing all available information about future inflation. The nominal interest rate reveals the slice of investors' rich information sets that is useful for forecasting inflation.

Gene's paper found these predictions held quite well in the data sample available at that time. Gene was also sensitive to what we now call "weak instruments." The most interesting sorts of tests are x variables that forecast inflation well in univariate regressions, but should be driven out by the nominal rate i in a multiple regression. Inflation must be forecastable for this test to have any bite. Fortunately, since inflation is persistent, past inflation rates serve quite well in this purpose. Past inflation forecasts future inflation, but interest rates forecast inflation better.

This simple article set off much of the subsequent investigation. For example, Hansen and Singleton (1982) wrote the relation between real and nominal interest rates as

$$1 = E\left[\beta \frac{u'(c_{t+1})}{u'(c_t)}(1+i_t)\frac{\Pi_t}{\Pi_{t+1}}\middle| I_t\right],$$

using the consumption-based model as the "model of market equilibrium." Writing their instruments as x_t, they tested the relation

$$0 = E\left[\left(\beta \frac{u'(c_{t+1})}{u'(c_t)}(1+i_t)\frac{\Pi_t}{\Pi_{t+1}} - 1\right) \times x_t\right].$$

This moment condition is the numerator of a Fama-like regression coefficient. That's how we've done asset pricing ever since.

"Forward Rates as Predictors of Future Spot Rates" really leads off the series of papers by finding stunning expected-return variation, and it investigates the term structure of interest rates.

We included "The Information in Long-Maturity Forward Rates," with Rob Bliss, once again as a paper that came late in the series, less famous and less well cited, but benefiting from nearly a decade of refinement and a much clearer first paper for a modern reader.

The log forward rate $f_t^{(n)}$ is the rate at which you can contract today t to borrow from $t + n - 1$ to $t + n$. The log spot rate is the rate at which you can contract at t to borrow from t to $t + 1$. The simple expectations hypothesis states that the forward rate equals the expected future spot rate, perhaps plus a constant risk premium,

$$f_t^{(n)} = E\left(y_{t+n-1}^{(1)} \,|\, I_t\right) + c.$$

Following the standard Fama idea of running the regression backwards, Fama and Bliss run

$$y_{t+n-1}^{(1)} - y_t^{(1)} = a_y + b_y\left[f_t^{(n)} - y_t^{(1)}\right] + \varepsilon_{t+n-1}^{y}$$

and they check for the "efficient market" prediction $b_y = 1$.

Subtracting today's spot rate $y_t^{(1)}$ on both sides is important. If you simply report today's weather as your forecast for tomorrow's weather, over the course of the year you will look like a pretty good weather forecaster. The coefficient in a regression of tomorrow's weather on your forecast will be one, with a very high R^2. To reveal the emptiness of this forecast, we run the regression of the *change* in weather from today to tomorrow on the *difference* between your forecast and today's weather. This is an important step because weather, like forward rates, is highly serially correlated.

The result of this regression is profoundly unsettling. Tables 1 and 2 of "Forward Rates as Predictors of Future Spot Rates" imply that ". . . the martingale model, which simply predicts that the interest rate will remain unchanged, does better in predicting future spot rates than forward rates" (p. 363). Crystallized in "The Information in Long-Maturity Forward Rates," the coefficient is not just a bit less than one. It is *zero*. The forward rate does, essentially, report today's weather as its forecast for tomorrow.

As for dividend yields, there is an identity at work here: the forward-spot spread must correspond to a change in yield or to an ex post excess return:

$$\left[y_{t+n-1}^{(1)} - y_t^{(1)}\right] + \left[p_{t+n-1}^{(1)} - p_t^{(n)} + p_t^{(n-1)}\right] = \left[f_t^{(n)} - y_t^{(1)}\right].$$

The term in the middle is the return from holding an n-period bond from time t to time $t + n - 1$, financed by holding an n–1 period bond for the same period. Running both sides of this identity on the forward spread, you will find $b_r + b_y = 1$ —exactly, not in expectation, in each sample—where b_r is the coefficient in a regression of excess returns on forward spreads,

$$\left[p_{t+n-1}^{(1)} - p_t^{(n)} + p_t^{(n-1)}\right] = a_r + b_r\left[f_t^{(n)} - y_t^{(1)}\right] + \varepsilon_{t+n-1}^{r}$$

The forward spread must, mechanically, reflect variation in risk premiums, or variation in expected spot-rate changes. If forward rates do not forecast changes in spot rates, then they correspond to risk premiums.

Comparing "Forward Rates as Predictors of Future Spot Rates" with "The Information in Long-Maturity Forward Rates," we can see Gene's thinking evolve and we see the message become clearer and clearer. The first paper doesn't actually have these regressions. It only interprets tables of variances and autocorrelations in terms that the regressions would later clarify.

In the first paper, you see Gene really trying hard to salvage the view that a lot of forward rate variation comes from variation in expected future spot rates. Noting that forward rates on their own aren't doing a good job, he states that risk premiums merely "obscure" the desired forecast power.

> Any variation over time in the expected premiums . . . tends to obscure the power of the forward rate . . . as a predictor of the spot rate.

He then constructs a model of expected return variation. "After some experimentation" he uses the "average of absolute values of the monthly changes in the spot rate during the year before month t + 1 and during the year following the month t + 1" (p. 366) to measure volatility, and posits expected returns are a linear function of this volatility, an idea that disappeared from later work. The bottom line sounds awfully comforting:

> When forward rates are adjusted for variation through time in expected premiums, they provide predictors of future spot rates as good as those obtained from the information in the time series of spot rates (p. 365).

> . . . The market reacts appropriately. . . . This evidence is consistent with the market efficiency proposition (p. 361 [abstract]).

> . . . This evidence is consistent with the market efficiency proposition that in setting bill prices, the market correctly uses at least the information in past spot prices. However, the best support for market efficiency is the direct evidence that in setting bill prices and forward rates the market reacts appropriately to the negative autocorrelation in monthly change in the spot rate and to changes through time in the degrees of this autocorrelation (p. 377 [summary]).

There is some hedging, but you don't come away from this with the idea that all variation in forward spreads comes from expected returns, and none from variation in expected interest rate changes.

"The Information in Long-Maturity Forward Rates," by contrast, cuts to the chase and leads off in Table 1 with return forecasting regressions. At a one year horizon, $b_r = 1$ and $b_y = 0$. End of story. The idea of proxying expected returns by volatility is gone. Similar summary quotes have a dramatically different flavor.

> Current 1-year forward rates on 1- to 5-year U.S. Treasury bonds are information about the current term structure of 1-year expected returns on the bonds (p. 680 [abstract]).

> We confirm that forward rate forecasts of near-term changes in interest rates are poor (p. 680).

> The slopes [b_r] in the term-premium regressions range from 0.91 to 1.42. All are within one standard error of 1.0. We can infer that the slopes (equal to 1-[b_r] . . .) in the complementary yield-change regression . . . are within one standard error of 0.0. The results suggest that . . . variation in current [forward-spot] spreads is mostly variation in the term premiums in current 1-year expected returns, and forward-spot spreads do not predict yield changes 1 year ahead (p. 684).

> . . . 1-year expected returns for U.S. Treasury maturities to 5 years, measured net of the interest rate on a 1-year bond, vary through time. . . . This variation in expected term premiums seems to be related to the business cycles (p. 689 [conclusion]).

Some of the difference in language reflects the change of emphasis from levels to differences across time and maturity. But much of the change simply reflects the experience of a whole slew of papers, and over and over again seeing price movements that correspond 100% to variation in expected returns.

"The Information in Long-Maturity Forward Rates," however, shows that at long horizons, the expectations hypothesis actually does work rather well. Five-year forward rates *do* correspond to expected changes in interest rates four years from now, and not to the corresponding risk premium. And Fama and Bliss relate this pattern to the clear cyclical behavior of interest rates. This term structure of risk premiums is now a new and exciting area of research, both in empirical work and in theory.

"Forward and Spot Exchange Rates" is worth reading for several reasons. Coming in between the last two papers, it shows the development of the identity linking the things that "should" be forecastable—yields, dividends, exchange rates—to the thing that "should not" be forecastable—excess returns. It started an immense literature and established one of the handful of founding facts that define international finance. And in doing so, it is a reminder of how pervasive the patterns are across many markets.

A forward exchange rate f_t contract is an obligation to buy foreign currency one period in the future at a set price. The spot exchange rate s_t is the price of the corresponding immediate purchase. Thus, the simple expectations hypothesis predicts that the forward rate is today's expectation of the future spot rate,

$$f_t = E(s_{t+1}|I_t) + c$$

where again I have allowed for a constant risk premium c. Following the usual Fama idea to run the regression backwards, we check if forward rates are set this way by running the change in the spot rate on the spread between today's forward rate and today's spot rate,

$$s_{t+1} - s_t = a_s + b_s(f_t - s_t) + \varepsilon^s_{t+1},$$

and checking for $b_s = 1$. In words, if the forward rate is higher than the spot rate, and expected returns are constant, we should see, on average, spot rates subsequently rise.

A trader who enters into a forward contract and then sells in the spot market earns an excess log return equal to $f_t - s_{t+1}$. Thus, the identity

$$\left[s_{t+1} - s_t\right] + \left[f_t - s_{t+1}\right] = f_t - s_t$$

says that the forward-spot spread must correspond to a change in spot rate or to an excess return. And running excess returns on the forward spot spread,

$$f_t - s_{t+1} = a_r + b_r(f_t - s_t) + \varepsilon^r_{t+1},$$

we once again have an identity $b_r + b_s = 1$. So a forward-spot spread must forecast the change in spot rate or it must forecast the excess return. The only question is which. (Does this sound repetitive? That's the point! We're seeing the same pattern over and over again.)

Here, Gene found an even more remarkable result. Not only do we not see $b_s = 1$, we see *negative* values of b_s! Correspondingly, b_r is even larger.

If forward rates are higher than spot rates—equivalently, if the foreign interest rate is lower than the domestic rate—it looks like an investment opportunity. Buy forward, or borrow abroad. Exchange rate changes "should" offset this profit opportunity. In fact, they go the "wrong" way and enhance the profit opportunity!

Again, there is a suggestive macroeconomic correlation. Domestic interest rates are low relative to foreign rates in the bottoms of recessions. These are times when all risk premiums are large, so perhaps it makes sense that the risk premium for holding exchange rate risk is large as well. A large number of macroeconomic models have been constructed following this insight to explain this exchange rate premium.

This remains a live field of research. Even the basic facts are not yet completely explored. Fama looked at each country in isolation. We are only starting to see work that ties countries together, considers the difference between time and country fixed effects in these regressions, and examines what factors in exchange rate risk correspond to these expected returns. We are only starting to link currency risk premiums, bond risk premiums, and stock risk premiums. (Lustig, Roussanov, and Verdelhan [2011] stand out in this effort. They form portfolios of countries based on forward-spot exchange rate spreads, they document the average returns of these portfolios, and they show those average returns line up with a slope factor. In doing so, they at last bring different countries together in the analysis. Their watershed paper amounts, however, essentially to implementing Fama and French's "Common Risk Factors in Returns on Stocks and Bonds" [1993] and "Multifactor Explanations of Asset Pricing Anomalies" [1996] techniques to the well-known currency puzzle. That this took 15 years suggests how much low-hanging fruit Fama has left on the trees.)

The patterns shown in these three examples—stocks, bonds, and foreign exchange—are pervasive in financial markets, extending to commodities, corporate bond spreads, real estate, and other assets. In each case, time-series variation in a price/x ratio should, if expected returns are constant over time, forecast changes in x. In each case, it doesn't at all. Instead it forecasts excess returns. Corporate spreads largely forecast returns to bondholders, not larger defaults. Price/rent ratios forecast returns to homeowners, not changes in rent. In each case the result is pretty dramatic, typically a 100%/0% split that turns out to be 0%/100% or more. In each case the risk premium varies slowly over time and is suggestively correlated with business cycles.

Furthermore, there are strong suggestions of common movement across asset classes. The variables that forecast stock returns also forecast bond returns, and the variables that forecast bond returns also forecast stock returns. The business cycle association, multivariate return forecasts, and common component are explored a bit in Fama and French (1989). Yet they have really only scratched the surface.

REFERENCES

Fama, Eugene F., and Kenneth R. French. 1989. "Business Conditions and Expected Returns on Stocks and Bonds." *Journal of Financial Economics* 25, 23–49.

Hansen, Lars Peter. 1982. "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica* 50 (4), 1029–1054. doi: 10.2307/1912775.

Hansen, Lars Peter, and Kenneth J. Singleton. 1982. "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models." *Econometrica* 50 (5), 1269–1286. doi: 10.2307/1911873.

Lustig, Hanno, Nikolai Roussanov, and Adrien Verdelhan. 2011. "Common Risk Factors in Currency Markets." *Review of Financial Studies* 24 (11), 3731–3777. doi: 10.1093/rfs/hhr068.

Shiller, Robert J. 1981. "Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends?" *American Economic Review* 71 (3), 421–436.