Sensible return forecasting for portfolio management
Gregory Connor
Financial Analysts Journal; Sep/Oct 1997; 53, 5; ABI/INFORM Global
pg. 44

Sensible Return Forecasting for Portfolio Management

Gregory Connor

Black and Litterman showed that a Bayesian adjustment to expected-return forecasts makes them more suitable for use in portfolio management. A new adjustment applies directly to return-forecasting models rather than to the forecasts they produce. This approach eliminates the need for an arbitrary adjustment when forecasts are inserted into a portfolio-choice model and integrates the return-forecasting and portfolio-choice steps of quantitative investment management.

In their simplest formulation, mean–variance portfolio-choice models perform very badly because they produce unbalanced "optimal" portfolios. Michaud (1989) showed that this result is attributable to the misuse of noisy forecasts as substitutes for true expected returns. The use of noisy forecasts gives rise to the *error-maximization problem*: A mean–variance optimizer tends to severely overweight those securities with positive estimation errors in their expected-return forecasts and severely underweight those with negative estimation errors.

Black and Litterman (1990) solved the error-maximization problem in an elegant way that works well in practice. They imposed a prior distribution on expected returns based on the efficient markets hypothesis and used that distribution to derive Bayesian-adjusted expected return forecasts.

In a common application, an investor's expected returns come from an estimated return-forecasting model. A simple and intuitive procedure can be used to adjust return forecasts in this case. Like the Black–Litterman model, this approach relies on the efficient market hypothesis. Unlike the Black–Litterman model, however, the adjustment applies to the forecasting model *coefficients* rather than to the forecasts the model produces. This new approach has the advantage of giving more-precise guidelines on the size of the appropriate adjustment. Also, it links the forecasting-model-estimation stage to the portfolio-choice stage, rather than treating them separately as in the original Black–Litterman model. The new adjustment is simple and easy to apply.

THE ONE-VARIABLE CASE

The new technique is particularly simple to apply in the case of a single lagged variable that is used to forecast returns on an asset (or asset class). Let r_t denote the Time t return on the asset and x_{t-1} denote a variable observable at Time $t-1$ that the investor believes can be used linearly to predict the Time t return. Let $\tilde{r}_t$ and $\tilde{x}_t$ denote the asset return and predictive variables minus their means. A typical estimated return-forecasting model assumes that returns respond linearly to the predictive variable; that is,

$$\tilde{r}_t = b\tilde{x}_{t-1} + \varepsilon_t, \tag{1}$$

It also assumes that the unexplained return, ε_t, is normally distributed with mean zero and variance σ_ε.

Let the estimation period for the return-forecasting model be from 1 to T, let σ_x denote the variance of x, and let $\hat{\sigma}_{xr}$ represent the sample covariance between x and r. The formula for the ordinary least squares estimate of b is $\hat{b}^{ols} = \hat{\sigma}_{xr}/\sigma_x$. Given the de-meaned value of the explanatory variable at the end of the estimation period, $\tilde{x}_T$, $\hat{b}^{ols}$ can be used as follows to forecast the expected return of the asset next period compared with its long-run expected return:

$$\hat{r}_{T+1} = \hat{b}^{ols}\tilde{x}_T. \tag{2}$$

If a portfolio manager applies forecasts from Equation 2 to a set of assets and then performs a mean–variance portfolio optimization, he or she will encounter the error-maximization problem. Although each asset's forecast is unbiased when viewed alone, the larger positive forecasts from the *set* of asset return forecasts will be biased upward. Similarly, the larger negative forecasts from the *set* of asset forecasts will be biased downward. The reason for the bias is that the portfolio manager has not included in the forecasts his or her prior knowledge that the true predictable deviations of returns from long-run expected returns tend to be distributed around zero. Therefore, the purported "optimal" portfolio will actually be extreme in its overweighting and underweighting of various assets.

Gregory Connor is Director of Research, Europe, at BARRA International and a visiting fellow at the London School of Economics.

The error-maximization problem can be solved in this situation by using the efficient markets hypothesis to set prior values on the forecasting model. Note that b equal to zero is equivalent to weak-form market efficiency because it implies that the forecasting model has no ability to predict returns. Therefore, the natural prior expected value for b is zero. It is also convenient to assume that b is normally distributed. Together with the assumption that ε_t is normal, this formulation transforms Equation 1 into a standard Bayesian regression problem with the following well-known solution (see Chow 1983):

$$\hat{b}^{bayes} = \left[\frac{T\sigma_x/\sigma_\varepsilon}{(T\sigma_x/\sigma_\varepsilon) + (1/\sigma_b)}\right]\hat{b}^{ols}, \tag{3}$$

where $\hat{b}^{ols}$ is the ordinary least squares estimate of b and σ_b is the prior variance of the unknown b.

To apply Equation 3, it is necessary to determine σ_b. This determination can be made using a transformed version of the R^2 of the forecasting model, which will be called *expected forecasting power* and denoted by ρ. Viewing the R^2 of the forecasting model as a random variable (because it depends upon the unknown true coefficient b) and taking the expectation of the ratio $R^2/(1 - R^2)$ gives the expected forecasting power of the model; that is,

$$\rho = E\left(\frac{R^2}{1 - R^2}\right).$$

The value of ρ is positive and less than 1; in practice, ρ is close in value to expected R^2. The determination of ρ will be discussed in more detail when we apply the model to actual data. As shown in the appendix, substituting the expected forecasting power ρ into Equation 3 and simplifying gives the following easy-to-apply version of the Bayesian-adjusted coefficient estimator:

$$\hat{b}^{bayes} = \left[\frac{T}{T + (1/\rho)}\right]\hat{b}^{ols}. \tag{4}$$

The Bayesian-adjusted coefficient on the left-hand side of Equation 4 equals the unadjusted ordinary least squares coefficient on the right-hand side multiplied by the quantity $T/(T + 1/\rho)$, which is greater than zero and less than 1. The "shrinkage factor," $T/(T + 1/\rho)$, pushes the Bayesian-adjusted coefficient toward zero, which is the prior expected value of the true (unobservable) coefficient.

Table 1 shows the magnitude of the shrinkage factor for typical values of T and ρ in monthly-return-forecasting models. In a return-forecasting model with one variable, an expected forecasting power of 1 percent is not particularly low. At this level of expected forecasting power, the predicted coefficient estimated over a two-year sample period should be shrunk to less than one-fifth of its unadjusted value! Even with a five-year estimation period, the shrinkage factors are quite small. The numbers in this table should instill an appropriate degree of humility in a researcher developing return-forecasting models based on relatively short sample periods. The sensible researcher will scale back the ordinary least squares estimated coefficient of such a model substantially before using it to choose a portfolio.

Table 1. Shrinkage Factors for Typical Monthly Sample Sizes and Expected Forecasting Powers with One Explanatory Variable

Expected Forecasting Power	Shrinkage Factor by Months in the Estimation Period			
	24	48	60	120
0.50%	0.11	0.19	0.23	0.38
0.75%	0.15	0.26	0.31	0.47
1.00%	0.19	0.32	0.38	0.55
2.00%	0.32	0.49	0.55	0.71
3.00%	0.42	0.59	0.64	0.78

THE MULTIVARIATE CASE

Extending this procedure to the case of a multivariate forecasting model is straightforward. Letting $\tilde{X}_t$ denote the k vector of explanatory variables minus their means and B denote the k vector of linear coefficients, the forecasting model is

$$\tilde{r}_t = B'\tilde{X}_{t-1} + \varepsilon_t. \tag{5}$$

Using the same approach as in the one-variable case, we let B have a multivariate normal prior distribution with zero means. As in the one-variable case, the assumption of zero means follows from the efficient markets hypothesis. Analogous to the one-variable case, the prior variances of B can be inferred from expected forecasting power. These variances, however, now depend upon the expected *marginal* forecasting power of each variable, which is similar to the marginal R^2 of the variable. Let R^2_j denote the marginal R^2 of explanatory variable j in Equation 5 (see the appendix for the definition of marginal R^2). Define the expected marginal forecasting power of the jth explanatory variable, ρ_j, from the equation $\rho_j = E[R^2_j/(1 - R^2)]$. As shown in the appendix, inserting ρ_j into the standard formula gives the following simple expression for the Bayesian regression coefficient j:

$$B_j^{bayes} = \left[\frac{T}{T + (1/\rho_j)}\right]\hat{B}_j^{ols},$$

where $\hat{B}_j^{ols}$ is the ordinary least squares estimate of the jth coefficient.

Table 2 shows the magnitude of the shrinkage factor $T/(T + 1/\rho_j)$ for typical sample sizes and expected forecasting powers in a model with four explanatory variables. The results shown are for total expected forecasting powers, ρ's, of 0.5 percent, 2 percent, and 5 percent, divided among the four explanatory variables in proportions 1/2, 1/4, 1/6, 1/12. By construction, the sum of the ρ_j's always equals the total ρ (note that the sum of the fractions equals 1). Table 2 makes clear that the methodology places a premium on model parsimony; explanatory variables with low expected marginal forecasting power have shrinkage factors close to zero.

APPLICATION TO INTERNATIONAL EQUITY ALLOCATION

To apply the described technique to dynamic asset allocation across international equity indexes, we estimated a return-forecasting model, generated out-of-sample expected-return forecasts, and used a mean–variance optimizer to choose a portfolio.

The return-forecasting model is estimated on the 48-month period from August 1991 to July 1995 for each of eight equity markets: Canada, France, Germany, Japan, the Netherlands, Switzerland, the United Kingdom, and the United States. The return series used is the local return to the Morgan Stanley Capital International equity index for each national market, minus the local risk-free rate. The forecasting model has the following form:

$$\tilde{r}_t = b_1\tilde{d}_{t-1} + b_2\tilde{y}_{t-1} + b_3\tilde{r}_{t-1} + \varepsilon_t,$$

where $\tilde{d}_{t-1}$ is the dividend yield on the national equity index, $\tilde{y}_{t-1}$ is the yield on long-term bonds minus the short-term rate, and $\tilde{r}_{t-1}$ is the lagged excess return to the national equity index (all the variables are de-meaned). The ordinary least squares estimation results are summarized in Table 3. We show both the in-sample adjusted R^2 and an out-of-sample R^2, which has better small-sample properties.[1]

To apply the Bayesian adjustment technique, we needed to set the expected marginal forecasting power, ρ_j, for each of the three regression coefficients in each of the eight markets. These values should reflect our prior beliefs about the likely fit of the models. This step is the most difficult in Bayesian estimation because one's prior beliefs about statistical models are difficult to state precisely.

The first, obvious step is to assume that all eight markets have the same expected total forecasting power, ρ. This assumption is justified by the symmetry of the model across markets; there is no a priori reason to expect one of these markets to have more predictability than another. Using this assumption, a sensible choice for ρ is $R^2/(1 - R^2)$, using the average of R^2 across the eight estimated models. We used the average out-of-sample R^2 in preference to the in-sample adjusted R^2. To complete the determination of expected marginal forecasting power, ρ_j, we needed to decide how to apportion the total expected forecasting power across the three explanatory variables (we assumed that the portions are the same across the eight markets). This apportionment was based on our a priori beliefs about the relative forecasting power of the three variables.

The dividend yield has the strongest a priori justification because there are good theoretical explanations for why dividend yields can predict returns. The bond yield variable also has some

Table 2. Shrinkage Factors for Typical Monthly Sample Size and Expected Forecasting Powers with Four Explanatory Variables

			Shrinkage Factor by Months in Estimation Period			
Expected Total Forecasting Power	Explanatory Variable	Expected Marginal Forecasting Power	24	48	60	120
0.5%	1	0.25%	0.0566	0.1071	0.1304	0.2308
	2	0.13	0.0291	0.0566	0.0698	0.1304
	3	0.08	0.0196	0.0385	0.0476	0.0909
	4	0.04	0.0099	0.0196	0.0244	0.0476
2.0%	1	1.00	0.1935	0.3243	0.3750	0.5455
	2	0.50	0.1071	0.1935	0.2308	0.3750
	3	0.33	0.0741	0.1379	0.1667	0.2857
	4	0.17	0.0385	0.0741	0.0909	0.1667
5.0%	1	2.50	0.3750	0.5455	0.6000	0.7500
	2	1.25	0.2308	0.3750	0.4286	0.6000
	3	0.83	0.1667	0.2857	0.3333	0.5000
	4	0.42	0.0909	0.1667	0.2000	0.3333

Table 3. Ordinary Least Squares Estimates for the Return-Forecasting Model

Country	Dividend Yield Coefficient	Bond Yield Coefficient	Lagged Return Coefficient	In-Sample Adjusted R^2	Out-of-Sample R^2
Canada	−0.0041	−0.0051	−0.2777	0.0168	−0.0860
France	0.0977	−0.0134	−0.0900	0.1300	0.1820
Germany	0.0322	−0.0248	−0.1272	0.0707	0.2050
Japan	0.1116	−0.0110	−0.0168	0.0157	−0.0900
Netherlands	0.0244	−0.0187	−0.2346	0.1000	−0.2200
Switzerland	0.0057	−0.0095	−0.0149	−0.0413	0.0260
United Kingdom	0.0635	−0.0200	0.0566	0.0321	0.3730
United States	0.0016	0.0032	−0.3326	0.0587	−0.3380
Average	0.0416	−0.0124	−0.1297	0.0478	0.0065

theoretical justification. The lagged return seems the weakest in its a priori justification.[2] Based on this comparison, we apportioned the total expected forecasting power across the three variables in ratios of 0.5, 0.3, and 0.2 for the dividend yield, bond yield, and lagged return variables, respectively.[3] Table 4 displays the values of ρ_j for each variable and the Bayesian-adjusted coefficients for all eight markets.

For simplicity, we constrained the chosen portfolio to be fully hedged against currency risk. We used a pound sterling perspective and set the unconditional expected returns of the assets (i.e., the expected returns without any forecasting model) so that the optimized portfolio using unconditional expected returns would be the benchmark portfolio.[4] The benchmark portfolio was set equal to the capitalization-weighted portfolio of the eight equity indexes, fully hedged against currency risk. We used the capitalization weights at the end of July 1995 and held those weights fixed throughout the forecasting/portfolio-optimization period. The risk-free rate was set to the U.K. rate in November 1995 and held fixed throughout the optimization period. The covariance matrix for the eight equity excess returns was set equal to the sample covariance matrix over the 48-month estimation period.

The estimated return-forecasting models were used to create one-month-ahead return forecasts for each of the eight equity indexes each month during the 12-month period from August 1995 to July 1996. The return forecast at time t equals the unconditional expected return plus $\hat{B}'\tilde{X}_{t-1}$. For comparison purposes, the return forecasts are shown in Table 5, which uses both the Bayesian-adjusted estimates of the coefficients and the ordinary least squares estimates. Also shown are the unconditional expected returns, which do not change from one month to the next. Note that the return forecasts from the adjusted model do not stray too far from the unconditional returns, whereas those using the ordinary least squares coefficients sometimes differ dramatically from the unconditional values. Two of the predictive variables (dividend yield and bond yield) change slowly through time, so the return forecasts within each market also tend to change slowly over the 12-month forecasting period.

Table 6 shows the optimal portfolios each month based on the return forecasts, using the unadjusted ordinary least squares coefficients and using the Bayesian-adjusted coefficients. The portfolio weightings based on the unadjusted forecasting model are extremely aggressive. In two months (January 1996 and July 1996), the chosen portfolio is entirely concentrated in a single asset; in most

Table 4. Bayesian-Adjusted Coefficient Estimates for the Return-Forecasting Model

Statistic	Dividend Yield	Bond Yield	Lagged Return
Expected marginal forecasting power	0.0033	0.0020	0.0013
Shrinkage factors	0.1349	0.0856	0.0587
Adjusted coefficients			
Canada	−0.0006	−0.0004	−0.0164
France	0.0132	−0.0012	−0.0053
Germany	0.0044	−0.0021	−0.0075
Japan	0.0151	−0.0009	−0.0010
Netherlands	0.0033	−0.0016	−0.0138
Switzerland	0.0008	−0.0008	−0.0009
United Kingdom	0.0086	−0.0017	0.0033
United States	0.0002	0.0003	−0.0196

Table 5. Return Forecasts Using the Adjusted and Unadjusted Models, Currency-Hedged One-Month Equity Return Forecasts
(based on pound sterling)

Date	Canada	France	Germany	Japan	Netherlands	Switzerland	United Kingdom	United States
Unconditional expected returns	0.0089	0.0100	0.0085	0.0137	0.0097	0.0087	0.0102	0.0088
Unadjusted return forecasts								
1995								
August	0.0102	0.0245	0.0152	0.0394	–0.0014	0.0152	0.0125	–0.0007
September	0.0210	0.0258	0.0423	0.0327	0.0111	0.0159	0.0126	0.0111
October	0.0168	0.0360	0.0218	0.0352	0.0098	0.0180	0.0110	–0.0047
November	0.0187	0.0473	0.0257	0.0369	0.0205	0.0199	0.0012	0.0097
December	0.0065	0.0416	0.0236	0.0379	0.0067	0.0227	0.0140	–0.0054
1996								
January	0.0157	0.0427	0.0244	0.0281	0.0118	0.0219	0.0107	0.0033
February	0.0029	0.0243	0.0187	0.0231	0.0090	0.0188	0.0072	–0.0037
March	0.0210	0.0266	0.0181	0.0203	0.0088	0.0184	–0.0028	0.0050
April	0.0136	0.0320	0.0162	0.0262	0.0040	0.0179	0.0089	0.0066
May	0.0055	0.0075	0.0189	0.0167	0.0011	0.0196	0.0002	0.0058
June	0.0113	0.0187	0.0190	0.0217	–0.0006	0.0169	0.0010	0.0025
July	0.0256	0.0180	0.0140	0.0199	0.0060	0.0172	0.0042	0.0100
Adjusted return forecasts								
1995								
August	0.0092	0.0119	0.0091	0.0163	0.0088	0.0092	0.0102	0.0081
September	0.0098	0.0117	0.0121	0.0154	0.0094	0.0093	0.0103	0.0089
October	0.0096	0.0129	0.0088	0.0154	0.0094	0.0094	0.0099	0.0079
November	0.0097	0.0149	0.0093	0.0157	0.0102	0.0096	0.0088	0.0087
December	0.0091	0.0139	0.0091	0.0158	0.0095	0.0098	0.0101	0.0078
1996								
January	0.0096	0.0139	0.0088	0.0147	0.0096	0.0097	0.0095	0.0083
February	0.0089	0.0117	0.0084	0.0140	0.0092	0.0095	0.0090	0.0079
March	0.0099	0.0114	0.0079	0.0137	0.0088	0.0094	0.0081	0.0084
April	0.0095	0.0124	0.0080	0.0144	0.0088	0.0094	0.0099	0.0086
May	0.0090	0.0092	0.0082	0.0133	0.0084	0.0095	0.0085	0.0085
June	0.0094	0.0103	0.0083	0.0138	0.0082	0.0093	0.0089	0.0083
July	0.0102	0.0104	0.0079	0.0136	0.0085	0.0094	0.0093	0.0088

months, it holds positive positions in only two assets. With the adjusted coefficients, the portfolio strategy is still quite active, but not nearly as extreme. The portfolio usually has nonzero weights on five of the eight assets. Table 7 shows the forecast returns for each optimal portfolio each month and the realized returns of these portfolios. Note the large average forecast error of the optimal portfolio using the unadjusted return-forecasting model, which helps explain the aggressive portfolio weights in Table 6: Using unadjusted forecasts, the investor is overly optimistic about potential outperformance and so implements an aggressive strategy (see Michaud).

CONCLUSION

The Bayesian adjustment for estimated return-forecasting models, presented here, is based on a prior hypothesis of efficient markets. The adjustment applies directly to the estimated return-forecasting model rather than to the forecasts the model produces. It has the effect of shrinking the estimated coefficients of the forecasting model toward zero, which is their expected value given the prior hypothesis of market efficiency.

This new approach has two advantages relative to the earlier work of Black and Litterman: It ties together two stages of quantitative investment management, return forecasting and portfolio construction, and it gives specific guidance on the size of the adjustment to use, based on the expected fit of the return-forecasting model. This technique performs well in an empirical application to international equity market allocation.[5]

APPENDIX

The model presented in this paper is based on standard Bayesian regression (see, e.g., Chow 1983).

Table 6. Optimal Portfolios Based on Adjusted and Unadjusted Return Forecasts

Date	Canada	France	Germany	Japan	Netherlands	Switzerland	United Kingdom	United States
Unconditional expected returns	0.0251	0.0416	0.0467	0.2750	0.0264	0.0350	0.1098	0.4404
Unadjusted return forecasts								
1995								
August	0.0000	0.0245	0.000	0.770	0.0000	0.000	0.0000	0.0000
September	0.0000	0.0258	0.957	0.044	0.0000	0.000	0.0000	0.0000
October	0.0000	0.0360	0.000	0.317	0.0000	0.000	0.0000	0.0000
November	0.0000	0.0473	0.000	0.042	0.0000	0.000	0.0000	0.0000
December	0.0000	0.0416	0.000	0.236	0.0000	0.000	0.0000	0.0000
1996								
January	0.0000	0.0427	0.000	0.000	0.0000	0.000	0.0000	0.0000
February	0.0000	0.0243	0.000	0.293	0.0000	0.060	0.0000	0.0000
March	0.3352	0.0266	0.000	0.038	0.0000	0.000	0.0000	0.0000
April	0.0000	0.0320	0.000	0.172	0.0000	0.000	0.0000	0.0000
May	0.0000	0.0075	0.277	0.138	0.0000	0.585	0.0000	0.0000
June	0.0000	0.0187	0.410	0.344	0.0000	0.149	0.0000	0.0000
July	0.0000	0.0180	0.000	0.000	0.0000	0.000	0.0000	0.0000
Adjusted return forecasts								
1995								
August	0.024	0.294	0.000	0.376	0.000	0.090	0.000	0.216
September	0.019	0.041	0.380	0.318	0.000	0.000	0.000	0.243
October	0.151	0.390	0.000	0.324	0.000	0.051	0.000	0.084
November	0.000	0.534	0.000	0.318	0.000	0.000	0.000	0.149
December	0.000	0.472	0.000	0.355	0.000	0.101	0.000	0.073
1996								
January	0.103	0.467	0.000	0.284	0.000	0.021	0.000	0.125
February	0.045	0.279	0.000	0.296	0.000	0.163	0.000	0.218
March	0.250	0.229	0.000	0.241	0.000	0.073	0.000	0.207
April	0.085	0.331	0.000	0.284	0.000	0.047	0.000	0.253
May	0.051	0.000	0.027	0.268	0.000	0.249	0.000	0.405
June	0.148	0.130	0.000	0.271	0.000	0.152	0.000	0.300
July	0.286	0.112	0.000	0.230	0.000	0.085	0.000	0.287

Table 7. Forecast Returns and Realized Returns of the Chosen Portfolios
(based on pound sterling)

	Unadjusted Return-Forecasting Model		Adjusted Return-Forecasting Model		Unconditional Return Forecasting	
Date	Forecast Return	Realized Return	Forecast Return	Realized Return	Forecast Return	Realized Return
1995						
August	0.036	0.045	0.012	0.020	0.010	0.015
September	0.042	–0.020	0.012	0.004	0.010	0.020
October	0.038	–0.002	0.013	–0.005	0.010	–0.009
November	0.047	0.010	0.014	0.028	0.010	0.042
December	0.041	0.030	0.014	0.035	0.010	0.027
1996						
January	0.043	0.082	0.013	0.055	0.010	0.032
February	0.024	–0.011	0.011	–0.006	0.010	–0.007
March	0.025	0.021	0.011	0.028	0.010	0.021
April	0.031	0.043	0.012	0.031	0.010	0.023
May	0.019	–0.017	0.010	–0.002	0.010	0.002
June	0.020	0.021	0.010	0.009	0.010	0.007
July	0.026	–0.028	0.011	–0.051	0.010	–0.050
Average	0.032	0.015	0.012	0.012	0.010	0.010
Average forecast error		0.01772		–0.00022		0.00018

The distinctive features of the model are the two assumptions used to determine the prior distribution of the regression coefficients—in particular:

- The prior expected values of the regression coefficients are set to zero, based on an efficient markets hypothesis.
- The prior variances of the regression coefficients are inferred from the model's expected forecasting power.

A Single Explanatory Variable

In the single-variable case, the return-forecasting model is

$$\tilde{r}_t = b\tilde{x}_{t-1} + \varepsilon_t. \quad \text{(A1)}$$

For a known b, the R^2 of the regression model shown in A1 is defined as the proportion of explained variance to total variance, $R^2 = b^2\sigma_x/\sigma_r$. Using $\sigma_r = b^2\sigma_x + \sigma_\varepsilon$ and rearranging gives $R^2/(1 - R^2) = b^2\sigma_x/\sigma_\varepsilon$. Taking the expectation over the random realizations of b gives $\rho = E[R^2/(1 - R^2)] = \sigma_b\sigma_x/\sigma_\varepsilon$. Substituting this formulation into the equation for the Bayesian coefficient estimate, Equation 2, gives

$$\hat{b}^{bayes} = \left[\frac{T}{T + (1/\rho)}\right]\hat{b}^{ols},$$

which is the formula shown in the text.

Multiple Explanatory Variables

With multiple explanatory variables, the return-prediction equation becomes

$$\tilde{r}_t = B'\tilde{X}_{t-1} + \varepsilon_t. \quad \text{(A2)}$$

As in the one-variable case, the coefficients B are treated as normally distributed random variables with zero means.

Let Σ_X denote the sample covariance matrix of the explanatory variables X, and let Σ_B denote the covariance matrix of the prior distribution for B. We assume that Σ_B can be expressed in the form $\Sigma_B = \Sigma_X^{-1/2}\, D\, \Sigma_X^{-1/2}$ for some diagonal matrix, D. The formula for the Bayesian regression coefficient estimates (see Note 3) is

$$\hat{B}^{bayes} = \left[(T/\sigma_\varepsilon)\Sigma_X + \Sigma_B^{-1}\right]^{-1}(T/\sigma_\varepsilon)\Sigma_X\hat{B}^{ols}. \quad \text{(A3)}$$

The R^2 of this model can be written as $R^2 = (1/\sigma_r)B'\Sigma_X B$. Note that we can re-express $B'\Sigma_X B$ as trace$(\Sigma_X^{1/2}BB'\Sigma_X^{1/2})$, where trace(•) is the trace operator. Taking expectations with respect to B gives $E[R^2] = (1/\sigma_r)$ trace(D).

Define the marginal R^2 of explanatory variable j as the full R^2 of the regression minus the R^2 when variable j is dropped from the model: $R_j^2 = R^2 - (1/\sigma_r)B_{(j)}'\Sigma_X B_{(j)}$, where $B_{(j)}$ is the vector of coefficients with the jth entry set to zero. We can also write $B_{(j)} = I_{(j)}B$, where $I_{(j)}$ is the $k \times k$ identity matrix, except with the jth diagonal component set to zero. Taking expectations of $R_j^2/(1 - R^2)$ with respect to B and simplifying gives $E[R_j^2/(1 - R^2)] = D_j/\sigma_\varepsilon$. Substituting this equation into the Bayesian regression formula (A3) and simplifying gives

$$\hat{B}_j^{bayes} = \left[\frac{T}{T + (1/\rho_j)}\right]\hat{B}_j^{ols},$$

which is the formula shown in the text.

NOTES

1. In small samples, this type of return-forecasting model has an upward bias in adjusted R^2. See Goetzmann and Jorion (1993) for a discussion. To eliminate this bias, we estimated the model on a 48-month period and then generated out-of-sample R^2 for the previous 12-month period. We used a pre-estimation period to generate the out-of-sample R^2 rather than a postestimation period so that we could save the postestimation period for the portfolio optimization tests.
2. See, for example, Fama and French (1989) for an argument that dividend yields and bond yields have theoretical validity as equity return-prediction variables.
3. Readers who are unfamiliar with Bayesian methods may object to the somewhat arbitrary assignment of these particular weights to the three variables. A less arbitrary alternative is to use equal weights of (1/3, 1/3, 1/3); this apportionment corresponds to uninformative prior beliefs about the relative power of the three explanatory variables. Someone committed to the Bayesian philosophy, however, uses whatever prior information is available, even if it is unfocused. This practice led me to use the particular weights (0.5, 0.3, 0.2) to reflect my prior expectations about the relative power of the three explanatory variables.
4. The unconditional expected excess return on each asset is set equal to its beta against the benchmark times the unconditional expected excess return of the benchmark I used the sample mean excess return of the benchmark for the unconditional expected excess return of the benchmark. The mean–variance objective function maximized is $E[r_x] + \lambda\text{Var}[r_x]$, where $E[r_x]$ and $\text{Var}[r_x]$ are the mean and variance of the excess return of the portfolio x. We used a value of $\lambda = 3.194$, which is chosen so that the unconditional return forecasts result in the benchmark portfolio being optimal. See Kahn, Roulet, and Tajbakhsh (1996) for a detailed description and justification of this approach.
5. I would like to thank Stan Beckers, Anthony Bercel, Lucie Chaumeton, Peter Muller, and Shanthi Nair for helpful comments, and especially Ross Curds for extensive research collaboration.

REFERENCES

Black, F., and R. Litterman. 1990. "Asset Allocation: Combining Investor Views with Market Equilibrium." Goldman Sachs Fixed Income Research Discussion Paper (September).

Chow, G.C. 1983. *Econometrics*. New York:McGraw-Hill.

Goetzmann, W.N., and P. Jorion. 1993. "Testing the Predictive Power of Dividend Yields." *Journal of Finance*, vol. 48, no. 2 (June):663–80.

Fama, E.F., and K.R. French. 1989. "Business Conditions and Expected Returns on Stocks and Bonds." *Journal of Financial Economics*, vol. 25, no. 1 (November):23–49.

Kahn, R.F., J. Roulet, and S. Tajbakhsh. 1996. "Three Steps to Global Asset Allocation." *Journal of Portfolio Management*, vol. 23, no. 1 (Fall):23–32.

Michaud, R.O. 1989. "The Markowitz Optimization Enigma: Is 'Optimized' Optimal?" *Financial Analysts Journal*, vol. 45, no. 1 (January/February):31–42.