American Statistical Association

1994 Proceedings
of the
Section on Bayesian Statistical Science

Papers presented at the
Annual Meeting of the American Statistical Association,
Toronto, Canada, August 13-18, 1994

ASA 1994 Winter Conference,
Atlanta, Georgia, January 6-9, 1994

Second North American Regional Meeting
of the International Society for Bayesian Analysis (ISBA),
Toronto, Canada, August 12-13, 1994

Second World Meeting of the International Society
for Bayesian Analysis,
Alicante, Spain, June 10-11, 1994

all under the sponsorship of the
Section on Bayesian Statistical Science

American Statistical Association □ 1429 Duke Street, Alexandria, VA 22314
NEW BAYESIAN STATISTICAL APPROACHES TO ESTIMATING AND EVALUATING MODELS OF EXCHANGE RATES DETERMINATION

Bluford H. Putnam and Jose Mario Quintana
Jose Mario Quintana, Bankers Trust Company, One Bankers Trust Plaza, New York, NY 10006

Key Words: Bayesian Forecasting, Bayesian Shrinkage, Exchange Rates, Portfolio Management

When one talks to currency dealers, it is not unusual to hear discussions of how the Sterling is weak, or the Yen is strong, or the Canadian Dollar is weak, with the intent to place the focus on one particular country, yet exchange rates are relative values between two currencies. Who is to say if the Canadian Dollar depreciates against the US Dollar that the problem lies more with the Canadian weakness than the US strength? In this research study, we present two new concepts for modeling exchange rates. First, we introduce a statistical technique for separating country effects in the analysis of exchange rates. Second, this technique requires that we analyze several exchange rates simultaneously, as in a portfolio performance criterion for evaluating models rather than traditional statistical tests. Both of these approaches are attempts to come down from the ivory towers of academia and to test and to learn from the ideas of currency traders in the dealing rooms. Ironically, this statistical formulation facilitates, via Bayesian shrinkage, the introduction of hypothesis postulated by economic theories of exchange rates. The impact of this input, in terms of portfolio performance, is quite promising.

The outline of the paper is as follows: Section 1 considers the intrinsic relative nature of exchange rates and motivates the use of multivariate models whose structure is invariant to a change in perspective. Section 2 introduces a tailored relative version of seemingly unrelated regressions (SUR) for exchange rate modeling that conforms with the concepts aforementioned. Section 3 discusses an analogue relative shrinkage formulation. Section 4 compares the exchange rate predictive ability of the SUR and shrinkage models in the context of portfolio performance and Section 5 summarizes the findings of this research.

1. RELATIVE EXCHANGE RATE MODELS

Understanding the dynamic path of an exchange rates between two currencies virtually requires a framework of analysis based on relative comparisons between the two countries that underlie the particular exchange rate in question. Basic exchange rate arbitrage conditions such as purchasing power parity and interest rate parity embody this approach. In purchasing power parity the continuously compounded growth (the natural logarithm of the ratio of a exchange rate over the previous exchange rate) in of an exchange rate is related to the difference in the (continuously compounded) inflation rates of the two countries being considered (Bilson, 1984). With interest rate parity, the relationship between the (logarithms of the) spot and forward exchange rates are related to the (continuously compounded) interest rate differential between the two countries (Levich, 1985). Exchange rate models based on a monetary analysis of inflation argue that the (logarithmic) change in an exchange rate is partly a function of the relative tightness (or ease) of monetary policy in the two countries (Putnam and Wilford, 1986). Trade theorists look to relative export and import performance to explain movements in exchange rates.

When an exchange rate theory is formulated based on comparisons between two similar independent variables in two different countries, issues are raised as to the most appropriate empirical methods to use when moving from theory to practice. One issue revolves around the question of whether modeling techniques should assume identical coefficients on similar variables or allow a divergence from the theory? For example, in the purchasing power parity case, one could postulate an equation,

$E_b z_k = \beta_b p_b - \beta_k p_k$ , \hspace{1cm} (1.1a)

where $E_b z_k$ equals the expected (logarithmic) change in an exchange rate for currency of country (k) in terms of the currency of country (b), and $p_k$, $p_b$ equal the (continuously compounded) inflation rates.

Alternatively, one could consider the following equation,

$E_b z_k = \beta (p_b - p_k)$ . \hspace{1cm} (1.1b)

In the first equation $\beta_b$ and $\beta_b$ do not necessarily have to be equal. In the second equation, $\beta_b$ is set equal to $\beta_k$ equal to $\beta$. The choice of the equation that is preferable depends on the purpose of the investigation. If one is testing the hypothesis that $\beta_b$ equals $\beta_k$ in a traditional sense, then the first equation, (1.1a) is going to be preferred. If one is concerned with producing exchange rate forecasts to be used in making
investment decisions in financial markets, then the decision as to which equation is preferred is much more controversial.

The returns from investing in exchange rates are highly volatile and contain a considerable amount of noise when attempts are made to explain exchange rate movements with fundamental economic variables. The low signal to noise ratio in the dependent variable (exchange rate returns) opens up wide possibilities for spurious correlation within any given estimation period. That is, the relationship among the dependent variable and the independent variables may not be robust for out of sample forecasting. The question then becomes whether the forecaster wants to put greater weight on empirical estimates from a noisy data series or to increase the weight placed on fundamental theory (Putnam, 1992).

In the particular case of exchange rates, this broad issue can be narrowed to the question of whether to enforce a strict version of relative independent variables (i.e., as in equation (1.1b)), or whether to allow more leeway for the estimation process (i.e., as in equation (1.1a)). In this paper, we provide a forecasting example which leans more heavily on the economic theory of exchange rates that argues for strict relativity in the formulation of the independent variables.

An extension of this issue of the appropriate use of theoretical ideas in an exchange rate forecasting context involves the choice of estimating each exchange rate equation separately or as a package. One set of exchange rate theories assume relatively free markets between countries in goods and capital, so that the exchange rate equation for each pair of countries is identical in terms of the set of independent variables. That is, the same exchange equation can be applied to any exchange rate given the critical assumptions about the free flow of goods and capital between the two countries. Where these types of theories are applied to forecasting exchange rates the issue then becomes whether or not to constrain coefficients across equations for similar independent variables.

For example, in the purchasing power parity case, one could be examining three different exchange rates.

\[ E_b z_j = \beta_b p_b - \beta_j p_j \]
\[ E_b z_k = \beta_b p_b - \beta_k p_k \]
\[ E_b z_m = \beta_b p_b - \beta_m p_m \]  

(1.2)

Should the process be constrained to force all of the \( \beta \) coefficients to be equal? This would be the answer from a strict interpretation of many exchange rate theories. Again, one needs to remember the problem that exchange rate data has a very low signal to noise ratio. Thus, there is a very large opportunity for spurious correlation in a given estimation period with little out of sample forecasting ability. In the case of the low signal to noise ratio, modeling the coefficients simultaneously and forcing equality on the coefficients for the relevant relative independent variables may give economic theory a much better chance to produce a profit-making forecasting model than single equation techniques with unconstrained coefficients. One can observe the conduct of successful currency traders and note that they analyze markets simultaneously, incrementally adding or subtracting positions from their portfolios, and that they do not behave as if each exchange rate is in a world all its own. In this paper, we investigate an approach that looks at exchange rates as a package and adheres to a strict interpretation for the coefficients of relative independent variables.

As a final point, exchange rates in theory can be decomposed into the forces pushing an exchange rate from each of the two countries involved. That is, since the independent variables are all relative to a counterpart in the other country, one can show whether it is the economic changes in one country versus the other that is pushing the exchange rate. Empirically, however, exchange rate analysis is usually conducted in a method that does not impose the symmetry implied by the relative nature of most exchange rate models. Therefore, in this paper we want to explore a statistical formulation that makes explicit the role of the base currency’s country versus the foreign currency’s country in the exchange rate forecasting process.

Our objective is to investigate an exchange rate forecasting system that embodies a strict view of the relative nature of exchange rate determination and takes us closer to economic theory in this regard. In addition, we want our forecasting approach to be consistent with a portfolio approach to investing. This requires that we use statistical techniques that treat a set of exchange rates as a package and not as independent equations. This should bring our quantitative approach more into line with our observations concerning what makes for successful investors and particularly currency traders. Finally, our models are evaluated using risk-return criteria rather than statistical criteria, since our goal is to produce attractive risk adjusted profits rather than to prove or disprove a given economic theory. We now turn to the development of our empirical approach, given these objectives and criteria.
2. RELATIVE DYNAMIC SEEMINGLY UNRELATED REGRESSIONS

Consider a set of auxiliary, not observable, set of random dependent variables representing respectively the strength (weakness) of a set of currencies. The relative dynamic SUR model for exchange rate determination makes the following assumptions: First, the continuously compounded rate of growth of a cash deposit in one currency relative to a cash deposit in another currency is the difference of the corresponding strengths. Second, the strengths follow a Dynamic SUR model (Putnam and Quintana, 1993). The rest of this section formalizes these concepts.

The continuously compounded rate of growth of a cash deposit in currency (k) relative to a cash deposit in currency (b) from time \( t-1 \) to time \( t \), denoted by \( y_{\text{kt}} \), is defined as,
\[
\log \left( \frac{S_{\text{kt}}(1+i_{\text{kt}})}{S_{\text{kt-1}}(1+i_{\text{kt-1}})} \right) = y_{\text{kt}} = \log \left( \frac{S_{\text{kt}}}{F_{\text{kt}}(1+i_{\text{kt}})} \right),
\]
where \( S_{\text{kt}} \) and \( S_{\text{kt-1}} \) the spot prices of one unit of currency (k) in terms of the currency (b) at times \( t-1 \) and \( t \) respectively, \( i_{\text{kt}} \) and \( i_{\text{kt-1}} \) are the interest rates for the corresponding period, and \( F_{\text{kt}}(1+i_{\text{kt}}) \) is the fair forward price of a unit of currency (k) at time \( t-1 \) to be delivered at time \( t \).

The model first assumption states,
\[
y_t = H_t y_t^*, \quad H_t = \begin{bmatrix} 0 & 0 \\ -1 & I \end{bmatrix},
\]
where \( y_t = (y_{t1}, \ldots, y_{tn}) \) is a \( (r \times 1) \) vector of observations, \( H_t \) is a \( (r \times r) \) matrix, \( 0 = (0, \ldots, 0) \) is a \( ((r-1) \times 1) \) zero vector, \( y_t^* \) is a \( (r \times 1) \) strength vector, \( 1 = (1, \ldots, 1) \) is a \( ((r-1) \times 1) \) unit vector. The continuously compounded rates in (2.2) are arbitrarily defined, without lost of generality, relative to the first currency.

The model second assumption asserts,
\[
y_t^* = X_t \Theta_t + e_t^*, \quad e_t^* \sim N(0, V_t \Sigma_t^2),
\]
\[
\Theta_t = G_t \Theta_{t-1} + f_t, \quad f_t \sim N(0, W_t \Sigma_t^2),
\]
\[
\sigma_t^2 \sim \chi^2 (\beta_t, s_{t-1}, \beta_t, d_{t-1}), \quad \text{and}
\]
\[
\theta_{t-1} | \sigma_{t-1}^2 \sim N(m_{t-1}, C_{t-1}, \sigma_{t-1}^2),
\]
where \( X_t \) is a \( (r \times p) \) block diagonal matrix of independent variables, \( \Theta_t \) is a \( (p \times 1) \) vector of system regression parameters, \( e_t^* \) is a \( (r \times 1) \) error vector, \( V_t \) is a \( (r \times r) \) variance matrix associated with \( e_t^* \), \( G_t \) is a \( (p \times p) \) evolution (trend) matrix, \( f_t \) is a \( (p \times 1) \) evolution noise vector, \( W_t \) is a \( (p \times p) \) variance matrix associated with \( f_t \), \( \sigma_t^2 \) is a system scale factor, and \( \beta_t \) is a discount factor associated to the system scale factor. The structure of the block diagonal matrix of independent variables is,
\[
X_t^* = \begin{bmatrix}
X_{t1}^* & 0 & \cdots & 0 \\
0 & X_{t2}^* & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & X_{tn}^*
\end{bmatrix}
\]
where the diagonal elements are typically country specific factor vectors that presumably explain their corresponding currency strengths. For details on the implementation of the Dynamic SUR model described by the equations (2.3) see Putnam and Quintana (1993).

The model defined by the set of equations (2.2) and (2.3) can be reformulated, due to the superposition principle, as,
\[
y_t = X_t \Theta_t + e_t, \quad e_t \sim N(0, V_t \Sigma_t^2),
\]
\[
\Theta_t = G_t \Theta_{t-1} + f_t, \quad f_t \sim N(0, W_t \Sigma_t^2),
\]
\[
\sigma_t^2 \sim \chi^2 (\beta_t, s_{t-1}, \beta_t, d_{t-1}), \quad \text{and}
\]
\[
\theta_{t-1} | \sigma_{t-1}^2 \sim N(m_{t-1}, C_{t-1}, \sigma_{t-1}^2),
\]
where \( X_t = H_t X_t^*, \quad e_t = H_t e_t^*, \quad \text{and} \quad V_t = H_t V_t \Sigma_t H_t^* \). The model (2.4) is a slight variation of the general multivariate DLM (West and Harrison, 1989, p. 598) where the matrices of independent variables have a specific structure.

3. RELATIVE DYNAMIC SHRINKAGE

The benefits of using Bayesian shrinkage techniques to improve the forecast of financial assets, including currencies, has been pointed out by Dumas and Jacquillat (1990) and Jorion (1991) among others. In this study we extend the investigation in two substantial ways: First, the models are formulated in relative terms. Second, the shrinkage models employed are based on a, more general, modern Bayesian shrinkage approach, along the lines of Zellner, Hong, and Ming (1991), that has proved successful in global macroeconomic forecasting.

The Dynamic shrinkage model analogue to the dynamic SUR model of Section 2 is defined by the equations (2.2) and (2.3) but replacing the equation (2.3a) by,
\[ y_i^* = X_i^* \beta_i + e_i^* , \quad e_i^* \sim N(0, \Sigma_i^*) , \quad (3.1a) \]
\[ \beta_i = B_i \theta_i + d_i , \quad d_i \sim N(0, \sigma_i^2) . \quad (3.1b) \]
The superposition principle can be applied, as in Section 2, to rewrite the model as,
\[ y_i = X_i \beta_i + e_i , \quad e_i \sim N(0, \Sigma_i^*) , \quad (3.2a) \]
\[ \beta_i = B_i \theta_i + d_i , \quad d_i \sim N(0, \Sigma_i^*) . \quad (3.2b) \]
\[ \theta_i = G_i \theta_{i-1} + f_i , \quad f_i \sim N(0, \Sigma_i^*) , \quad (3.2c) \]
\[ \sigma_i^2 \sim \chi^2(\beta_i, s_{i-1}, \beta_i, d_{i-1}), \quad \text{and} \]
\[ \theta_{i-1} | \Sigma_{i-1} \sim N(M_{i-1}, C_{i-1} \sigma_{i-1}^2) , \quad (3.2d) \]
\[ \sigma_{i-1}^2 \sim \chi^2(s_{i-1}, d_{i-1}) . \]

Different settings of the link matrix \( B_i \) and the noise matrix \( U_i \) allow for a variety of shrinkage methods. Henceforth we focus on the straightforward case of common coefficients, in other words, the instance where equation (3.2b) becomes,
\[ \beta_i = \begin{bmatrix} \beta_{i1} \\ \vdots \\ \beta_{in} \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \theta_i . \quad (3.3) \]
The rationale behind this particular choice is the assumption, induced by economic theory, that the sensitivity of the strength of the currencies to each similar factor, say, the local short interest rates, is the same across different countries. Therefore, by reducing the number of regression coefficients the information coming from the observations is used more efficiently resulting in better predictive ability.

4. IMPLEMENTATION AND PERFORMANCE

The implementation of the Dynamic SUR model of Section 2 and the Dynamic shrinkage model of Section 3 require the specification of the variance matrices \( \Sigma_i \) and \( \Sigma_i^* \). We deal with the problem of specifying \( \Sigma_i \) by estimating it on-line in a manner similar to the traditional static Bayesian SUR model (Zellner, 1971, pp. 240-246); see Quintana and Putnam (1993) for a description of the process. To overcome the specification of \( \Sigma_i^* \) we followed the alternative method of discounting information advocated by West and Harrison (1989). A technical difficulty also arises when attempting to implement the dynamic SUR and the dynamic shrinkage models defined, respectively, by the set of equations (2.4) and (3.2). The reason is that they do not conform with the traditional DLM and hierarchical DLM models in the sense that the defining matrices are singular. There are several ways to overcome this problem, we use the sweep operator approach described in Quintana (1987, Chapter 4).

We used the following method to assess the performance of the currency models. First, we applied the two currency models of Section 2 and Section 3 to forecast one-month ahead the rate of growth for a set of currencies. Second, we constructed mean-variance efficient portfolios based on the models predictions. Third, we compared the simulated results in terms of risk-return characteristics for the testing period.

The US dollar, British Pound, Japanese Yen, Deutschemark, Swiss Franc, French Franc, Dutch Guilder, Australian Dollar, and Canadian Dollar defined the dependent variables, described by equations (2.1) and (2.2), for the period from January 1980 through December 1993. Notice that, by convention, all these rates were taken relative to the first currency (the US Dollar). The cumulative sum of these variables, is shown in Figure 4.1. Each time series represent, in semi-logarithmic form, the ratio of wealth of a foreign currency deposit over to a US Dollar deposit, per unit of capital. Notice that both the effect of the exchange rate and the effect of compounding interest rates are taking into consideration.

![Figure 4.1 Wealth of Foreign Currency Relative to the US Dollar](image)

Five local independent variables were used for each currency equation: Short Term (continuously compounded) Interest Rates, (a measure of) Change in Long Term Interest Rates, Yield Curve (the difference between Long and Short Interest Rates), (a measure of) Money Supply Acceleration, and Year-Over-Year Trade Balance Change. All the independent variables were lagged to insure that the information was available at the time when the forecast was made.

The one-month ahead forecast information of the models was used to buy and sell one-month forward contracts (i.e. to construct portfolios of forward
contracts). Forward currency contracts are agreements to exchange, in the future, fixed amounts of two currencies at prices set today. No money changes hands until the contract expires or is offset.

Mean-variance efficient portfolios with constant risk-aversion were constructed using the predictive mean and variance-covariance of the forward contract returns. This strategy aims to obtain attractive returns while controlling the risk. This process can be formally justified as a multi-period maximization of an additive quadratic utility function; full details are found in Quintana and Putnam (1994). The key ingredients of the process are the predictive moments of the forward contract returns. These were implied from the relationship,

$$b_{kt} = \frac{S_{kt}}{F_{k(t-1)}} - 1 = \exp(b_{y_{kt}}), \quad (4.1)$$

where the $b_{kt}$ denotes the return of a forward contract (the payoff resulting from buying forward currency (k) in exchange for one unit of currency (b) at the forward exchange rate $F_{k(t-1)}$).

The monthly simulations discussed next are all from the US investor viewpoint for the period from January 1988 through December 1993 (Eight years were used as an early training period). Figure 4.2 depict the performance of the Dynamic SUR and Dynamic Shrinkage models along with the performance of the (future contracts of the) Standard & Poor's 500 stock index (S&P500). To put these performances in perspective the risk-aversion parameter was set ex-post to match the historical variance of the S&P500.

![Figure 4.2 Model Driven Portfolio Performance](image)

Figure 4.2 Model Driven Portfolio Performance

It is apparent that both, the SUR and Shrinkage driven strategies easily outperform the S&P 500 (which is considered a highly efficient portfolio of US stocks). The performance of the model driven strategies is very similar from the start until the end of 1991, from that point forward the Shrinkage continues delivering strong returns whereas the SUR performance wanes in comparison.

The analysis of the model driven portfolio strategies presented so far are traditional in the sense that ignore transactions costs both, in the optimization process and in the simulated returns. It is well known that in this framework the portfolios often exhibit large allocations that are difficult to implement in practice due to high transaction costs. In addition, they may appear too risky to the seasoned practitioner ultimately responsible for the investment decision.

Figure 4.3 shows the effect of taking into account transaction costs both, in the construction of portfolios and in the calculation of the portfolio returns. For the optimization process we employed a quadratic penalty for taking large positions; see Quintana and Putnam (1994) for details. For the calculation of returns we assumed a flat one-way cost of five basis points (five percent of one percent) per forward contract.

![Figure 4.3 Cost Effect in Portfolio Performance](image)

Figure 4.3 Cost Effect in Portfolio Performance

Table 4.1 summarizes the performance of the different portfolio strategies in terms of monthly forward return statistics. It is important to keep in mind that forward returns do not require the investors to put any money down, in other words, do not require any initial capital; capital may be needed at the settlement date to cover for potential losses, but could

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P 500</td>
<td>0.66</td>
<td>3.47</td>
<td>0.19</td>
</tr>
<tr>
<td>SUR</td>
<td>1.48</td>
<td>3.47</td>
<td>0.43</td>
</tr>
<tr>
<td>Shrinkage</td>
<td>1.89</td>
<td>3.47</td>
<td>0.55</td>
</tr>
<tr>
<td>Shrinkage with Cost</td>
<td>1.72</td>
<td>3.30</td>
<td>0.52</td>
</tr>
</tbody>
</table>

Table 4.1 Model Driven Portfolio Monthly Forward Return Statistics
be invested, say in a cash deposit, and earn interest in the mean time. This assessment in terms of forward returns is typically used for strategies based on derivative instruments (such as forward or future contracts). However, the average reader presumably is more familiar with traditional investment programs where, say, one unit of capital is invested, and the proceeds are reinvested for the duration of the program. Figure 4.4 displays the cumulative wealth in this traditional terms taking into account not only the cash flows coming from the forward return, but also including the interest accrued by the capital.

![Figure 4.4 Cumulative Wealth per unit of Capital](image)

5. CONCLUSIONS

Our conclusions from these investigations are threefold. First, the Bayesian shrinkage technique is a promising avenue of research when applied to economic ideas that can be generalized across countries, such as exchange rate forecasting. Second, we believe that exchange rate theories are more efficiently implemented in a portfolio context and that generalized procedures are to be preferred over a set of disconnected single equations which are estimated independently. That is, investors think in terms of portfolios and so should applied financial statistical approaches. Finally, we feel the combination of Bayesian shrinkage and simultaneous equation estimation gives generalized economic theories an improved chance to produce risk-adjusted profit-making strategies. That is, this approach helps to reduce spurious correlation and the probability of rejecting an economic theory that is actually appropriate.

ACKNOWLEDGMENTS

We are grateful to Arnold Zellner for drawing the Shrinkage techniques to our attention, to Allen Hunt and Stella Kang for research assistance, to D. Sykes Wilford for reviewing the paper and providing many useful comments as part of the peer review process.

REFERENCES


