Journal of Multinational Financial Management
12 (2002) 107–133

Journal of
MULTINATIONAL
FINANCIAL
MANAGEMENT

www.elsevier.com/locate/econbase

Uncertainty of trading rules in currency markets: an application of non-parametric bootstrapping

Paolo Pasquariello *

Stern School of Business, Kaufman Management Education Center, New York University, Suite 9-180, 44 West 4th Street, New York, NY 10012-1126, USA

Received 15 April 2000; accepted 15 January 2001

Abstract

The problem of measuring the precision of signals generated by fundamental macroeconomic models is not trivial. In this paper, we suggest three different approaches for the estimation of the true and unknown distribution of the population signal. We apply the bootstrapping procedure described by Efron and Tibshirani (Stat. Sci. 1 (1986) 54) to estimate the empirical distribution of the signal and measure its precision at a specific point in time with confidence intervals. Direct and indirect bootstrapping methods are devised as a way to capture the unknown variability of the signal without altering the information content of the available data. This framework is then implemented for a simple fundamental model for the CAD/$ exchange rate. We find that accounting for skewness and prediction bias affects significantly the shape and width of the estimated confidence intervals around the estimated directional signal, and that the two proposed forms of bootstrapping are more satisfactorily than a naïve Historical approach in highlighting the uncertainty surrounding the model's predictions, and generating a measure of precision in the resulting model's recommendations. © 2002 Elsevier Science B.V. All rights reserved.

JEL classification: F31; F37

Keywords: Exchange rate forecasting; Bootstrapping; Trading rules

* Tel.: +1-212-998-0369; fax: +1-240-526-7168. PhD candidate at the Stern School of Business. The author composed this paper as a summer associate at J.P. Morgan. The views expressed in the paper are primarily theoretical and academic in nature and do not represent the views of J.P. Morgan or its affiliates. All opinions, theories and errors, as such, are purely those of the author.

E-mail address: ppasquar@stern.nyu.edu (P. Pasquariello).

PII: S1042-444X(01)00049-4

1. Introduction[1]

Trading rules based on fundamental economic relations are widely adopted by professional investment managers in the currency market. Deviations from equilibrium prices implied by the purchasing power parity (PPP), the covered interest rate parity (CIRP), and other country-specific models are exploited to generate extra profits with respect to specific ex-ante buy-and-hold strategies or benchmarks.

The ability of these strategies to generate excess returns consistently over significantly long periods of time is still an object of dispute in the literature.[2] Fundamental signal models (FSM), as we call them in this paper, can nevertheless offer important insights in understanding some of the economic trends that influence the behavior of exchange rates. In any case, the proper use of FSMs requires more statistical analysis than the one usually performed to test their effectiveness. Whether, after having observed the output of the model at a certain point in time, the analyst follows precisely the resulting positioning signal or simply takes it into consideration in forming his opinion about the future direction of the currency, the precision of that signal is going to play a relevant role in the decision process. How *confident* can the analyst be that the observed signal number is close to the *true* signal, the one generated by the *true* and unobservable multivariate distribution of the factors and variables entering his model?

A potential answer to such a question lies in the construction of confidence intervals for the signals generated by a FSM model. An example will help clarify this point. Assume that, at a certain date, a properly calibrated FSM is suggesting a short position, -5%, in a currency, with respect to a pre-specified benchmark. The analyst looks at the signal, evaluates independently other variables not entering the model framework, i.e., not fully captured by the signal itself,[3] and eventually takes a positioning decision. Throughout this process, the degree of precision of the signal generated by the FSM is certainly going to play a role for the analyst in comparing the model's suggestion with other inputs. Hence, a confidence band of $\pm 1.5\%$ around the observed -5%, for a selected significance level, will give the analyst a clearer picture of where the *true* and *unobservable* signal should be. However, consider the case in which the confidence band is of $\pm 9\%$. For the given significance level, the *true* signal could even be positive, with the observed -5% representing simply a random fluke.[4]

[1] I would like to thank Stephen Brown, New York University, and Arun Muralidhar, J.P. Morgan Investment Management, for their helpful comments and advices. The suggestions of an anonymous referee considerably improved the exposition of this paper. The views, opinions and all remaining errors in this article are naturally mine.

[2] For a review of the most recent tests of efficiency in the currency markets see Levich (1985, 1989) and Levich and Lee (1993).

[3] A classical example would be political instability, market momentum, or other non-systematic factors for the currency.

[4] While the analysis that follows focuses on the ex-ante signal uncertainty for models attempting to generate positive excess returns, some risk-management literature has been exploring the issue of determining ex-post how much of the observed excess return generated by active trading can be explained by skills (of the model and/or the analyst) rather than by noise. For more details, see Muralidhar and Mala Khin (2001).

For all its apparent importance, the problem of deriving confidence intervals for the signals generated by a FSM model is not trivial. If the FSM were linear in the factors and variables entering its formulation, assumptions about their true probability distributions would put us closer to a solution. Unfortunately, these distributions are themselves unknown, usually different from each other, and in most cases not easily approximated as normal.[5] Even if such an approximation were possible, it would not help the analyst for the case in which the factors and the variables enter the FSM in a nonlinear fashion. Moreover, for additive FSMs, i.e. for models averaging out different components, no '*cross-sectional*' version of the Central Limit Theorem applies, given the limited number of factors typically considered in a fundamental currency-positioning model.

The claim we make here is that the solution to this complicated issue lies in the application of a bootstrapping methodology. Bootstrapping, as we shall see in the remainder of the paper, substitutes considerable but easily manageable amounts of computation in place of theoretical analysis, when the latter is either impossible, as our case seems to suggest, or extraordinarily cumbersome. Bootstrapping involves drawing a certain number of samples from the empirical distributions of the variables entering the FSM. Each sample is called a *bootstrap sample*. Because the extractions happen from the empirical distribution of the data, a bootstrap sample turns out to be the same as a random sample of a specified size drawn with replacement from the actual sample.[6] The bootstrapping algorithm proceeds in three steps:

1. Using a random number generator, a large number of bootstrap samples is independently drawn for each of the variables entering the selected FSM;
2. The statistic of interest, e.g. the FSM signal, is then evaluated for each bootstrap sample;
3. The sample standard deviation of the statistic is calculated and confidence intervals are specified.

The paper is organized as follows. In Section 2, a simple structure for a general FSM is described. The proposed model does not pretend to exploit excess return opportunities in the currency market, but rather represents a general example of the class of linear models widely adopted by practitioners in currency management. Section 3 illustrates the specific non-parametric bootstrapping techniques suggested in the article. Section 4 shows an application of the techniques to the model of Section 2 using real market data. Section 5 concludes.

[5] Assumptions of IID Normality for the variables included in the model, although unrealistic, might still not help, for what follows in Section 2.

[6] The statistical analysis that follows relies on the fundamental contributions of Efron (1981), Efron and Gong (1983), Efron (1984), and Efron and Tibshirani (1986).

2. A simple linear FSM model

Currency-positioning models attempt to capture the effect of macroeconomic variables over the expected trend of the foreign exchange rate. The most frequently adopted macro-factors result from several (and fully established in the literature) equilibrium relationships between currency rates, price indexes, nominal and real interest rates, and flows of international trade.

We assume that the currency for which we want to construct a FSM model is the Canadian dollar versus the US dollar, from now on CAD.[7] CAD are thus units of Canadian dollar per one US dollar. The reasons why the CAD has been selected for our study are threefold. First, the relatively small size of the Canadian economy vis a vis the rest of the industrialized world makes the Canadian case a good proxy for the small country–big country Mundell–Fleming setting (see Mundell, 1961a,b; Fleming, 1962). This widely embraced model explains the relation between interest rate differentials and exchange rate fluctuations through the channel of international capital flows. Second, most of the Canadian trade with the rest of the world is characterized by an easily identifiable, and at the same time highly significant component, commodities. Raw materials are less related to sudden swings of tastes and preferences in the international markets than other tradable goods and services. This in turn, as it will be shown in this section, allows us to specify a trade factor in terms of deviations from the equilibrium in the commodities markets induced by the world business cycle. Third, the Canadian dollar shows a definite propensity to trend, hence it appears to represent, in our judgment, a well-suited object for the application of a fundamental macroeconomic model of exchange rate fluctuations. It is important to emphasize again that the model described in this section makes no attempt of originality and does not represent a *Holy Grail* for profiting in the currency markets. The following linear structure is simply an example of a FSM structure generating positioning signals for which we want to specify some useful statistical properties. We identify four main economic relations:

2.1. The purchasing power parity (PPP) factor

The *absolute* PPP paradigm asserts that in the long term the real exchange rate between Canadian dollars and US dollars has to be equal to unity. In other terms, this implies that $R = (P \cdot \text{CAD})/P^* = 1$, where P is the Canadian producer price index (PPI) and P^* is the US PPI. For the purposes of our analysis, in order to capture the long-term reversion of the real exchange rate toward its equilibrium, we instead focus on the following *adjusted* PPP variable:[8]

[7] However, our FSM model can be easily extended to any currency of interest.

[8] In short, the *adjusted* version of the PPP determines the equilibrium spot rate that would be necessary at time t to preserve the same level of real competitiveness existing at the time selected by the analyst (BASE), if a certain inflation differential had been observed between the BASE year and t. Consequently, the structure proposed in Eq. (1) signals an unfavorable trend for the CAD when the domestic inflation is relatively high and the Canadian dollar is expected to depreciate to restore previous levels of competitiveness, and vice-versa when the domestic inflation is relatively low.

$$\mathrm{PPP}_t = \frac{\mathrm{PPI}_t^{\mathrm{CAD}}/\mathrm{PPI}_{\mathrm{BASE}}^{\mathrm{CAD}}}{\mathrm{PPI}_t^{\mathrm{US}}/\mathrm{PPI}_{\mathrm{BASE}}^{\mathrm{US}}}\,\mathrm{CAD}_{\mathrm{BASE}}$$

$$F_{\mathrm{t}}^{\mathrm{PPP}} = \alpha\left(\ln\left(\frac{\mathrm{CAD}_t}{\mathrm{PPP}_t}\right) - \beta I_1 + \ln\left(\frac{\mathrm{CAD}_t}{\mathrm{PPP}_t}\right) + \beta I_2\right) \quad (1)$$

$$I_1 = \begin{cases} 1 & \text{if } \ln\left(\frac{\mathrm{CAD}_{\mathrm{t}}}{\mathrm{PPP}_{\mathrm{t}}}\right) \succ \beta \\ 0 & \text{otherwise} \end{cases}, \qquad I_2 = 1 - I_1$$

In Eq. (1) PPP_t is the equilibrium spot nominal exchange rate, BASE is the base month selected for the construction of the PPP factor, β is a real exchange barrier trigger, and α is a calibration parameter.[9] A positive value of the factor signals a favorable trend for CAD.

2.2. The flows of funds factor

This component is built on the assumption that, when the Canadian real interest rate is *unusually* lower than the US real interest rate, an expected *depreciation* of the Canadian currency is needed to balance outstanding and future flow of funds.[10] The factor translating this hypothesis into a trend signal is the following:

$$F_t^{\mathrm{REAL}} = \gamma((r_t^{\mathrm{CAD}} - r_t^{\mathrm{US}}) - (r_{\mathrm{Ht}}^{\mathrm{CAD}} - r_{\mathrm{Ht}}^{\mathrm{US}})) \quad (2)$$

Here r_t is simply a short-term moving average of the money-market real interest rate, calculated *a la Fisher* as a difference between the nominal cash rate and the corresponding ex-post inflation rate; r_{Ht} is a long-term moving average of money-market real interest rates, attempting to capture the equilibrium component.[11] Finally, γ is a calibration parameter selected by the analyst. A negative value for the

[9] As PPI data are released monthly, daily PPP values are inferred through interpolation. Different versions of the basic PPP hypothesis use different price variables, in particular the GDP Deflator, CPI or even *expected* CPI. We selected the PPI indicator as a good proxy for the relative competitiveness of the Canadian economy, given the relevant weight that Bank of Canada attributes to commodity prices in its formulation.

[10] Some literature, see Isard (1995) for a review, rejects the hypothesis that the covered interest rate parity (CIRP) is an adequate description of currency dynamics. In short, the CIRP assumes that flows of funds and exchange rates are immensely more rapid than sticky nominal interest rates in adjusting to interest rate differentials, and that markets are continuously and instantaneously in equilibrium. Hence, a *positive* spread between Canadian and US interest rates is justified in equilibrium just if investors expect a *depreciation* of the Canadian dollar by the same amount. For example, a positive interest rate spread would imply a *negative* signal for the domestic currency, in this case CAD. Relying on the available empirical evidence, we decided to adopt instead the flows-of-funds approach described in the text.

[11] The choice of an ex-post measure of the real interest rate is still controversial in the economic literature, more than a century after the publication of the seminal work by Fisher (1892) on monetary topics. For an overview of the issue of proper specification of measures of real interest rates, see Pasquariello (1994).

factor represents an expected depreciation of the CAD, thus reflects a positive view of the model for the US dollar.

2.3. The market segmentation factor

This component tries to capture the existing market segmentation between short-term traders and long-term-oriented investors. We measure short-term versus long-term fluctuations of the interest rates to capture their effect on the CAD through the movement of flows of funds looking for more rewarding investment opportunities. Hence, the factor is:

$$F_t^{\mathrm{YIELD}} = \delta((i_t^{\mathrm{CAD}} - i_t^{\mathrm{US}}) - (i_{\mathrm{Ht}}^{\mathrm{CAD}} - i_{\mathrm{Ht}}^{\mathrm{US}})) \tag{3}$$

In this case, i_t is a short-term moving average of a nominal interest rate that is representative of the long-end of the Canadian yield curve, while i_{Ht} represents its long-term moving average. δ is a calibration parameter selected by the analyst. Again, a negative value for the signal invites the investment manager to reduce a long position in CAD.

2.4. The industrial factor

The last component of this *stylized* fundamental model attempts to capture the portion of long-term currency variability explained by the build-up in demand for Canadian commodities, as measured by a world industrial production index. An expansion in the world economy is likely to determine higher demand for raw materials, higher commodity prices, and eventually an upward pressure over CAD, given the presumed significance of commodity trade for the Canadian economy. Hence, the factor is:

$$F_t^{\mathrm{RAW}} = \varepsilon(I_t^{\mathrm{WORLD}} - I_{\mathrm{Ht}}^{\mathrm{WORLD}}) \tag{4}$$

In this case, I_t is a short-term moving average of a world industrial production index, while I_{Ht} is its long term moving average. *Historically* high demand for Canadian raw materials will translate in expected appreciation for the currency, and consequently a positive factor signal. ε is, as usual, a calibration parameter selected by the analyst.[12] The resulting FSM signal will be positive when the model believes it is appropriate for the investment manager to be longer in CAD, negative when it suggests a relative short position in CAD.

Calibration of the model is critical to its use as a trading rule. The analyst chooses values for the calibration parameters α, β, γ, δ, and ε that make the factors additive and comparable. Moreover, those parameters implicitly attribute weights to each of the factors in the total signal. Different weighting techniques are

[12] This signal is generated every month, thus posing a specific problem for the bootstrap procedure. In Section 4.2, we illustrate a possible solution to this issue.

available. Most practitioners select the weights on the factors on the basis of their relative volatility. The weights can also be adjusted to take into account the existence of statistically significant correlation among factors. The additive FSM signal will then be:

$$F_t^{\mathrm{TOT}} = F_t^{\mathrm{PPP}} + F_t^{\mathrm{REAL}} + F_t^{\mathrm{YIELD}} + F_t^{\mathrm{RAW}} \tag{5}$$

Calibration is also important for this fundamental macroeconomic model to generate specific positioning signals. Any of the four factors in Eqs. (1)–(4) will be negative (or positive) by construction when the corresponding economic model suggests that a depreciation (appreciation) of the CAD is going to occur, hence that CAD should be sold (bought). It is then possible to interpret the factor signals as (sometimes contrasting) indicators of the expected relative strength (or weakness) of CAD with respect to the US dollar. An investor can therefore trade on them, not differently from the case of buy–sell signals resulting from *technical analysis* of the time series of CAD. Moreover, as in *technical* trading models, practitioners use the relative intensity of the negative (or positive) strength signal, and some pre-specified measure of risk to derive more specific positioning suggestions from the FSM. Consider for example the case of a US-based portfolio manager who wants to hedge the currency exposure of an underlying long position in CAD-denominated assets with the directional signals generated by a fundamental macro-model. He could then use historical data on CAD to choose the set of parameters α to ε that maximizes the Sharpe ratio of the performance differential versus a benchmark passive hedging strategy, given a target level for the standard deviation of that performance differential. With the proper calibration and the selected risk measure with respect to the benchmark strategy, the FSM model can specify by how many percentage points the manager should deviate from the selected passive hedge.

Now the analyst's question suggested in Section 1 arises: how to build a confidence interval for the signal generated by the FSM model at a certain date t? Next section proposes a simple bootstrapping methodology that provides the analyst with a satisfactory answer.

3. The bootstrapping methodology

3.1. Confidence intervals for the FSM signal F_t^{TOT}

The major theoretical task in computing the standard error (SE) of a generic FSM signal F_t, and confidence intervals for given levels of significance lies in the correct estimation of its true population distribution ψ *at a certain point in time t*. As already emphasized in Section 1, the problem of specifying an exact expression for ψ is analytically intractable, unless some very restrictive assumptions are made with respect to the distributions of the variables entering the FSM signal, but also with respect to the functional form of the signal itself.

The only viable alternative remains to specify an empirical distribution for the signal, $\hat{\Psi}$. Given an estimated $\hat{\Psi}$, the literature, especially in Efron and Tibshirani

(1986), has identified different techniques apt to measure the statistical error or better the accuracy of a given statistic ϑ, in our case the signal F_t, at a certain point in time t. More specifically, the following expression defines a measure $R(F,\hat{\Psi})$ of the precision in the estimate of the true signal as:

$$R(F_t,\hat{\Psi}_t) = F_t - \mu(\hat{\Psi}) \tag{6}$$

Here μ is the sample average for the signal F at time t estimated from the sample distribution of the signal, $\hat{\Psi}$. Hence, $R(F_t,\hat{\Psi}_t)$ provides a rough measure of the accuracy of our signal F_t in measuring the mean of the true signal, as inferred from the sample distribution $\hat{\Psi}_t$. The standard error for the estimate F_t of the true signal is then given by:

$$\mathrm{SE}(F_t,\hat{\Psi}_t) = \left[\frac{\sum_{i=1}^{N} (\hat{R}(F_i,\hat{\Psi}_t))^2}{N(N-1)}\right]^{1/2} = \left[\frac{\sum_{i=1}^{N} (F_i - \bar{F}(\hat{\Psi}_t))^2}{N(N-1)}\right]^{1/2} \tag{7}$$

The specification of the mean signal $\bar{F}$ and N depends on the methodology adopted to estimate ψ, as it will be made clearer later in the exposition. Finally, the confidence interval for the estimated signal F_t can be set according to the following alternative but closely related methods:

3.1.1. The standard approach

A range centered at the estimated signal is specified as:

$$[F_t \pm \mathrm{SE}(F_t,\hat{\Psi}_t) z^{(\alpha)}] \tag{8}$$

where $z(\alpha)$ is the $100-(\alpha/2)$ percentile point of a standard normal distribution.

The Standard approach is clearly not appropriate when the underlying population distribution for F_t, or our estimate $\hat{\Psi}$ is not symmetric, but skewed. In that case, a central sample interval clearly ignores the degree of asymmetry in the true distribution. Efron and Tibshirani (1986) show that the following two methods perform more satisfactorily when the analyst has reason to believe that the underlying ψ is strikingly asymmetric.

3.1.2. The percentile approach

If we define $\hat{\Psi}_t$ to be the estimated cumulative density function of F_t, then this approach identifies the following endpoints for a confidence interval around F_t:

$$\begin{aligned} &\hat{\Psi}_t(s) = \Pr\{F_t \prec s\} \\ &\lambda(\alpha/2) = \hat{\Psi}_t^{-1}(\alpha/2) \quad \lambda(1-\alpha/2) = \hat{\Psi}_t^{-1}(1-\alpha/2) \\ &[\lambda(\alpha/2),\lambda(1-\alpha/2)] \end{aligned} \tag{9}$$

In other terms, the confidence interval calculated through the percentile approach is simply the interval between the $100(\alpha/2)$ and $100(1-\alpha/2)$ percentiles of the estimate for the distribution ψ.

Another possible reason why the standard methodology can be misleading is that it ignores the possibility that $\hat{F}_t$, our observed signal, contains a prediction bias, as described in Eq. (6). The simple percentile approach does not take this bias into account in the construction of a viable confidence interval. The following procedure makes an adjustment for this type of bias.

3.1.3. The bias-corrected percentile approach

Let:

$$z_0 = \phi^{-1}(\hat{\Psi}_t^{-1}(F_t)), \tag{10}$$

where ϕ is the standard Normal cumulative density function. Then the confidence interval for F_t, is derived as:

$$\begin{aligned} &\hat{\Psi}_t(s) = \Pr\{F_t \prec s\} \\ &\pi(\alpha/2) = \hat{\Psi}_t^{-1}(\phi(2z_0 + z^{(\alpha/2)})) \quad \pi(1-\alpha/2) = \hat{\Psi}_t^{-1}(\phi(2z_0 + z^{(1-\alpha/2)})) \\ &[\pi(\alpha/2), \pi(1-\alpha/2)] \end{aligned} \tag{11}$$

The confidence interval generated by Eq. (11) can be further adjusted for skewness, by implementing the following correction:

$$\begin{aligned} &\mathrm{SKEW}_t = \mathrm{SKEW}(\partial \log(\hat{\Psi}_t(F_t))/\partial F) \\ &b = \mathrm{SKEW}_t/6 \\ &\pi^*(\alpha/2) = \hat{\Psi}_t^{-1}\left(\phi\left(z_0 + \frac{z_0 + z^{(\alpha/2)}}{1 - b(z_0 + z^{(\alpha/2)})}\right)\right) \\ &\pi^*(1-\alpha/2) = \hat{\Psi}_t^{-1}\left(\phi\left(z_0 + \frac{z_0 + z^{(1-\alpha/2)}}{1 - b(z_0 + z^{(1-\alpha/2)})}\right)\right) \\ &[\pi^*(\alpha/2), \pi^*(1-\alpha/2)] \end{aligned} \tag{12}$$

The statistic SKEW_t depends on the degree of skewness bias existing in the true population distribution for F_t. As suggested by Efron (1984), SKEW_t is approximately equal to z_0 under reasonable asymptotic conditions.

3.2. Bootstrapping

We now turn to the issue of determining an empirical distribution $\hat{\Psi}_t$ for the signal statistic F_t^{TOT}. We propose three different procedures to obtain $\hat{\Psi}_t$ from available historical data for both the signal realizations and the variables entering F_t^{TOT}.

3.2.1. ψ from History

A certain number of past observations for the signal F_t^{TOT} are considered to estimate its mean and standard deviation, as well as $\hat{\Psi}_t$. In Eq. (7), N becomes the distance, in number of trading days, between the current observation t and the first observation in the time-range selected by the analyst.

As the empirical application in the next section makes clear, the *Historical* method is not appealing, as it does not capture the variability of the signal.

Consequently, the estimated standard deviation does not measure accurately the effective degree of uncertainty related to the signal provided by the FSM. Non-parametric bootstrapping from historical data presents more attractive features. This procedure involves drawing a certain number of samples from the *empirical* distribution of the statistic of interest. Past observations contain information about the true and unknown distribution of the data and true and unknown characteristics of the underlying population. Hence, the empirical distribution obtained through a cross-sectional bootstrap sample drawn *with replacement* from the actual series turns out to be the non-parametric maximum likelihood estimate (MLE) of the unknown distribution ψ. Consequently, Efron and Tibshirani (1986) show that the standard error of the signal F_t^{TOT} is the non-parametric MLE of the true standard error.

The bootstrapping procedure can also be carried out parametrically. However, this would imply for the analyst to assume ex-ante a specific form for the true sampling distribution ψ. Such an assumption may limit the effectiveness of the technique and ignore some population characteristics not fully described by any available theoretical distribution. These are exactly the characteristics that the non-parametric bootstrapping technique preserves by extracting $\hat{\Psi}$ from the available data. As we shall see, the cross-sectional sampling may involve the sample F_t^{TOT} series or the series of factors entering F_t^{TOT}. In the first case we have the following.

3.2.2. ψ from Indirect bootstrapping

Given a time-range selected by the analyst, at time t N bootstrapped unit-samples for each of the factors in the signal F_t^{TOT} are independently drawn with a random number generator. For each of the samples, the resulting signal is computed. Then, the empirical distribution for $\hat{\Psi}_t$ is derived from the N signals F_t^{TOT} and formulas (6)–(12) are implemented.

To ensure internal consistency at a specific point in time among macroeconomic variables that the analyst assumes to be related to each other, a special grouping procedure can be implemented. By using the same random number generator output in drawing a sample for the selected variables, it is in fact possible to preserve the underlying simultaneous relationship among them. For example, if the random number generator output corresponds to date $(t-k)$ for the nth bootstrap unit-sample at time t, values from date $(t-k)$ are used for each of the *grouped* variables, and the nth bootstrapped F_t^{TOT} is calculated according to those variables. Finally, if sampling involves the total FSM signal defined in Eq. (5), we have the following.

3.2.3. ψ from Direct bootstrapping

In this case, given a time-range selected by the analyst, N bootstrapped unit-samples for F_t^{TOT} are generated with random and independent drawings from the observed sample of past total signals. Then, the empirical distribution $\hat{\Psi}_t$ is derived accordingly.

Indirect bootstrapping with *grouping* appears to be the most appealing procedure to an analyst on an ex-ante basis. The reasons for this bias are twofold. First, the proposed methodology substantially replicates N times the same event occurring at time t, when the FSM model generates the signal. This ensures consistency between the random generation process and the true, and in this case known, generation process for the signal. Second, using the same random generator output for each of the variables entering the model guarantees internal consistency among the data at a specific point in time. This procedure preserves cross-sectional correlation among factors, i.e., information contained in the relationship among the variables that would otherwise be lost in a complete randomization of the simulation process.[13] In other terms, that consistency derives from basic and unknown economic relationships among the different factors. Grouping selected variables in the bootstrapping approach simply allows the analyst to preserve these unobservable characteristics of the data while trying to define the degree of precision of the signal. The procedure does not come without cost, given the trade-off between the gain in internal coherence among the randomly extracted variables and the loss of variability in the simulation process. Direct bootstrapping is an extreme form of grouping, for it preserves completely the cross-sectional consistency among the components of the FSM signal, at the cost of reduced variability in the simulated generation of the factors.

It is important to observe that, in the light of the evidence reported in footnote 13, the presence of serial correlation for some of the variables to be bootstrapped, but not in the total signal, does not help us selecting among the criteria. In fact the autocorrelation should not affect significantly the results of the bootstrapping method when it is both very significant or so low that can be ignored. Why? Suppose that the serial correlation for a time series to be bootstrapped is very high, explaining more than 90% of the series variability for lags up to 30 months, and more than 75% of the same variability for lags up to 100 months, as in the case of the PPP series for Canada. Then, the value assumed by the variable at time t will be a good proxy for values assumed by the same variable at time $(t-j)$ or $(t+j)$. If the serial correlation for the variable is instead very low and decreasing for higher lags, then the direct bootstrapping procedure is clearly neutral and harmless. The

[13] Note that any serial correlation of the factors is pulled out from the data through bootstrapping. This is the reason why no technical signal has been introduced as part of the basic FSM model, although Levich and Lee (1993) show that the profitability of technical trading rules in the foreign exchange market is statistically significant even when the underlying time series of currency futures prices is bootstrapped. Nonetheless, we maintain that the exclusion of mechanical trading rules from the factor-construction process is more appropriate, as the procedure adopted in this paper does not preserve any time-series property of each of the bootstrapped variables. This limitation could be particularly severe in the case of price-related series, where persistence is much more pronounced. For example, the AR(1) coefficient for the Canadian versus US real interest rate differential over the time frame adopted for our analysis is 0.9973 and does not reduce below 0.90 up to 30 lags. The nominal yield differential shows similar values for AR(1) to AR(5) components, but there the autocorrelation fades away much faster. Not surprisingly, autocorrelation is most evident in the PPP series, and least evident for the industrial factor. However, interestingly enough, autocorrelation is much less evident for the total signal generated by the model. In the interval for which, as described in Section 4, we generated confidence intervals for the total signal generated by the model, the AR(1) coefficient for F_t^{TOT} is equal to 0.5105 and not statistically significant for lags of 5 months or more.

Section 4 is devoted to an empirical application of the three techniques here described to the CAD exchange rate and the FSM model of Eqs. (1)–(5).

4. An application of non-parametric bootstrapping to the CAD FSM model

4.1. Model building and bootstrapping procedures

The model of Section 2 offers the opportunity to apply the methodologies proposed in this paper and compare their results. Daily time series for each of the variables in Eqs. (1)–(5) are used to build a FSM signal for the Canadian dollar from January 1, 1980 to April 30, 1998.[14] Short- and long-term moving averages in Eqs. (2)–(4) are computed over three-month and ten-year intervals, respectively. As previously mentioned, calibration of the model implicitly weights the contribution of each factor to the total signal F^{TOT}. Consistently with a widely adopted strategy among practitioners, we compute the daily performance differential, Δ_t, an analyst would have experienced by passively hedging (i.e. selling) 50% of a hypothetical position in Canadian dollars held by a US-based investor, instead of hedging the same position by following the deviations from the benchmark suggested by the FSM model. Then, we choose the set of values for the various parameters in Eqs. (1)–(4) that maximizes the average annual Sharpe ratio for Δ_t, subject to the constraint that each of the factors contributes equally to a target level for the standard deviation of that performance differential.[15]

Fig. 1A shows the behavior of the CAD/$ spot exchange rate and the FSM signal from January 1, 1993 to April 30, 1998.

As expected, the signal becomes negative in periods of US dollar strength, when it suggests selling more than 50% of the underlying CAD asset exposure, positive in periods of CAD appreciation, when it instead suggests reducing the benchmark hedged position. Fig. 1B illustrates the performance of each of the four factors entering the FSM signal according to Eq. (5). F^{YIELD} and F^{REAL} appear to have directed the model to increase the hedge for the underlying CAD asset exposure in 1996. A decline in F^{RAW} seems instead to have supported the more recent negative view on the Canadian dollar in spite of an increase in F^{YIELD} and F^{PPP}. F^{PPP} is always positive in our sample, suggesting a relatively virtuous behavior of the Canadian inflation rate between 1993 and 1998.

---

[14] The interest rate and price data have been kindly provided by J.P. Morgan Investment Management and originate from Bank of Canada and the Federal Reserve Bank. The spot rates correspond to the noon buying rate for cable transfers payable in foreign currencies as registered by the Federal Reserve Bank of New York each business day. The World Production Index is calculated monthly by the IMF.

[15] The standard deviation of the performance differential between an actively managed portfolio and a benchmark is known among practitioners as '*tracking error*'. The constraint that each factor contributes equally to the total tracking error arises from a common ex-ante assumption among currency portfolio managers about the relative importance of the signals. The (rounded) parameter values resulting from this procedure and adopted in this section are $\alpha = 0.75$ and $\beta = 0.05$ in Eq. (1), $\gamma = 3.85$ in Eq. (2), $\delta = 14$ in Eq. (3) and $\varepsilon = 76.5$ in Eq. (4).

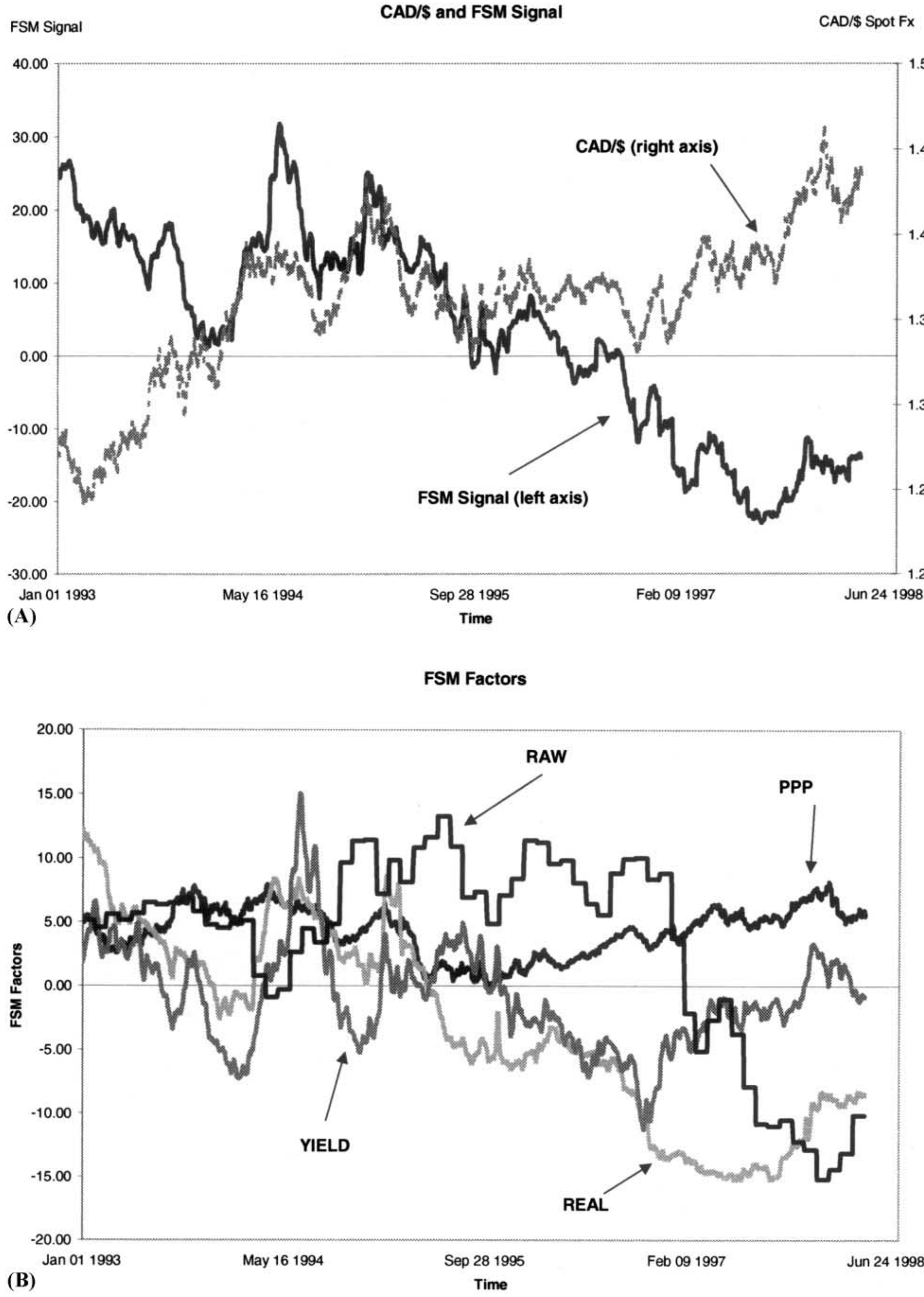


Fig. 1. (A) CAD/$ spot exchange rate and the FSM signal from January 1, 1993 to April 30, 1998. The spot rate corresponds to the noon buying rate for cable transfers payable in foreign currencies as registered by the Federal Reserve Bank of New York each business day. The FSM signal is calculated according to Eqs. (1)–(5) in the text. (B) FSM Factors from January 1, 1993 to April 30, 1998. The factors are computed according to Eqs. (1)–(4) in the text. The data series utilized in the computations were kindly provided by J.P. Morgan Investment Management.

Table 1 reports summary statistics for the FSM signal F^{TOT} and for each of its components. F^{RAW} and F^{REAL} appear to be the most volatile factors, while F^{PPP}, as suggested above, is the most stable.

It is worth observing that the FSM signal may take contrarian views: at the end of 1997 for example, F^{TOT} was suggesting a reduction of the benchmark 50% hedge in spite of a continuing depreciation of the Canadian dollar.[16] It is in similar circumstances that the precision of the signal is particularly important for the analyst: again, how confident, he might ask, was the model that the passive hedge position should have been reduced? How precise was the signal in that specific time frame? To explore these issues, we focus on the following sub-interval, May 15, 1997 to April 30, 1998. Using each of the procedures described in the past section to extract the empirical distribution of the signal F_t^{TOT}, $\hat{\Psi}_t$, we calculate bias, standard errors (SE) and confidence intervals (CI) according to Eqs. (6)–(12).

The implementation of the Indirect bootstrapping approach requires additional attention for the following reasons. First, one of the factors, F^{RAW}, relies on data that are released by the IMF at the beginning of each month. Thus, the resulting signal does not proceed smoothly along the sample but jumps in the first trading day of each month, according to the data release schedule. To reflect this process, we randomly draw the variables entering F^{RAW} just at the beginning of each month, in order to maintain the same internal consistency advocated for in Section 3. As we shall see, this decision affects, although not significantly, some of the results that follow. Second, we specify F^{YIELD} and F^{REAL} as the variables to be grouped in the Indirect bootstrapping procedure. Although some economic considerations make this choice reasonable, the selection is still somehow arbitrary, and it should reflect the analyst's view of the interaction among the components of F_t^{TOT}.

We carry out the bootstrapping methods described in Section 3 by independently drawing (with replacement) N separate time t unit-samples for each of the factors described by Eqs. (1)–(4) from a sample of 252 trading days before time t. We then

Table 1
Sample statistics for the FSM signal and its four components—January 1, 1993 to April 30, 1998

Statistics	F^{TOT}	F^{PPP}	F^{REAL}	F^{YIELD}	F^{RAW}
Mean	3.40	4.32	−3.33	−0.89	3.30
Median	4.52	4.60	−4.27	−1.26	5.60
S.D.	13.86	1.93	7.40	4.00	7.58
Absolute deviation[a]	11.67	1.61	6.38	3.12	6.02

[a] The absolute deviation statistic is computed as the average of the absolute deviations of each observation from its mean.

[16] It is important to note that capturing the directionality of the exchange rate does not necessarily make the model more profitable (or less unprofitable) than the passive hedge strategy. The returns potentially generated by the model's suggested deviations from the benchmark have to be eventually adjusted by transaction costs and cost-of-carry, if the positions are implemented on a forward basis.

compute time t sample statistics and confidence intervals using the resulting N observations for each of the factors and the total signal F_t^{TOT}. The Historical method is instead implemented using 504 trading days immediately preceding time t. We finally repeat this procedure for each day t of our selected sub-sample, May 15, 1997 to April 30, 1998. As an example, Table 2 reports values for bias, SE and confidence intervals (CI) for each of the suggested approaches for $t =$ April 30, 1998, and $N = 1000$.[17]

Table 2
Precision measures of the FSM signal F_t^{TOT} for each of the bootstrapping[a] techniques proposed in the text—April 30, 1998

Statistics (and the corresponding equation in the text)	ψ from History[b,c] [1]	ψ from Direct bootstrapping[d] [2]	ψ from Indirect bootstrapping —grouping[e] [3]	ψ from Indirect bootstrapping—no grouping[e] [4]
Observed FSM signal (5)	−14.02	−14.02	−14.02	−14.02
Bias measure (6)	−1.83	+3.22	+2.45	+2.34
Standard error (7)[f]	0.313	0.103	0.196	0.177
SKEW[g]	+0.493	−0.144	+0.414	+0.474
Standard CI (8)[h]	−14.54 to −13.51	−14.19 to −13.85	−14.34 to −13.70	−14.31 to −13.73
Percentile CI (9)[h]	−21.91 to +0.30	−22.26 to −12.16	−25.01 to −4.78	−24.07 to −6.60
Bias percentile CI (10–11)[h]	−21.92 to +0.23	−22.06 to −12.20	−25.23 to −4.35	−24.35 to −5.75
Skew percentile CI (10–12)[h]	−21.68 to +0.81	−22.07 to −12.19	−25.22 to −4.31	−24.34 to −5.72

[a] Independent drawings are performed from a sample of 252 days before time t.
[b] The 'historical' ψ has been computed using the past 504 trading days' observations for the signal F^{TOT}.
[c] The empirical $\hat{\Psi}$ is derived from past observations for the signal $F(t)$, as described in the text.
[d] The empirical $\hat{\Psi}$ is derived from Direct bootstrapping of the signal $F(t)$, as described in the text.
[e] The empirical $\hat{\Psi}$ is derived from Indirect bootstrapping of the variables entering the signal F_t^{TOT}, with our without grouping, as described in the text.
[f] In Eq. (7), $N = 504$ in [1], as the variance is computed from the past 504 trading days, while $N = 1000$ for both Direct [3] and Indirect [2] bootstrapping.
[g] SKEW is a measure of the asymmetry of the distribution of the signal. For a symmetric distribution of ψ, SKEW would be zero. If SKEW is positive, the 'long tail' of the distribution is in the positive direction.
[h] Confidence intervals (CI) are calculated for a 10% significance level.

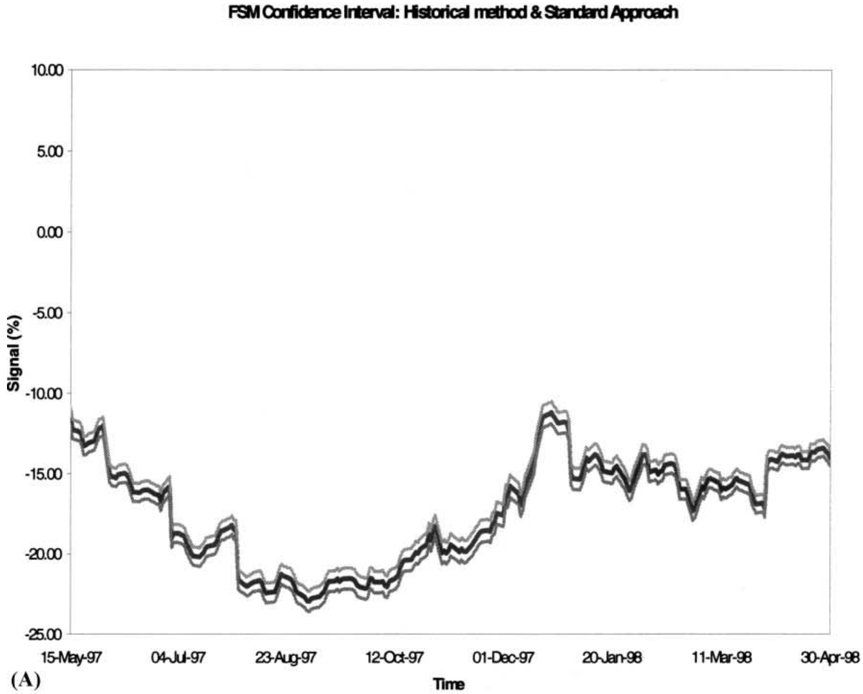
[17] The issue of establishing the optimal bootstrapping sample size has not been solved yet by the literature. Efron and Tibshirani (1986) assert that there is little improvement in the estimation of the standard error of the relevant statistic past $N = 100$. The situation seems to be quite different, they argue, for setting bootstrap confidence intervals, with Efron (1984) claiming that $N = 1000$ is a rough minimum for the number of samples necessary to compute percentile intervals. However, bootstrap experiments implemented by the author for $N \geqslant 1500$ did not show any significant improvement in the results, with respect to our initial assumption of $N = 1000$. As suggested by Efron (1984), the coefficient b in Eqs. (10)–(12) is estimated as SKEW(F_t^{TOT})/($6N^{0.5}$). Again, N is equal to 504 trading days for the Historical method.

The symmetric confidence intervals computed using the standard approach are very tight, as they do not take into account any form of skew or prediction bias possibly affecting the model's directional signal F^{TOT}. However, when the percentile approach is used and the correction for bias and skewness are accounted for, the confidence intervals computed from Eqs. (9)–(12) are very different from the intervals resulting from the standard approach, although very similar to each other, independently from the procedure used to estimate ψ. All lower bands are significantly smaller, although do not seem to be significantly different whether estimated with the History or Indirect bootstrapping method, with and without grouping (i.e. completely at random). This is not the case for the upper band, as long as bias and skew corrections are accounted for in the interval calculation procedure. In fact, for the Historical method the adjusted upper bands become even positive, for the existing bias measure was negative, as $\mu(\psi) > F_t^{TOT}$. This means that the corresponding empirical distribution for the total FSM signal implies a slightly less pessimistic view for the CAD than in the case of Indirect bootstrapping. Thus, the resulting estimated upper bands are higher. Direct bootstrapping offers a different picture. The confidence intervals so estimated are *tighter* and shifted *downward*, thus signaling a stronger confidence of the model regarding the suggested increase in the hedge for the underlying CAD asset than in the case of the other two alternative procedures described in the paper. Interestingly enough, the estimated skew for F_t^{TOT} is lower than the one computed with indirect bootstrapping or simply from historical data; however, the bias measure is generally higher.

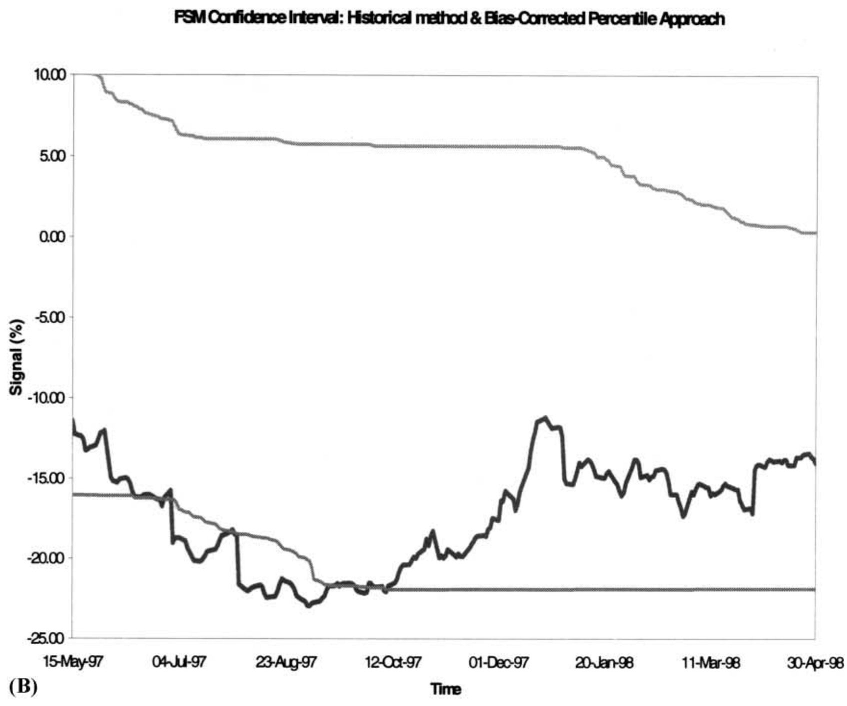
To better evaluate the performance of the three methodologies presented in the paper, we repeat the same exercise of Table 2 for each date between May 15, 1997 and April 30, 1998 and report our results in Figs. 2 and 3. Fig. 2A displays the 10% significance-confidence interval for the signal F^{TOT} when the distribution ψ is estimated from past observations, i.e. according to Section 3.2.1, and the endpoints of the CI are computed using the standard approach of Eq. (8). As already mentioned, the intervals are relatively tight in nature and do not take into account model's uncertainty, skewness and bias-prediction errors in the estimation of F^{TOT}.

When the skewness bias is considered and a bias-corrected percentile confidence interval is calculated according to Eq. (12) of Section 3.1, the resulting band, as shown in Fig. 2B, is significantly wider and fluctuates less over time, because the historical percentiles contain less information about the variability of the true and unobservable signal. In other terms, the 90% interval adjusts less rapidly to the

Fig. 2. Confidence intervals for F^{TOT} from May 15, 1997 to April 30, 1998. F^{TOT} is computed according to Eqs. (1)–(5) in the text. In (A) the distribution for F_t^{TOT} has been estimated using an interval of 504 trading days immediately preceding time t. The confidence interval is computed according to the standard procedure described in the text, Section 3.1 Eq. (8). In (B) the distribution for F_t^{TOT} has been estimated using an interval of 504 trading days immediately preceding time t. The confidence interval is computed according to the bias-corrected percentile procedure described in the text, Section 3.1 Eqs. (10)–(12). The data series utilized in the computations were kindly provided by J.P. Morgan investment management.



(A)



(B)

Fig. 2. (*Continued*)

amount of information contained in the most recent signal data. Nonetheless, it also appears to account more explicitly for the uncertainty surrounding the model's predictions, as it suggests a bigger range of values in which the true and unobservable total signal is likely to be found, given the information contained in the past values for the observed directional signal.

More importantly, in the second half of 1997 the observed signal frequently crosses its confidence band. This is made possible by the more explicit recognition of the asymmetry of the unknown distribution ψ. How to interpret this result? As observed by Wang (1992), the bias-correction percentile method generally performs well when the bootstrapped samples are approximately normal, but might produce extremely asymmetric confidence intervals when the empirical distribution of the statistic F^{TOT} is highly skewed. The average skewness measure for the interval during which the bootstrapped total FSM signal crosses the lower end of the confidence interval is in fact negative (-0.11) and statistically significant. This finding implies that the empirical distribution of F^{TOT} was characterized by an asymmetric tail extending toward more negative values, but also that the degree of skewness in the bootstrapped distribution $\hat{\Psi}$ for F^{TOT} was small. Moreover, the confidence interval in Fig. 2B is already adjusted for the presence of skewness bias, as in Efron and Tibshirani (1986).

These considerations suggest us a more natural interpretation for F_t^{TOT} being lower than $\pi^*(\alpha)$. The confidence interval is defined, at the chosen significance level α, by the probability $1-\alpha$ that the unobservable true FSM signal F^{TOT} lies in the estimated band. In other terms, this means that in repeated sampling an interval constructed in this fashion would contain the true parameter $100(1-\alpha)$ percent of the times. Hence, the fact that the observed F^{TOT} does not fall into the band indicates that the value observed for the FSM signal simply represents a tail-event. In other terms, the model registered an unlikely value for the signal. The analyst would then be *less* confident in the precision of the model's estimate for the true (and unobservable) F^{TOT}, and consequently *less* confident in the resulting suggestion to increase the hedge for the underlying position in CAD assets. The larger bandwidth for the confidence interval at or around the time when $F^{TOT} < \pi^*(\alpha)$ suggests in fact that more noise was conditioning the precision of the signal-gener-

Fig. 3. Confidence interval for F^{TOT} from May 15, 1997 to April 30, 1998. F^{TOT} is computed according to Eqs. (1)–(5) in the text. In (A) the distribution for F_t^{TOT} has been estimated by Indirect bootstrapping of the variables entering the FSM signal. YIELD and REAL variables are grouped, according to the analysis of Section 3. The confidence interval is computed according to the bias-corrected percentile procedure described in the text, Section 3.1 Eqs. (10)–(12). In (B) the distribution for F_t^{TOT} has been estimated by Direct bootstrapping of the variables entering the FSM signal. Variables are not grouped. The confidence interval is computed according to the percentile procedure described in the text, Section 3.1 Eq. (9). In (C) the distribution for F_t^{TOT} has been estimated by Direct bootstrapping of the variables entering the FSM signal. Again, variables are not grouped. The confidence interval is computed according to the bias-corrected percentile procedure described in the text, Section 3.1 Eqs. (10)–(12). The data series utilized in the computations were kindly provided by J.P. Morgan Investment Management.

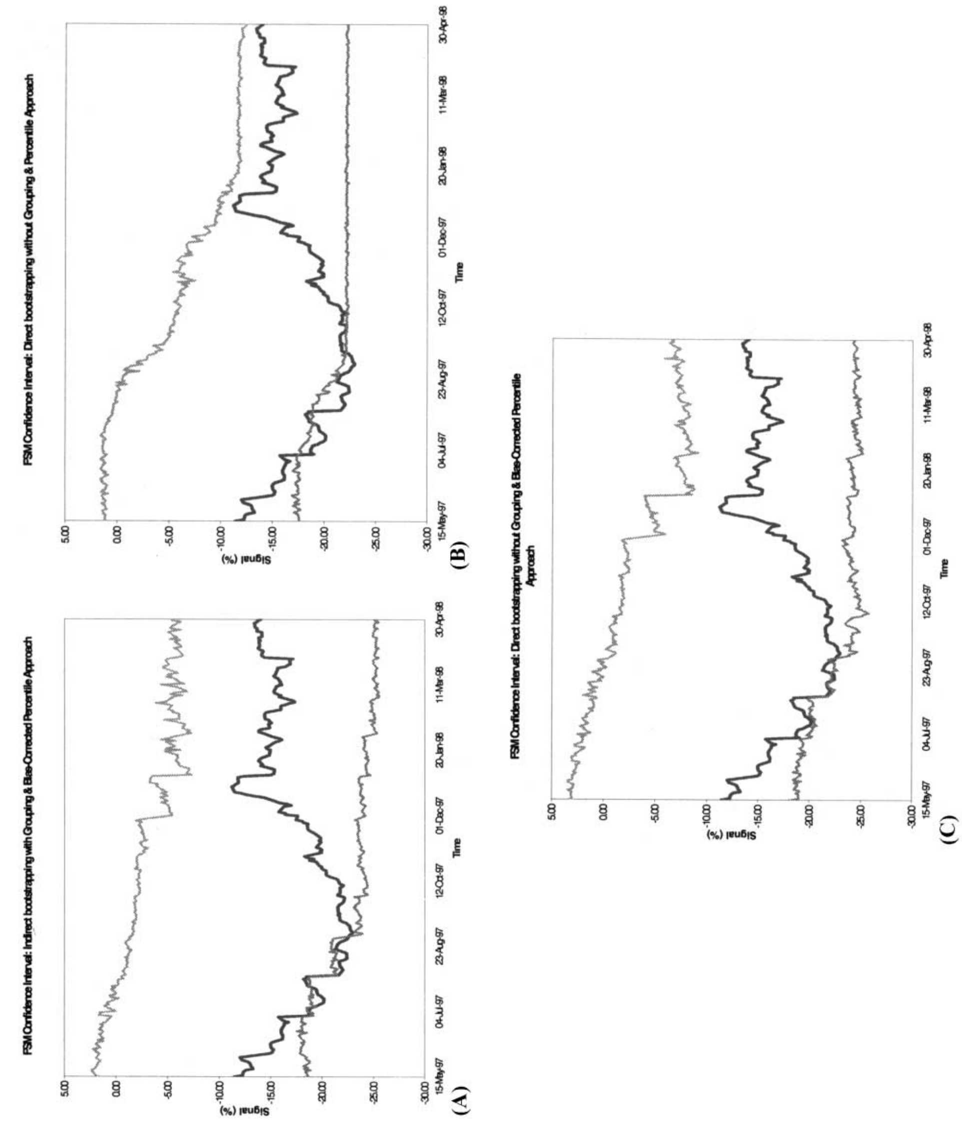


Fig. 3. (*Continued*)

ation process of Eqs. (1)–(5) than in subsequent periods. When the confidence interval is obtained from the Indirect bootstrapping procedure with *grouping*, as in Fig. 3A, F^{TOT} still crosses the lower endpoint of its confidence interval several times between July and September of 1997.

Next, the Indirect bootstrapping procedure *without* grouping is considered. As explicitly stated in Section 3.2, the rationale for this approach lies in the attempt to reproduce as accurately as possible the experiment of the formation of the signal, at the cost of losing the internal consistency among the macroeconomic data entering the model. The expected gain is in an estimated distribution function of the signal, $\hat{\Psi}_t$, that better reflects the variability of each of the variables selected by the analyst to be relevant for the Canadian dollar. And in fact the resulting CI (not reported here) shows that complete randomization seems to generate a slightly tighter and smoother band around the observed F_t^{TOT}, with respect to the interval obtained with Indirect bootstrap with grouping in Fig. 3A. In other words, independently bootstrapping F^{REAL} and F^{YIELD} reduces the uncertainty of the model in generating a directional signal, and the resulting confidence interval is tighter.

Finally, the least randomized procedure, Direct bootstrapping, is examined. Direct bootstrapping, as such, groups all the variables entering the selected model for the currency, thus preserving the cross-sectional consistency of the observations at the cost of a supposedly serious loss of variability in the simulation process. How costly is this loss? From Fig. 3B it appears that, for the last part of the sample interval, as anticipated by the analysis in Table 2, the confidence interval around F^{TOT} estimated using the percentile approach of Eq. (9) is generally tighter than the ones generated by Historical and Indirect bootstrapping methods. If however, the CI is corrected for potential bias and skewness, according to Eqs. (10)–(12), the resulting gain in precision with respect to Indirect bootstrapping disappears, as evident from comparing Fig. 3A and 3C. Clearly, the impact of the adjustment for prediction bias and skewness in ψ on the width of the confidence interval for F_t^{TOT} may be significant.

4.2. Procedure selection and model analysis

The theory presented in Section 3 does not prescribe exactly how to choose the endpoints for a confidence interval. An obvious criterion would be to select the technique that minimizes the width of the interval. If the sampling distribution is symmetric, the symmetric interval is always the best one, and the analyst will choose the tightest among the intervals provided by Eqs. (8)–(12) and the sampling approaches of Section 3.2. However, if the sampling distribution is not symmetric, then this criterion won't be optimal. When this is the case, again no clear-cut answers are available in the theory. In such a circumstance, we believe the choice of the analyst should fall over the technique that best reflects the estimated degree of skewness for the empirical distribution of the FSM signal and the bias contained in the observed F_t^{TOT}. The resulting corrections can in fact be important, as we have seen in the examples of Figs. 2 and 3. Moreover, our empirical analysis suggests that alternatively Direct and Indirect bootstrapping with grouping generate the lowest standard error for the observed F_t^{TOT} and the tightest confidence intervals in different sub-periods, although still accounting for prediction bias and skewness. From a different perspective, the evidence presented in Section 4.1 also reveals that the bandwidth of the confidence intervals for the FSM signal varies over time, independently from the procedure adopted to compute them.

These considerations lead us to the following question: can the analyst explain why most of the confidence intervals estimated above, and reported in Figs. 2 and 3, become tighter in the last few months of the sample? In other terms, why does the model become more precise in generating a directional signal in the first few months of 1998? An apparently satisfying answer to this question lies in the analysis of the rolled volatility of the CAD daily returns, of the signal itself and of the factors that enter the FSM model over time. When the return, the signal and/or the factors' volatility increase, it should be more difficult for the model to discern a direction for the currency. Consequently, the estimated band for the true signal is expected to be wider. Vice-versa for the case when volatility decreases: in that circumstance, the model should be able to point more confidently toward a specific direction.

Nonetheless Fig. 4A, with the confidence interval width (obtained with direct bootstrapping) on the left axis and the rolled annualized CAD daily return volatility[18] on the right axis, suggests a strikingly different and prima facie counterintuitive story. While the confidence interval band has been tightening by at least 10 percentage points since August 1997, with a pronounced decline between the end of 1997 and the beginning of 1998, the currency volatility increased by more than 30% on an annual basis.

There are two equally reasonable explanations for this fact. The signal seems to become more precise when strongly heading in a specific direction, and this usually happens when the underlying spot rate, here the CAD, moves away from its historical mean, thus implying by definition an increase in the volatility. Moreover, the higher degree of precision for the total signal might have been generated by lower volatility of the signal itself and of the factors that compose it, or from a change in the correlation among these factors. And in fact, as Fig. 4B reveals, this is what we observe for the selected time frame. We calculate the volatility of each of the factors and the FSM signal at time t as the corresponding annualized rolled standard deviation for a window of 252 trading days. The volatility of the total signal starts to drop at the beginning of August 1997. The decline seems to be induced mostly by the industrial production factor F^{RAW} and the real rate spread. Both F^{REAL} and F^{YIELD} factor-volatility first decrease and then increase sharply by the end of the year, but the total rolled standard

[18] The rolled volatility is calculated over a window of 252 trading days. Then, the daily CAD volatility is annualized by multiplying the resulting numbers by the square root of 252.

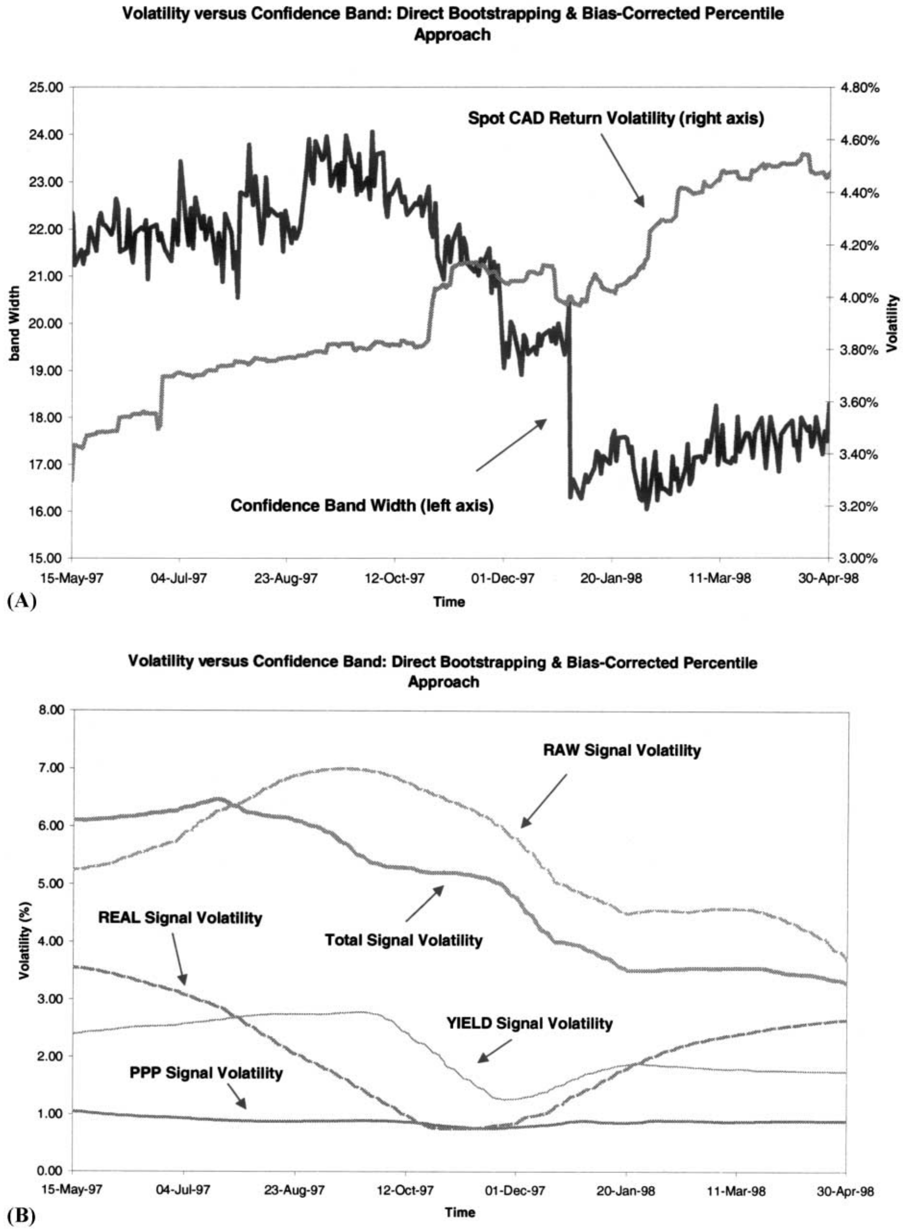


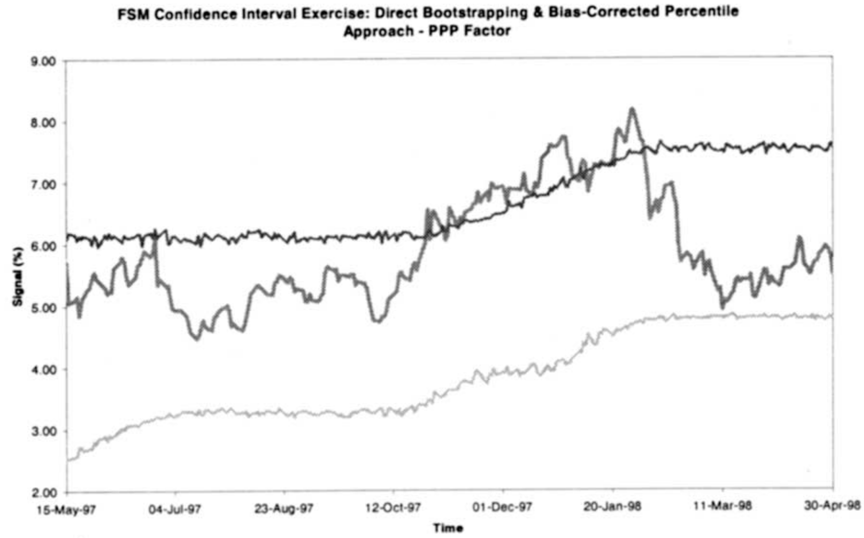
Fig. 4. (A) Confidence interval width for F^{TOT} from May 15, 1997 to April 30, 1998 and CAD annualized return volatility. F^{TOT} is computed according to Eqs. (1)–(5) in the text. The distribution for F_t^{TOT} has been estimated by direct bootstrapping of the variables entering the FSM signal. The confidence interval is computed according to the bias-corrected percentile procedure described in the text Section 3.1 Eqs. (10)–(12). The rolled volatility is calculated over a window of 252 trading days and consequently annualized by multiplying the resulting values by the square root of 252. (B) Rolled volatility of each of the factors and the FSM signal at time t. The rolled volatility is calculated over a window of 252 trading days and consequently annualized by multiplying the resulting values by the square root of 252. The data series utilized in the computations were kindly provided by J.P. Morgan Investment Management.

deviation of F^{TOT} keeps moving downward because of the assumed weightings and the structure of the covariance among the signals.

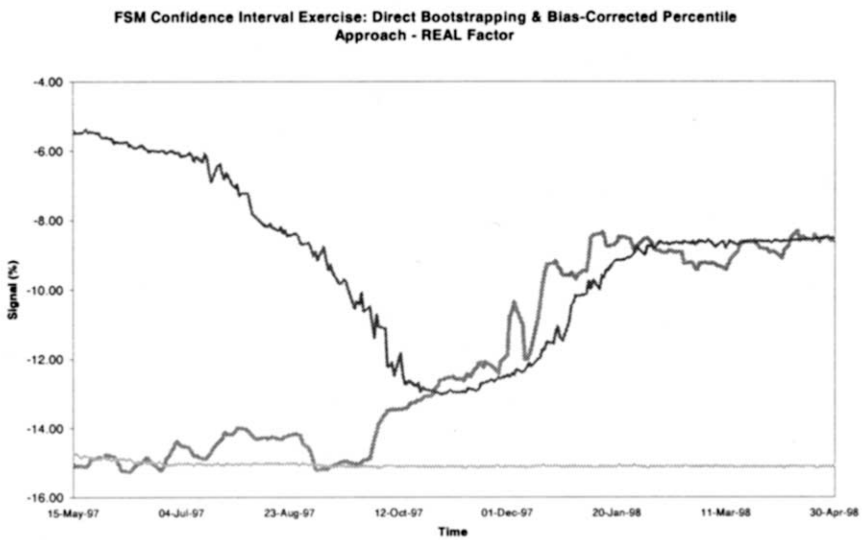
Hence, the precision of the variables entering the total FSM output seems to play an important role for the analyst to examine critically the precision of his model and to properly use its indications. Thus, given the degree of accuracy of the signal estimated through one of the suggested procedures, it appears legitimate to ask which factors contribute the most to the relative statistical strength or weakness of F^{TOT}. In other terms, can we say something about the precision of each of the factors in the model of Eqs. (1)–(5)? The resulting information would be as precious for the analyst as the confidence interval for the global directional signal. Why? Because less than significant or uncertain signals might be excluded from the computation of the total directional output of the model, F^{TOT}, thus improving its precision. The procedures that are necessary to identify confidence intervals for each of the variables entering the selected FSM follow most of the analysis developed in the last section. However, the distinction between Direct and Indirect bootstrapping is now redundant, for each factor results from a single time series of observations. Fig. 5A–5D show the behavior of non-symmetric confidence intervals for the four factors included in the model, computed with the bias-corrected percentile approach of Eqs. (10)–(12).

How to explain the less-than-favorable and *lower-than-usual* total CAD signal between July and September of 1997 we observe in Figs. 2 and 3? As evident from Fig. 5A and B, F^{PPP} (5A) and F^{REAL} (5B) were contributing significantly to the less than usual (for a 90% level of confidence) CAD-favorable signal. F^{RAW} (5D) is *unusually* negative, i.e. crosses the lower endpoint of its confidence interval for most of the sample interval considered in the analysis, while F^{YIELD} (5C) ends roughly centered at zero, with a wide band oscillating from $+3$ to -3. Clearly, a wise manager could exclude from time to time the factors that are not informative enough (when they cross the endpoints of their corresponding CI) or simply not precise enough (when the bandwidth of their CI is too large) to give a reasonable contribution to a more precise total directional signal. Finally, the sudden downward shift in the width of the CI for F^{TOT}, displayed in Fig. 4A, appears to be due to the dynamics of the two interest rate factors in the trading rule of Eqs. (1)–(5). In fact, the confidence intervals for both F^{REAL} (5B) and F^{YIELD} (5C) shrink dramatically in December of 1997, after a period of declining rolled volatility, as previously observed in Fig. 4B. The CI for F^{RAW} (5D) tightens only by the end of January 1998, while the uncertainty surrounding F^{PPP} (5A) is substantially unchanged over the sample period.

A word of *caveat* is necessary. The confidence interval decomposition ignores, for its nature, the existence of co-movements among factors. Hence, any conclusion implying from it has to be attenuated by the acknowledgement that two mutually dependent and equally imprecise factors may still generate a more precise confidence interval for the aggregate signal. Nonetheless, confidence interval decomposition helps the analyst in discerning the elements that contribute to

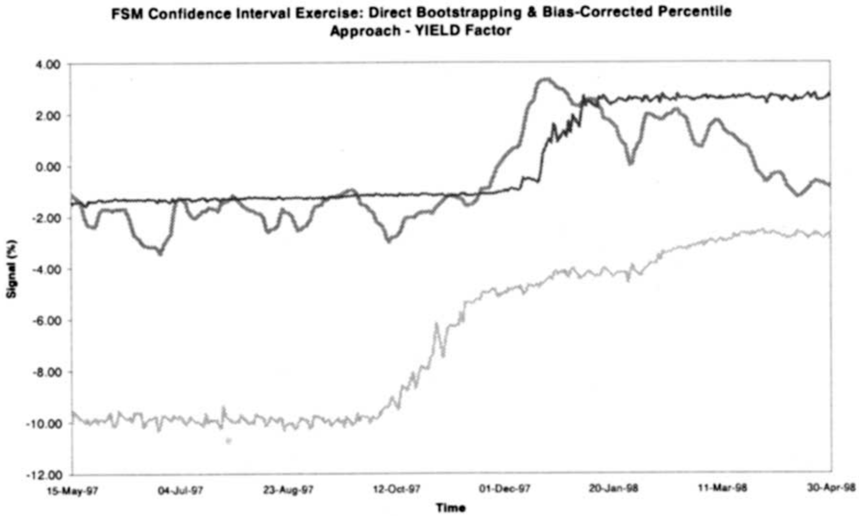


A

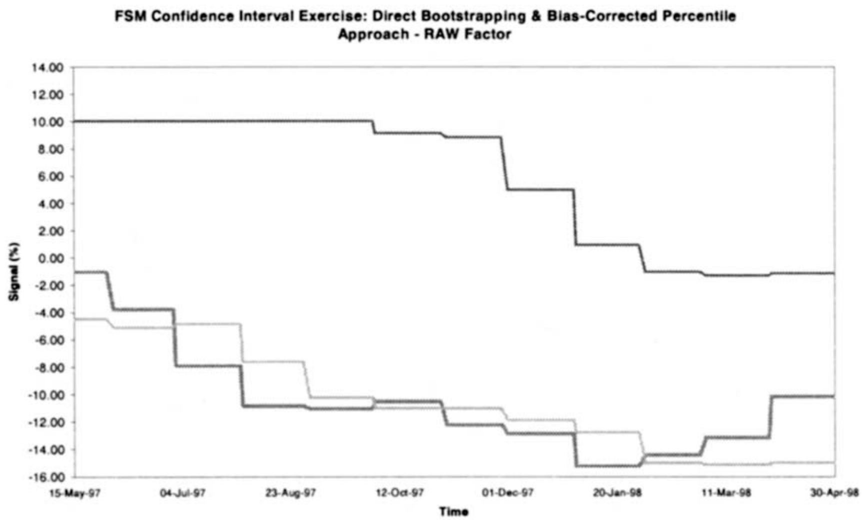


B

Fig. 5. Confidence interval for F^{PPP} (A), F^{REAL} (B), F^{YIELD} (C) and F^{RAW} (D) from May 15, 1997 to April 30, 1998. Each factor is computed according to Eqs. (1)–(4) in the text. The distribution for each factor has been estimated by Direct bootstrapping. The confidence intervals are computed according to the bias-corrected percentile procedure described in the text, Section 3.1 Eqs. (10)–(12). The data series utilized in the computations were kindly provided by J.P. Morgan Investment Management.



C



D

Fig. 5. (*Continued*)

the degree of precision of the model he aims to use profitably in the currency markets. The knowledge of the estimated precision of each factor at a specific point in time could also be utilized to update the calibration parameters, i.e. to make the factor-weights endogenous, and hence to generate a more precise directional signal as an output of a fundamental model of exchange rate behavior.[19]

5. Conclusions

The problem of measuring the precision of signals generated by fundamental macroeconomic models is not trivial. In this paper, we suggested three different approaches for the estimation of the true and unknown distribution of the population signal. We then used established statistical techniques and the information contained in the empirical distribution of the signal to measure its precision at a specific point in time. Direct and Indirect bootstrapping are devised to capture the unknown variability of the signal without altering the information content of the historical data.

We implement this framework for a simple fundamental model for the CAD/$ exchange rate. We find that accounting for skewness and prediction bias affects significantly the shape and width of the estimated confidence intervals around the estimated directional signal. When this is done, the results lead us to exclude the naïve Historical approach from the analyst's set of options, for it does not satisfy the need for a measure of precision in the model's recommendations. No clear-cut selection criteria seem to be available in choosing between the two proposed forms of bootstrapping. Nonetheless, factor-decomposition and statistical analysis of the building blocks for F^{TOT} may enhance the analyst's understanding of the degree of precision observed for the total factor, facilitate a more precise selection of *significant* signals, and determine a more effective factor-weighting in forming a directional view on the currency.

References

Efron, B., 1981. Nonparametric estimates of standard error: the jackknife, the bootstrap, and other resampling methods. Biometrika 68, 589–599.

Efron, B., Gong, G., 1983. A leisurely look at the bootstrap, the jackknife, and cross-validation. Am. Statistician 37, 36–48.

Efron, B., Tibshirani, R., 1986. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. Stat. Sci. 1, 54–77.

Efron, B., 1984. Better bootstrap confidence intervals. J. Am. Stat. Assoc. 82, 171–185.

Fleming, J.M., 1962. Domestic financial policies under fixed and under floating exchange rates. Intl. Monetary Fund Staff Papers 12, 369–380.

Fisher, I., 1892. Mathematical Investigations in the Theory of Value and Prices. Yale University.

Isard, P., 1995. Exchange Rate Economics. Cambridge Surveys of Economic Literature.

[19] The precision of each of the factors could be measured for example by the relative width of the corresponding confidence interval, i.e. computed as the ratio between the width of the confidence interval and the factor at each point in time. Weights for the individual factors in Eq. (5) could then be derived from these comparable precision indices.

Levich, R., Lee, T., 1993. The significance of technical trading rule profits in the foreign exchange market: a bootstrap approach. J. Intl. Money Finance 12, 451–474.

Levich, R., 1985. Empirical studies of exchange rates: price behavior, rate determination and market efficiency. In: Jones, G., Kenen, G. (Eds.), Handbook of International Economics, vol. 1. North-Holland, Amsterdam, pp. 255–261.

Levich, R., 1989. Is the foreign exchange market efficient. Oxford Rev. Econ. Policy 5 (Fall issue), 40–60.

Mundell, R.A., 1961a. The international disequilibrium system. Kyklos 14, 153–172.

Mundell, R.A., 1961b. A theory of optimum currency areas. Am. Econ. Rev. 51, 657–665.

Muralidhar, A., Mala Khin, U., 2001. Establishing a peer comparator universe for an institutional investor. In Pension Fund Management, Stanford University Press, forthcoming.

Pasquariello, P., 1994. Real Interest Rates in the Eighties. Unpublished Laurea dissertation, Universita' Commerciale Luigi Bocconi, Milano.

Wang, M., 1992. The analysis of retrospectively ascertained data in the presence of reporting delays. J. Am. Stat. Assoc. 87, 397–406.