International Journal of Forecasting 33 (2017) 894–914

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Selecting exchange rate fundamentals by bootstrap

Pinho J. Ribeiro [1]

Adam Smith Business School, University of Glasgow, UK

ARTICLE INFO

Keywords:
Bagging
Bumping
Combined forecasts
Economic evaluation of exchange rate models
Exchange rate forecasting

ABSTRACT

Research shows that the predictive ability of economic fundamentals for exchange rates varies over time; it may be detected in some periods and disappear in others. This paper uses bootstrap-based methods to select time-specific conditioning information for the prediction of exchange rates. By employing measures of the predictive ability over time, along with statistical and economic evaluation criteria, we find that our approach based on pre-selecting and validating fundamentals across bootstrap replications leads to significant forecast improvements and economic gains relative to the random walk. The approach, known as bumping, selects parsimonious models that have out-of-sample predictive power at the one-month horizon; it is found to outperform various alternative methods, including Bayesian, bagging, and standard forecast combinations.

1. Introduction

The identification of macroeconomic fundamentals that can predict fluctuations in exchange rates reliably is a topic that has long attracted intense research efforts in exchange rate economics. At the core of this research is the evidence, first put forward by Meese and Rogoff (1983), that a random walk model often provides more accurate forecasts than fully-fledged models of exchange rate determinants (e.g., monetary models). Meese and Rogoff (1983) attributed their findings to small sample estimation biases, model misspecification (including unexplained nonlinearities), and parameter instability. Three decades later, in spite of the availability of larger samples and more sophisticated econometric methods, the evidence increasingly suggests that Meese and Rogoff's (1983) findings have not yet been overturned; see for example Rossi's (2013) survey of the literature. One main unresolved issue relates to the time-varying predictive content of economic fundamentals, which manifests itself in accurate exchange forecasts in certain specific periods, but not in others (Cheung, Chinn, & Pascual, 2005; Rossi, 2013).[2]

This paper employs bootstrap-based methods to reveal the set of exchange rate fundamentals that apply at each point in time. Bootstrapping is a technique for drawing random multiple samples from the existing data with replacement. Thus, it allows us potentially to make sharper inferences about model attributes and other quantities of interest, by examining these quantities across sample replications. Exploiting this feature allows us to infer, for instance, how well our estimation method fits multiple areas of an exchange rate model, defined according to the fundamental that it includes. The strength of the inference is generally strong when the observed sample is a good approximation of the true unobserved population.

E-mail address: pinho.ribeiro@bancomoc.mz.

[1] Present address: Banco de Mocambique, Av. 25 de Setembro n^o 1695, Maputo, Mozambique.

[2] Cheung and Chinn (2001) report survey evidence which suggests that this time-evolving relationship reflects the market participants' changing views on the factors that drive exchange rate movements. Bacchetta and Van Wincoop (2004, 2013) formalize the idea in a scapegoat model of exchange rate determination. Fratzscher, Rime, Sarno, and Zinna (2015) find mixed support for the out-of-sample forecasting ability of the scapegoat model, with the random walk producing more accurate forecasts according to the mean-squared prediction error metric, but failing on the basis of the direction of change metric. Berge (2014) also considers an idea implied by the scapegoat model, but fails to uncover predictive ability at the one-month horizon.

http://dx.doi.org/10.1016/j.ijforecast.2017.05.004
0169-2070/

Given both the documented time-varying predictive ability of exchange rate fundamentals and concerns about model misspecification, bootstrapping per se may not provide a complete econometric solution for pinning down informative exchange rate fundamentals. We deal with this by embedding a model selection and validation procedure in our bootstrap methods. One variant of the bootstrap method that we employ produces forecasts using the models selected in the various bootstrap samples, then averages the forecasts across sample replications. This approach was introduced by Breiman (1996), who called it bootstrap aggregation or bagging. An alternative bootstrap method that we adopt produces forecasts using the single best model revealed, which is then trained across bootstrap samples. We term it bumping, following its initial advocates (Tibshirani & Knight, 1999). To the best of our knowledge, we are the first to apply these techniques in exchange rate economics. However, we are aware that bagging has previously been applied to the forecasting of inflation, unemployment, stock returns, and hedge funds; see Inoue and Kilian (2008), Jin, Su, and Ullah (2014), Panopoulou and Vrontos (2015), and Rapach and Strauss (2010, 2012).[3]

Despite their mutual foundation on the bootstrap, it is quite different aspects of these two methods that have the potential to improve the forecast accuracy. The bagging technique is designed to improve the performances of unstable forecasting procedures, defined as those in which the forecasts differ substantially across sample replications (Breiman, 1996; Buhlmann & Yu, 2002). Bumping, on the other hand, is intended for procedures that yield many local optima for a specific target criterion (Tibshirani & Knight, 1999). Under a minimum squared prediction error target, for instance, bumping can lead to improvements if the prediction errors from many exchange rate models differ by only a narrow margin across samples.

In our application, we juxtapose the performances of our bootstrap-based methods and a set of other competing methods. In this set, all of the methods rely on the single sample realization for forecasting, or for selecting and combining forecasts. These include: (i) simple linear regressions that are conditioned on each fundamental, (ii) combination methods based on the mean, the median, the trimmed mean, the discounted mean squared prediction error (DMSPE), and log-score weights, (iii) Bayesian model averaging (BMA) or selection (BMS) of the sort considered by Wright (2008), (iv) shrinkage estimators such as the LASSO (Tibshirani, 1996) and the elastic net (Zou & Hastie, 2005), and (iv) the kitchen-sink regression of Welch and Goyal (2008). To facilitate our comparison, the forecasts from all methods are normalized to those of the driftless random walk (RW). As Rossi (2013) noted, the RW constitutes the toughest benchmark to beat in the exchange rate literature.

Focusing on five OECD bilateral currency rates against the U.S. dollar, and a monthly dataset spanning the period from January 1989 to May 2013, we use our methods to forecast recursively at the one-month horizon. As Rossi's (2013) survey reveals, exchange rate predictability is challenging to detect at this specific horizon. Our study employs standard and more recently propounded exchange rate fundamentals, including (i) those from the Taylor (1993) rule, (ii) the Nelson and Siegel (1987) relative factors from yield curves, and (iii) factors extracted from exchange rates as per Engel, Mark, and West (2015). We evaluate the statistical performances of our methods relative to the RW using the root mean squared forecast error (RMSFE), complemented with the Cheung et al.'s (2005) direction of change statistic. Inference on statistical significance is based on the Clark and West (2006) test for RMSFE, and a studentized version of the Diebold and Mariano (1995) test for the direction metric.

This paper further assesses whether our methods generate economically significant gains in a stylized dynamic asset allocation strategy. Inspired by Della Corte, Sarno, and Sestieri (2012) and Li, Tsiakas, and Wang (2015), we compute the fee that a risk-averse investor with a quadratic utility would be willing to pay to use our methods instead of the RW. In addition, we implement both the performance measure of Goetzmann, Ingersoll, Spiegel, and Welch (2007) and the Sharpe ratio. We investigate whether the differences in Sharpe ratios are significant by employing the bootstrap method of Ledoit and Wolf (2008). Finally, we consider Han's (2006) break-even transaction costs that render an investor indifferent as to whether they use our methods or the RW.

In our main findings, bumping reveals sets of economic fundamentals that have strong and significant predictive power for exchange rates. From a statistical perspective, it is the only method that improves upon the RW for at least three exchange rates, irrespective of the measure used. Moreover, the improvement is not ephemeral, but is detectable throughout the forecast sample. In terms of economic gains, bumping leads to Sharpe ratios of up to 0.79, which is significantly different from those of the RW (0.37). The gains associated with bumping are also verifiable over the out-of-sample period. None of the other methods surpass the performance of bumping, although BMA and the Nelson-Siegel relative curvature factor tend to outperform the RW according to the direction of change metric and economic measures. However, BMA's performance weakens after the 2008 financial crisis. The bagging method improves upon the RW for up to two currencies, but not significantly so. An inspection of the performances of our bootstrap-based approaches shows that although they both benefit from the bootstrap's ability to sharpen the inference, bagging tends to over-fit, whereas bumping robustly selects parsimonious models with good out-of-sample predictive power.

The next section sets out our econometric methodology, focusing on (i) our bootstrap-based methods, (ii) the competing approaches that we consider, (iii) the set of fundamentals or predictors, (iv) the data and forecasting approaches (v) the metrics for statistical evaluation, and (vi) the criteria for economic evaluation. Section 3 reports the empirical results, including the main features that underlie the performance of bumping. We then conduct robustness checks in Section 4, and conclude in Section 5.

[3] Applications of the bumping method are more common in machine learning, statistics, and medical science.