Discussion of "Can Parameter Instability Explain the Meese-Rogoff Puzzle", by Bacchetta, van Wincoop and Beutler

Domenico Giannone
Université Libre de Bruxelles, ECARES and CEPR

September 16, 2009

1 Introduction

The paper by Bacchetta, van Wincoop and Beutler (BWB) brings fresh air on a long debated issue in international macroeconomics.

BWB investigate whether exchange rate unpredictability is due to instabilities of the relationship between exchange rates and its fundamentals. In their exercise, BWB calibrate on actual data a model where the parameters linking exchange rates and fundamentals are allowed to change over time. BWB find that the pattern of out-of-sample (un)predictability assessed on data generated from a fixed coefficient model can roughly reproduce the features observed in the data. In addition, BWB find no significant differences in out-of-sample accuracy when data are generated by time varying rather than fixed coefficients. Significant impact of parameter instability is found only when shift in parameters are persistent but in this case parameters instability increases rather than reducing predictability. On the basis of these findings the authors conclude that exchange rate unpredictability is due to the weakness rather than to the instability of the relationship between exchange rates and fundamentals.

I like the idea of the paper, but I am unconvinced about the authors' conclusions. My main point is that Bacchetta, van Wincoop and Beutler (BWB) do not properly account for parameters instability.

BWB calibrate the model used for simulations in a way that does not allow for permanent shifts in the parameters. In their set up, the variance of parameter's innovations (σ^2_β) tends to become smaller when parameter's autocorrelation (ρ^2_β) increases. Permanent shifts in the parameters are ruled out by construction because the unconditional variance of the parameters ($\frac{\sigma^2_\beta}{1-\rho^2_\beta}$) is kept fixed and hence the model collapses to fix coefficients when structural changes tend to have permanent effects (i.e. $\sigma_\beta \to 0$ when $\rho_\beta \to 1$). In this set up the authors show that persistent parameter instability does not worsen forecasting accuracy but on the contrary it induces substantial improvements.

However, there is no reason to think that structural changes are temporary. On the contrary, significant changes in the macroeconomic environment are very likely to last. In fact, in empirical works there is a widespread consensus to model time varying parameters as random walks (see Cogley and Sargent, 2001, 2005; Primiceri, 2005). It would be a mistake to immediately conclude from the results of BWB that the Meese-Rogoff puzzle is explained exclusively by the small sample estimation bias. Before drawing this conclusion, it is necessary to appropriately study the case in which parameters shift permanently. This is the task I will undertake in this discussion.

I will consider a model in which structural changes are permanent by assuming that the parameters linking exchange rates and fundamentals evolve as random walks. Instead of resorting to calibration, which relies on ad-hoc assumptions, the time varying model used for simulation will be fully estimated. The design of any model with time varying coefficients is rather problematic since it is hard to distinguish between strength and instabilities of the relationship between exchange rates. To overcome this problem the estimation is performed using Bayesian techniques and the allowed amount of time variation will be controlled for by setting the prior variance on the coefficient's innovations.

I will focus on forecasting the Euro/US dollar exchange rate using relative prices as fundamentals. Since fundamentals are likely to play an important role in explaining medium to long term fluctuations, the analysis will be performed by using annual data and taking into account the eventual dynamic adjustment to PPP equilibrium. BWB, instead, focus on short horizon forecasts (one-month-ahead) which are produced by exploiting only contemporaneous (within the month) correlations and neglecting the dynamic adjustment to long run equilibria.

As a consequence BWB are likely to overemphasize the weakness of the relationship between the exchange rates and fundamentals.

2 Forecasting the Euro

Let s_t be the (log) exchange rate between the Euro and the US dollar[1] and $\tilde{p}_t = p_t^{ea} - p_t^{us}$ be the (log) relative consumer prices between the Euro Area (ea) and the United States (us)[2]. In order to take into account the common trend between relative prices and the nominal exchange rate I will consider a bivariate vector autoregressive (VAR) model for the real exchange rate $q_t = s_t - \tilde{p}_t$ and the inflation differential $\tilde{\pi}_t = \tilde{p}_t - \tilde{p}_{t-1}$. The sample ranges from 1975 to 2007. Prior to 1999 I will consider Germany and the Deutsche Mark instead of the Euro Area and the Euro. Data are plotted in Figure 1.

Denoting, $y_t = [q_t, \tilde{\pi}_t]'$ we have:

$$y_t = A_0 + A_1 y_{t-1} + e_t, \quad \varepsilon_t \sim N(0, \Sigma)$$

For simplicity, I will focus on the forecast of the real exchange rate q_t. Qualitative results are confirmed when forecasting the exchange rate itself. The exercise goes as follows. Let us forecast first the Euro/Dollar exchange rate in 1999, the year of the introduction of the Euro. Parameters are estimated using data up to 1998 and samples of different length L. The shorter sample includes $L = 10$ years of data, from 1989 until 1998. The longest estimation sample starts in 1975 and includes $L = 24$ years of data. The estimated parameters are used to compute exchange rate forecasts. As in BWB we focus on predictions that are conditional on actual future fundamentals, i.e. assuming that relative prices, $\tilde{p}_t$, from 1999 onward were perfectly foreseen.[3] The forecast are compared with the actual value of the real exchange rate in 1999. The same exercise is repeated every year to produce one year ahead forecasts. Accuracy is measured by averaging the square forecast errors over the evaluation sample $1999 - 2007$.

[1]The source is OECD, National Accounts.
[2]The source is OECD, Main Economic Indicators (MEI).
[3]Qualitative results are confirmed when looking at unconditional predictions.

Figure 2 reports the mean square forecast error produced by the model relative to the random walk forecast (rMSFE). Results are plotted against the length L of the estimation sample. Numbers smaller than one indicate that model forecasts are more accurate that the random walk forecasts. Numbers larger than one indicate that forecast accuracy cannot be improved relative to the Naive benchmark by exploiting information contained in the fundamentals. Looking at the performances for different estimation samples provides interesting insights on the trade off between estimation error and structural instability since estimating the model using a few (many) years of data provides an insurance against model instability but at the same time it implies larger (smaller) parameter uncertainty.

With the shortest sample length the mean squared error of the model based forecasts is twice the mean squared error of the random walk. Expanding the estimation sample first improves forecasting accuracy indicating a reduction of parameter's estimation error. With a sample between 16 and 22 years around years model based forecast outperform the random walk with maximum improvements of 40% when including 20 years of data for the estimation. When the sample is further increased forecast accuracy deteriorates and the advantage of model based forecast relative to the random walk are lost, suggesting that the gains from reduced estimation error are counterbalanced by losses due to the presence of structural instabilities.

In summary, results indicate the out-of-sample accuracy does not improve monotonically when increasing the estimation sample but has a u-shape signaling the presence of a trade-off between parameter instability and parameter uncertainty. In addition, for some estimation window the model based forecasts are more accurate than the random walk forecasts.[4]

3 Inspecting the role of structural instabilities

In order to assess the role played by structural instabilities in accounting for exchange rate unpredictability, I will estimate a VAR model with time varying coefficients developed by Primiceri (2005). The model offers a parsimonious representation of prominent features of structural changes since it provides reliable descriptions of key macroeconomic aggregates (see Cogley, Primiceri, and Sargent, 2008) and accurate out-of-sample predictions (see D'Agostino,

[4]These results are in line with Molodtsova and Papell (2009) who also find some evidence of exchange rate predictability using a wider range of models and countries.

Gambetti, and Giannone, 2008).

I assume that y_t admits the following time varying coefficients VAR(1) representation:

$$y_t = A_{0,t} + A_{1,t} y_{t-1} + \varepsilon_t \tag{1}$$

where $A_{0,t}$ contains time-varying intercepts, $A_{1,t}$ are matrices of time-varying coefficients, and ε_t is a Gaussian white noise with zero mean and time-varying covariance matrix Σ_t. Let $A_t = [A_{0,t}, A_{1,t}]$, and $\theta_t = vec(A_t')$, where $vec(\cdot)$ is the column stacking operator. Conditional on such an assumption, θ_t is assumed to follow a random walk:

$$\theta_t = \theta_{t-1} + \omega_t \tag{2}$$

where ω_t is a Gaussian white noise with zero mean and covariance Ω. Let $\Sigma_t = F_t D_t F_t'$, where F_t is lower triangular, with ones on the main diagonal, and D_t a diagonal matrix. Denote by σ_t the vector of the diagonal elements of $D_t^{1/2}$ and $\phi_{i,t}$, $i = 1, ..., n-1$ the column vector formed by the non-zero and non-one elements of the $(i+1)$-th row of F_t^{-1}. The standard deviations, σ_t, are assumed to evolve as geometric random walks, belonging to the class of models known as stochastic volatility. The simultaneous relations ϕ_{it} in each equation of the VAR are assumed to evolve as independent random walks.

$$\log \sigma_t = \log \sigma_{t-1} + \xi_t \tag{3}$$

$$\phi_{i,t} = \phi_{i,t-1} + \psi_{i,t} \tag{4}$$

where ξ_t and $\psi_{i,t}$ are Gaussian white noises with zero mean and covariance matrix Ξ and Ψ_i, respectively. Let $\phi_t = [\phi_{1,t}', \ldots, \phi_{n-1,t}']$, $\psi_t = [\psi_{1,t}', \ldots, \psi_{n-1,t}']$, and Ψ be the covariance matrix of ψ_t. $\psi_{i,t}$ is assumed to be independent of $\psi_{j,t}$, for $j \neq i$. In addition, ξ_t, ψ_t, ω_t, ε_t are assumed to be mutually uncorrelated at all leads and lags.

The model is estimated using Bayesian methods. The prior densities are set by following Primiceri (2005). Details are reported in the appendix. Time variation is controlled for by setting a prior model in which the standard deviation of parameter's innovation is assumed to be a given percentage λ of the standard deviation of the coefficients estimated by maximum likelihood using a pre-sample including the first 10 years of data (1975-1984). In order to study the effects of parameter instabilities I will work with two prior models: 1) a prior of moderate time variation ($\lambda = 10\%$); 2) a prior of substantial time variation ($\lambda = 50\%$).

Figures 3 and 4 report the posterior mode of the autoregressive coefficients $\hat{A}_{1,t}$ and the 68% coverage intervals. When the prior allows for moderate time variation the estimated coefficients do not vary substantially along the sample. Significant time variations are found when more substantial time variation is allowed. The estimated coefficients are most of the time significantly different from zero indicating that there are significant dynamic linkages between the exchange rate and relative prices.

I draw 1000 times the model parameters from their posterior density. For each parameter's draw I simulate the path the real exchange rate and relative prices. Using the simulated data I perform an out-of-sample forecasting evaluation by mimicking the out-of-sample real time forecasting exercise performed in the previous section.

Let us consider first results when simulating the data from the model estimated using a prior of moderate time variation are reported in Figure 5. I report the median, the 16th and 84th percentiles, across simulations, of the relative mean squared forecast error. The relative mean square forecast error obtained on actual data is reported for comparison. It is evident that moderate time variation has some difficulties in replicating the pattern of the relative mean square forecast error obtained when using actual data. The simulation model cannot account for the u-shape since with simulated data forecast accuracy monotonically improves when the estimation sample become longer. For short and long estimation windows the relative mean square error obtained using actual data is at the boundary of the bands indicating that the model implies a higher predictability than the one observed in the data. A similar pattern is obtained when the simulation is based on the model with fixed, instead of moderately time varying, coefficients.

The features of the data are better captured when we simulate the data from the model estimated using a prior that allows for a substantial amount of parameter instabilities (Figure 6). In particular, the mean square forecast error is now well in the middle of the bands. The simulated model is also able to partially reproduce the deterioration of forecasting accuracy for large estimation samples. Comparing figure 5 and figure 6, it is evident that the out-of-sample forecasting accuracy deteriorates when data are generated by more substantial time variation. This is in contrast with BWB who claim that persistent time variation improves forecast accuracy.

In summary, results point out that estimation uncertainty alone cannot explain the pattern of forecast accuracy in the actual data. To match the data a substantial amount of time variation is needed.

4 Conclusions

In this discussion I have proposed an alternative empirical exercise to that presented by the authors where I consider permanent rather than temporal structural change.

Contrary to what found by the authors, I find some predictability for Euro-Dollar exchange rate using relative prices as fundamentals. Parameter instability and estimation uncertainty are both relevant since the accuracy of exchange rate forecasts tends to deteriorate when the estimation sample becomes too large. In addition, only when allowing for substantial parameter instability it is possible to match the patterns of forecast accuracy found in actual data.

These features have been overlooked by BWB since their simulation exercise has not been properly designed for investigating long lasting structural changes and does not take into account dynamic linkages over the medium to long term, the horizon at which fundamentals are expected to play a more relevant role in exchange rate determination.

The exercise performed in this discussion is rather stylized and a number of issues are still open. Using the TV-VAR model for forecasting is an interesting and promising route for improving the accuracy of exchange rate predictions.

References

Cogley, T., G. E. Primiceri, and T. J. Sargent (2008): "Inflation-Gap Persistence in the U.S.," NBER Working Paper 13749, NBER.

Cogley, T., and T. J. Sargent (2001): "Evolving Post WWII U.S. Inflation Dynamics," in *Macroeconomics Annual*, pp. 331–373. MIT Press.

——— (2005): "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII US," *Review of Economic Dynamics*, 8, 262–302.

D'Agostino, A., L. Gambetti, and D. Giannone (2008): "Macroeconomic Forecasting and Structural Change," manuscript, Université Libre de Bruxelles.

Molodtsova, T., and D. H. Papell (2009): "Out-of-sample exchange rate predictability with Taylor rule fundamentals," *Journal of International Economics*, 77(2), 167–180.

Primiceri, G. E. (2005): "Time Varying Structural Vector Autoregressions and Monetary Policy," *Review of Economic Studies*, 72, 821–852.

5 Appendix

The model setting and on the estimation are accurately described in Primiceri (2005). This appendix briefly describes the specification of our priors. First, the coefficients of the covariances of the log volatilities and the hyperparameters are assumed to be independent of each other. The priors for the initial states θ_0 of the time varying coefficients, simultaneous relations ϕ_0 and log standard errors $\log\sigma$ are assumed to be normally distributed. The priors for the hyperparameters, Ω, Ξ and Ψ are assumed to be distributed as independent inverse-Wishart. More precisely, the priors are as follows:

- Time varying coefficients: $P(\theta_0) = N(\hat{\theta}, \hat{V}_\theta)$ and $P(\Omega) = IW(\Omega_0^{-1}, \rho_1)$;
- Stochastic Volatilities: $P(\log\sigma_0) = N(\log\hat{\sigma}, I_n)$ and $P(\Psi_i) = IW(\Psi_{0i}^{-1}, \rho_{3i})$;
- Simultaneous relations: $P(\phi_{i0}) = N(\hat{\phi}_i, \hat{V}_{\phi_i})$ and $P(\Xi) = IW(\Xi_0^{-1}, \rho_2)$;

where the scale matrices are parameterized as follows $\Omega_0^{-1} = \lambda_1\rho_1\hat{V}_\theta$, $\Psi_{0i} = \lambda_{3i}\rho_{3i}\hat{V}_{\phi_i}$ and $\Xi_0 = \lambda_2\rho_2 I_n$. The hyper-parameters are calibrated using a time invariant recursive VAR estimated using a pre-sample consisting of the first ten years of data (1975-1984). For the initial states θ_0 and the contemporaneous relations ϕ_{i0}, the means, $\hat{\theta}$ and $\hat{\phi}_i$, and the variances, $\hat{V}_\theta$ and $\hat{V}_{\phi_i}$, are set to be the maximum likelihood point estimates and four times its variance. For the initial states of the log volatilities, $\log\sigma_0$, the mean of the distribution is chosen to be the logarithm of the point estimates of the standard errors of the residuals of the estimated time invariant VAR. The degrees of freedom for the covariance matrix of the drifting coefficient's innovations are set to be equal to 10, the size of the pre-sample. The degrees of freedom for the priors on the covariance of the stochastic volatilities' innovations, are set to be equal to the minimum necessary for insuring the prior is proper. In particular, ρ_1 and ρ_2 are equal to the number of rows Ξ_0^{-1} and Ψ_{0i}^{-1} plus one respectively. The parameters λ_i are very important since they control the degree of time variations in the unobserved states. The smaller such parameters are, the smoother and smaller are the changes in coefficients. The results reported in the paper are obtained by setting: a) $\lambda_1 = 1/10^2$, $\lambda_2 = 1/10$ and $\lambda_3 = 1/10^2$ in the prior with moderate time variation ($\lambda = \frac{1}{10}$); b) $\lambda_1 = 1/2^2$, $\lambda_2 = 1/2$ and $\lambda_3 = 1/2^2$ for the prior with substantial time variation ($\lambda = \frac{1}{2}$)

Figure 1: ***The data***

Figure 2: ***The relative Mean Square Forecast Error***

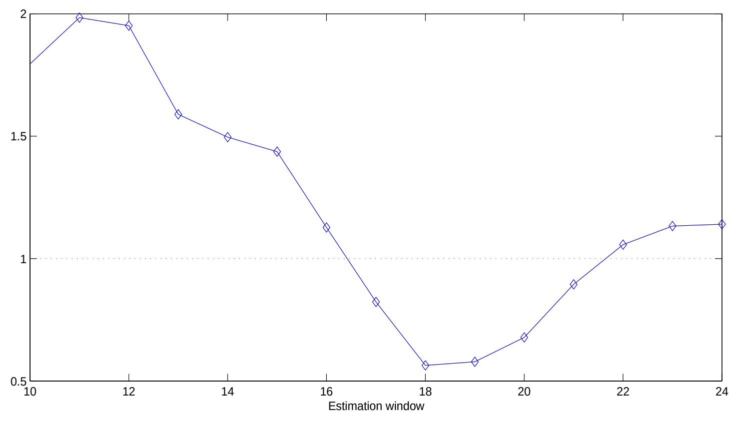

Figure 3: ***The estimated time varying coefficients** ($\lambda = 10\%$)*

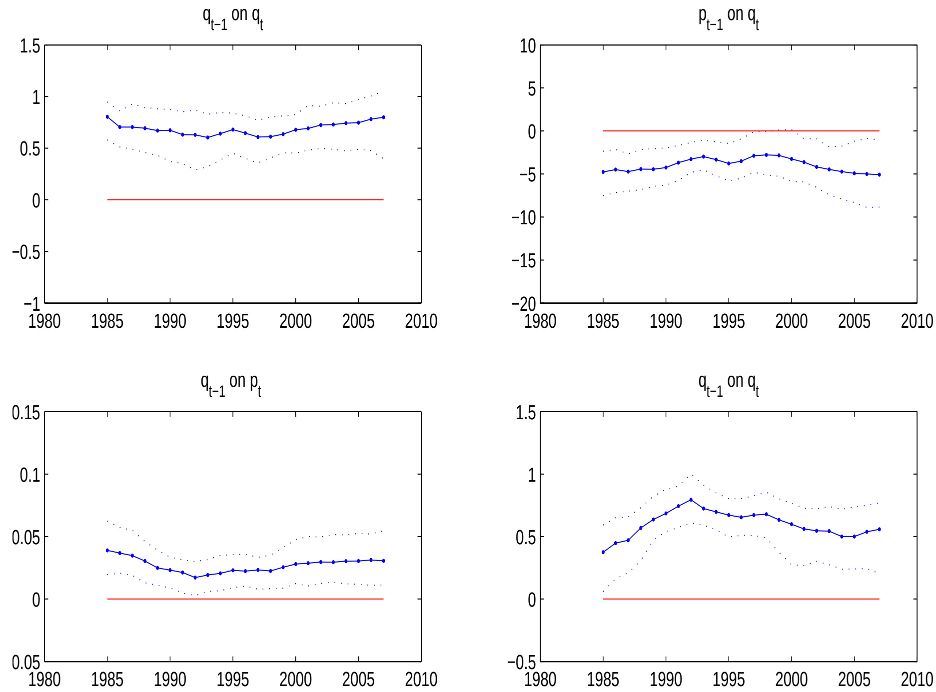

Figure 4: ***The estimated time varying coefficients** ($\lambda = 50\%$)*

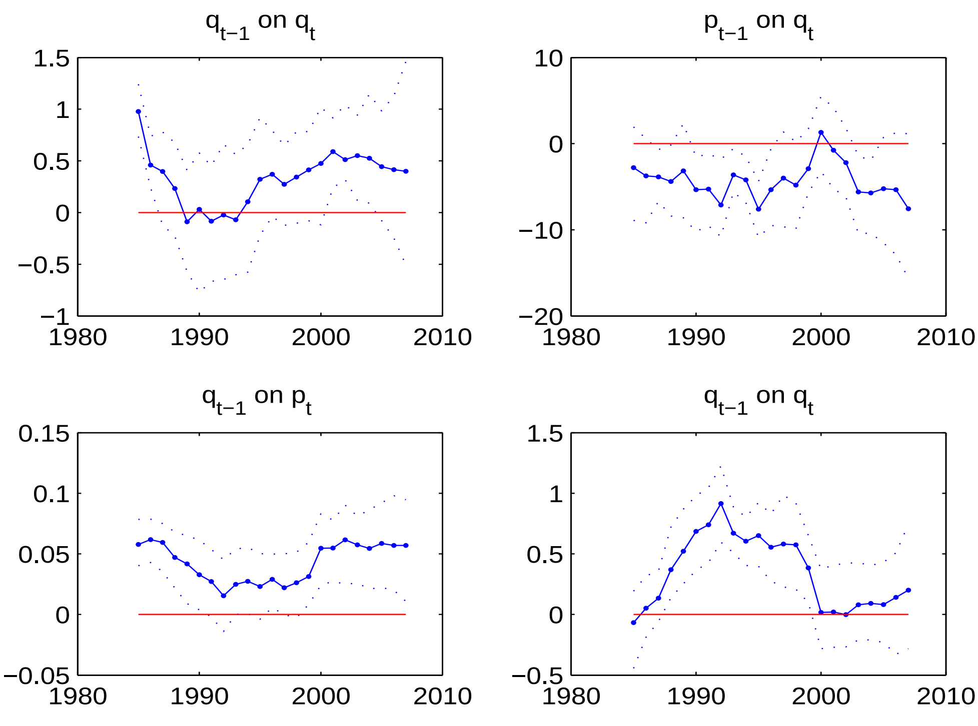

Figure 5: ***The relative Mean Square Forecast Error***: *actual data and data simulated from the time varying model estimated using a prior of moderate time variation.* ($\lambda = \frac{1}{10}$)

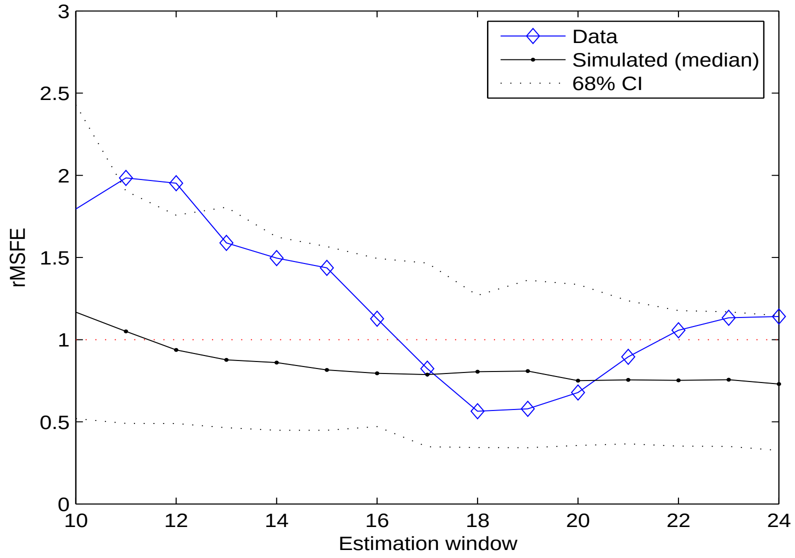

Figure 6: ***The relative Mean Square Forecast Error***: *actual data and data simulated from the time varying model estimated using a prior of substantial time variation.* ($\lambda = \frac{1}{2}$)