ISSN 1518-3548

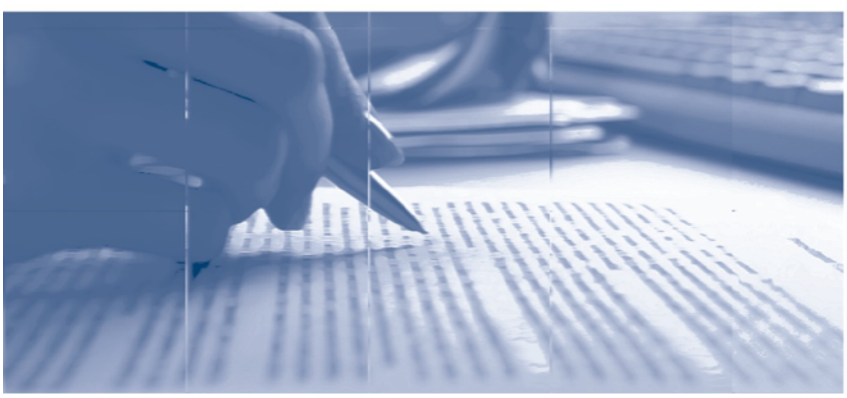
BANCO CENTRAL DO BRASIL

Risk Assessment of the Brazilian FX Rate

Wagner Piazza Gaglianone e Jaqueline Terra Moura Marins

January, 2014

Working Papers



344

ISSN 1518-3548
CGC 00.038.166/0001-05

Working Paper Series	Brasília	n. 344	January	2014	p. 1-43

Working Paper Series

Edited by Research Department (Depep) – E-mail: workingpaper@bcb.gov.br

Editor: Benjamin Miranda Tabak – E-mail: benjamin.tabak@bcb.gov.br
Editorial Assistant: Jane Sofia Moita – E-mail: jane.sofia@bcb.gov.br
Head of Research Department: Eduardo José Araújo Lima – E-mail: eduardo.lima@bcb.gov.br

The Banco Central do Brasil Working Papers are all evaluated in double blind referee process.

Authorized by Carlos Hamilton Vasconcelos Araújo, Deputy Governor for Economic Policy.

General Control of Publications

Banco Central do Brasil

Comun/Dipiv/Coivi

SBS – Quadra 3 – Bloco B – Edifício-Sede – 14º andar

Caixa Postal 8.670

70074-900 Brasília – DF – Brazil

Phones: +55 (61) 3414-3710 and 3414-3565

Fax: +55 (61) 3414-1898

E-mail: editor@bcb.gov.br

Citizen Service Division

Banco Central do Brasil

Deati/Diate

SBS – Quadra 3 – Bloco B – Edifício-Sede – 2º subsolo

70074-900 Brasília – DF – Brazil

Toll Free: 0800 9792345

Fax: +55 (61) 3414-2553

Internet: <http//www.bcb.gov.br/?CONTACTUS>

Risk Assessment of the Brazilian FX Rate*

Wagner Piazza Gaglianone†

Jaqueline Terra Moura Marins‡

Abstract

In this paper, we construct several multi-step-ahead density forecasts for the foreign exchange (FX) rate based on statistical, financial data and economic-driven approaches. The objective is to go beyond the standard conditional mean investigation of the FX rate and (for instance) allow for asymmetric responses of covariates (e.g. financial data or economic fundamentals) in respect to exchange rate movements. We also provide a toolkit to evaluate out-of-sample density forecasts and select models for risk analysis purposes. An empirical exercise for the Brazilian FX rate is provided. Overall, the results suggest that no single model properly accounts for the entire density in all considered forecast horizons. Nonetheless, the GARCH model as well as the option-implied approach seem to be more suitable for short-run purposes (until 3 months), whereas the survey-based and some economic-driven models appear to be more adequate for longer horizons (such as one year).

Keywords: Exchange rate, Density Forecasts, Risk Management.

JEL Classification: C14, C15, C53, E37, F31.

*The authors thank Emanuel Kohlscheen, Luiz Renato Lima, Osmani Guillén, Ricardo Schechtman and Waldyr Areosa for their helpful comments and suggestions. The opinions expressed in this article are those of the authors and do not necessarily reflect those of the Banco Central do Brasil. Any remaining errors are ours.

†Research Department, Banco Central do Brasil. E-mail: wagner.gaglianone@bcb.gov.br

‡Research Department, Banco Central do Brasil. E-mail: jaqueline.terra@bcb.gov.br

1 Introduction

The foreign exchange (FX) rate market is one of the most important in the financial system. According to the recent report of the Bank for International Settlements (BIS, 2013), trading in foreign exchange markets averaged \$5.3 trillion per day in April 2013.[1] Besides its huge trading volume, it also represents the largest asset class in the world leading to high liquidity.[2] Other features of this market are the high volatility and the potential variety of factors that might affect exchange rates (e.g. economic fundamentals, speculative transactions and currency interventions, among many others).

Forecasting exchange rates is of great importance for economic agents, in particular, for investors and policy makers. Accurate forecasts of FX rates allow investors, for instance, to design adequate trading strategies and to hedge against market risk. On the other hand, central banks worldwide closely monitor the daily FX movements, since it can impact future price dynamics and, thus, help setting the appropriate interest rate policy (Groen and Matsumoto, 2004). Besides, it is a useful information for central bankers to decide for interventions.

In practical terms, however, accurately forecasting the FX rate has proved to be a nontrivial exercise. The failure of standard economic theory to explain foreign exchange rate behaviour using key economic fundamentals (such as the money supply, trade balance and national income) has prevailed in the international economics literature since the classical papers of Meese and Rogoff (1983a,b). The authors investigated the out-of-sample forecasting performance of standard exchange rate models during the post-Bretton Woods period and concluded that such models do not perform better than a naive random walk (RW) forecast.[3] Indeed, the macroeconomic theory has proposed several potential predictors of exchange rates (usually based upon the Purchasing Power Parity (PPP) hypothesis, the Uncovered Interest Rate (UIP) parity condition and the monetary model). However, the forecasting contribution of such approaches has been under question since the highly influential findings of Meese and Rogoff. In this sense, Bacchetta and van Wincoop (2006) describe the RW paradigm as "*...the major weakness of international macroeconomics.*"[4]

As a consequence, an extensive literature has studied the forecasting performance of empirical exchange rate models and several (potential) explanations have been put forward. Just to mention a few papers: Mark (1995) finds evidence of greater predictability of economic exchange rate models at longer horizons, although these findings have been

[1] According to the same report, it is up from \$4.0 trillion in April 2010 and \$3.3 trillion in April 2007.

[2] Nonetheless, in the long run, the attractiveness of carry trade strategies relative to other investments is not clear. Indeed, there is a large literature that started with Burnside et al. (2006), which suggests that market frictions greatly reduce the profitability of currency speculation strategies.

[3] The random walk forecast is such that the (log) level of the nominal exchange rate is predicted to remain at the current (log) level (also known as the "no change" forecast).

[4] On Meese and Rogoff, see also the recent working paper by Rossi (2013).

questioned later by Kilian (1999). Kilian and Taylor (2003) argue that exchange rates can be predicted from economic models after taking into account the possibility of nonlinear exchange rate dynamics. Cheung et al. (2005) examine the out-of-sample performance of the interest rate parity, monetary, productivity-based and behavioral exchange rate models and conclude that (indeed) none of these models consistently beats the RW forecast at any horizon. The authors argue that even if a particular macroeconomic "fundamental" has some level of predictive power for a bilateral exchange rate (at a certain horizon), the same variable may show no predictive power at different horizons or for other bilateral exchange rates. On the other hand, Engel and West (2005) argue that it is not surprising that a random walk forecast outperforms fundamental-based models under some circumstances. The argument is based on the behavior of the exchange rate as an asset price within a rational expectation present-value (Taylor rule) model with a discount factor near one.[5]

In a distinct but complementary approach, several papers in the late 90s started investigating the random walk paradigm from a different view: out-of-sample density forecasting. For instance, Diebold, Hahn, and Tay (1999) use the RiskMetrics model to compute half-hour-ahead density forecasts for Deutschmark–dollar and yen–dollar returns. Christoffersen and Mazzotta (2005) construct option-implied density and interval forecasts for four major exchange rates. Clews et al. (2000) describes a non-parametric method to forecast risk neutral densities, based on the smile interpolation of option prices. Boero and Marrocu (2004) obtain one-day-ahead density forecasts for euro nominal effective exchange rate using self-exciting threshold autoregressive (SETAR) models. Sarno and Valente (2009) use information from the term structure of forward premia to evaluate the exchange rate density forecasting performance of a Markov-switching vector equilibrium correction model. Hong et al. (2007) construct half-hour-ahead density forecasts for euro–dollar and yen–dollar exchange rates using a broad set of univariate time series models that capture fat tails, time-varying volatility and regime switches.

In general, these previous studies on exchange rate density forecasting use high frequency data, which are not available for most conventional economic fundamentals. In addition, these studies quite often do not consider multi-step-ahead forecasts and, generally, assume that conditional densities are analytically constructed (i.e. based on parametric densities). Wang and Wu (2012) tackle these issues by using a semiparametric method, applied to a group of exchange rate models, to generate out-of-sample exchange

[5] As complementary lines of research, see also the following papers: Wu (2008) studies the importance of the order flows at short horizons, within the "microstructure approach". Engel et al. (2009) based on a panel of exchange rates argue that in the presence of stationary, but persistent, unobservable fundamentals, long-horizon predictability prevails in FX rate forecasting. Della-Corte et al. (2009) discuss the forward premium and its promising results in a portfolio allocation framework. Chen and Tsang (2009) find that cross-country yield curves are useful in predicting exchange rates. Molodtsova and Papell (2009) extend the standard set of exchange rate models by incorporating Taylor rule fundamentals. More recently, Fratzscher et al. (2012) investigate the scapegoat theory (as an attempt for explaining the poor performance of traditional models), and Morales-Arias and Moura (2013) explore forecast combination based on panel data and adaptive forecasting.

rate interval forecasts. The authors suggest that economic fundamentals might provide useful information in (out-of-sample) forecasting FX rate distributions. Based on forecast intervals for ten OECD exchange rates, the authors find that, in general, FX models generate tighter forecast intervals than the random walk, given that their intervals cover out-of-sample exchange rate realizations equally well. Moreover, the results suggest a connection between exchange rates and fundamentals: economic variables (indeed) contain information useful in forecasting distributions of exchange rates. in this sense, the Taylor rule model (Molodtsova and Papell, 2009) performs better than the monetary, PPP and forward premium models, and its advantages are more pronounced at longer horizons.

In this paper, we also go beyond point forecasting and follow the previous strand of literature focused on density forecasting. We address the subject by considering statistical approaches (such as GARCH), economic-driven models, and also setups based on financial data (treating the exchange rate as an asset price). Instead of using high frequency data and focusing on very short-run horizons, we employ monthly data, that enable us to investigate standard macroeconomic models of (point and density) forecast, constructed here from both parametric and semiparametric setups.

In addition, based on a set of density forecasts, generated for a full range of forecast horizons (from one to twelve months), we go a step further and ask the following question: Which model is the best one (among the considered set of models) for a given forecast horizon, and a given part of the conditional distribution of the FX rate? The objective here is to increase our understanding of exchange rate dynamics from a risk-analysis perspective. In other words, our objective is to investigate risk measures of FX-rate generated from distinct approaches, which may reveal potential links between exchange rates and economic fundamentals (or financial variables) that a simple point forecast evaluation may neglect. This way, our main contribution is to bring together a whole set of statistical tools, from distinct strands of the literature (e.g. international economics, density forecasting and risk management) to investigate FX-rate tail risk, through the lens of competing models.

This paper is organized as follows: Section 2 presents the density forecast models and the respective estimation schemes. Section 3 presents our empirical exercise to investigate the Brazilian FX rate, based on a set of (pseudo) out-of-sample multi-step-ahead density forecasts. The estimated densities are investigated both from a global (entire density) as well as a local perspective (based on value-at-risk analysis and risk management tools). Section 4 concludes.

2 Methodology

2.1 Density Forecast Models

Along this paper, we investigate $m = 1, ..., 5$ models to construct the point (and density) forecasts for the nominal exchange rate (s_t) of the Brazilian Real with respect to the U.S. dollar (R$/US$).[6] Following the notation of Wang and Wu (2012), a general setup of the (point forecast) model m takes the form of:

$$s_{t+h} - s_t = \mathbf{X}'_{m,t}\boldsymbol{\beta}_{m,h} + \varepsilon_{m,t+h} \tag{1}$$

in which $s_{t+h} - s_t$ is the h-periods change of the (log) exchange rate and $\mathbf{X}'_{m,t}$ is a vector with economic variables used in model m. Regarding multi-period ahead forecasts ($h > 1$), notice that we follow the "direct forecast" approach, in contrast to the "recursive (or iterated) forecast" route. See Marcellino, Stock and Watson (2006)[7] for a good discussion.

Table 1 - Exchange rate (point and density) forecast models

Model	Label	Covariate vector $X'_{m,t}$	Density Forecast
1	GARCH - Monte Carlo simulation	$[1; \Delta s_t]$	parametric
2	Option-implied (RND-RWD)	$-$	nonparametric
3	Random walk (without drift)	$[1]$	semiparametric
4	Survey-based expectation	$[1; s^e_{t+1\|t}]$	semiparametric
5	Relative PPP model	$[1; \Delta q_t]$	semiparametric

Notes: Covariate vectors are presented for h=1. RND stands for risk-neutral density and RWD to real world density. The $s^e_{t+1|t}$ term refers to the survey-based expectation of the exchange rate of period t+1 formed at period t, and the real exchange rate q_t is defined as $q_t \equiv s_t + p^*_t - p_t$ in which $p_t(p^*_t)$ is the (log) consumer price index in the home (foreign) country.

Model 1 is a GARCH-Monte Carlo simulation model. It is a backward-looking form, improved by some variance reduction techniques employed over the traditional random sampling simulation method. After the estimation of different specifications[8], the one that has better adjusted the data was the AR(1)-Gaussian GARCH(1,1), with "Descriptive

[6] The term "model" is used throughout this paper in a broad sense that includes forecasting methods.

[7] "Iterated" multi-period ahead time series forecasts are made using a one-period ahead model, iterated forward for the desired number of periods, whereas "direct" forecasts are made using a horizon-specific estimated model, where the dependent variable is the multi-period ahead value being forecasted. Which approach is better is an empirical matter: in theory, iterated forecasts are more efficient if correctly specified, but direct forecasts are more robust to model misspecification.

[8] AR(1)-GARCH(1,1) t-student, random walk with drift and gaussian white noise, random walk with gaussian GARCH and random walk with t-student GARCH. The sampling simulation methods combined with each one of these models were Simple Random Sampling, Simple Random Sampling with runs, Latin Hypercube and Descriptive Sampling.

Sampling" as the simulation method. It can be represented as below ($h = 1$):

$$\Delta s_t = \alpha + \beta \Delta s_{t-1} + \eta_t \tag{2}$$

$$h_t^2 = \omega + \gamma h_{t-1}^2 + \delta \eta_{t-1}^2, \tag{3}$$

where s_t is the log of the nominal exchange rate, h_t^2 is the conditional variance and η_t is the input variable of the simulation model assumed to be standard normally distributed and descriptive sampled instead of randomly sampled.[9]

Model 2 is based on financial data and the extraction of information from option prices.[10] It consists basically of two major steps: (i) obtaining risk-neutral densities (RND); and (ii) transforming these densities into real world densities (RWD). The RND for an asset price gives the set of probabilities that investors would attach to the future asset prices in a world in which they were risk-neutral. But if investors are risk-averse (as they usually are), risk premia will drive a wedge between the probabilities inferred from options (RND) and the true probabilities they attach to alternative values of the underlying asset price (RWD).

The first step follows a method proposed by Shimko (1993) which is a non-parametric technique for extracting RND from option prices based on the construction of a implied volatility curve for the option via interpolation of its strike prices (the "smile volatility curve"). Shimko's method was developed for stock option prices and we had adapted it for exchange rates. The adaptation used the Black Model for pricing future price options (Black, 1976).[11] Breeden and Litzenberger (1978) derived an explicit relationship between the risk-neutral density of an asset and the price of the option on that asset, as expression below:

$$\frac{\partial^2 C_t}{\partial K_t^2} = e^{-r_t T} f(s_t), \tag{4}$$

in which C_t is the (call) option price of an underlying asset s_t, K_t is the respective exercise (strike) price of the referred option, r_t is the risk-free interest rate, T denotes the number of days to maturity, and $f(s_t)$ is the risk-neutral probability density (RND) of the underlying asset s_t. Shimko obtained the densities from this formula by interpolating the calculated implicit volatilities for the same maturity and different exercise prices. To do so, one

[9] For more details about "Descriptive Sampling" and other sampling methods for variance reduction, see Saliby (1989) and Glasserman (2004).

[10] The main idea is that options are contracts giving the right (but not the obligation) to buy or sell an asset at a given point in the future at a price set now (the strike price). Options to buy (call options) are only valuable if there is a chance that when the option comes to be exercised, the underlying asset will be worth more than the strike price. This way, if one considers options to buy a particular asset at a particular point in the future but at different strike prices, the prices at which such contracts are trade now provides some information about the market's view of the chances that the price of the underlying asset will be above the various strike prices. Therefore, options tell us something about the probability the market attaches to an asset being within a range of possible prices at some future date.

[11] If the underlying asset of the future contract is the exchange rate, the Black Model becomes equivalent to the Garman-Kohlagen Model for pricing exchange rate options.

must generate an entire continuum of values for the relation of the option price versus its exercise price, given that only a few points of this curve are indeed known.

In this sense, Shimko proposes a quadratic interpolation of the implied volatilities associated with each existing exercise price. From this new curve of implied volatilities, the continuum of values for the option price is obtained, allowing the calculation of second derivatives and, thus, the respective densities. In this paper, risk neutral densities for future exchange rates were generated only for one, two and three months ahead ($h = 1, 2, 3$), due to the low liquidity of exchange rate-based options (and the lack of available data) for longer maturities.

The second step follows a method proposed by Vincent-Humphreys and Noss (2012). Instead of the commonly used method of applying utility-function transformations to the RND, these authors propose an empirical and less restrictive methodology. They use a parametric Beta distribution function to calibrate the difference between the RND and the RWD. According to the authors, although the Beta distribution is parsimonious, as it depends on only two parameters, it nests many simple forms of transformation, such as mean shift, mean-preserving changes in variance and changes involving mean, variance and skewness.

Models 3 is a benchmark model of random walk without drift. **Model 4** is a forward-looking survey-based expectation (i.e., consensus forecast), with a bias correction device as proposed by Capistrán and Timmermann (2009). **Model 5** is a weaker version of the PPP model (so-called relative PPP). The density forecasts of models 3, 4 and 5 are generated by using a semiparametric approach based on quantile regression, as suggested by Gaglianone and Lima (2012).[12] The idea is to use a location-scale specification to construct density forecasts from the covariate vector $\mathbf{X}'_{m,t}$. This way, the conditional specification for models 3, 4 and 5 with mean and variance dynamics is defined as

$$s_{t+h} - s_t = \mathbf{X}'_{m,t}\boldsymbol{\alpha}_{m,h} + \left(\mathbf{X}'_{m,t}\boldsymbol{\beta}_{m,h}\right)\zeta_{t+h}, \tag{5}$$

$$\zeta_{t+h}|\mathcal{F}_t \sim F_{\zeta,h}\left(0,1\right) \tag{6}$$

where $F_{\zeta,h}\left(0,1\right)$ is some distribution with mean zero and unit variance, which depends on h but does not depend on the information set $\mathcal{F}_t$. $\mathbf{X}'_{m,t} \in \mathcal{F}_t$ is a $k \times 1$ vector of covariates, and $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are $k \times 1$ vectors of parameters which include the intercepts α_0 and β_0. This class of data-generating process (DGP) is very broad and includes most common volatility processes such as ARCH and stochastic volatility. Based on the previous model and using standard quantile regression techniques (see Koenker, 2005), the conditional quantiles of $s_{t+h} - s_t$ are given by

$$Q_{m,\tau}(s_{t+h} - s_t \mid \mathcal{F}_t) = \mathbf{X}'_{m,t}\boldsymbol{\theta}_{m,h}(\tau) \tag{7}$$

[12] The authors generate multi-step-ahead conditional density forecasts for the unemployment rate in the U.S. from (point) consensus forecasts and quantile regression; which is a setup that do not impose any parametric structure on the shape of the conditional distributions.

where, for a given quantile level $\tau \in [0;1]$, it follows that $\boldsymbol{\theta}_{m,h}(\tau)$ is a $k \times 1$ vector of parameters of the form $\theta_i(\tau) = \left(\alpha_i(\tau) + \beta_i(\tau) F_{\zeta,h}^{-1}(\tau)\right); i = 1, ..., k$. Given a family of estimated conditional quantiles $Q_{m,\tau}(\cdot)$, the conditional density of $s_{t+h} - s_t$ can be easily estimated by using the Epanechnikov kernel, for instance, which is a weighting function that determines the shape of the bumps.

To guarantee monotonicity of the conditional quantiles (and the validity of the related conditional distribution), by avoiding possible crossing of quantiles, some rearrangement procedure (e.g., He, 1997; Chernozhukov et al., 2010) could be further employed.[13]

Finally, it is worth mentioning that, besides the five models above, an extended set of economic-driven models is also considered (see Table A.1 in appendix), based on Molodtsova and Papell (2009) and Wang and Wu (2012), although their results are not reported in this paper to save space.[14] The objective here is not to propose the best model to forecast the FX rate, but rather to discuss risk management implications of a given set of available models to (density) forecast the foreign exchange rate, within a range of forecast horizons.

2.2 Forecast Evaluation

The forecast evaluation is conducted in this paper throughout distinct perspectives: Firstly, for illustrative and comparison purposes, we do a standard point forecast evaluation, focused on the forecast performance of the central part (location) of the estimated densities. To do so, we compute the root mean squared error (RMSE) and check whether it is possible to beat the random walk forecast for a given forecast horizon, based on the Diebold-Mariano-West (1995, 1996) test and also on the Harvey et al. (1997) modified test.

Secondly, the density forecast evaluation is conducted along two dimensions: (i) Global analysis, which is a shape evaluation based on the entire estimated density. Following the literature on density forecast evaluation, we investigate coverage rates; the Berkowitz (2001) test; the model ranking from the log predictive density scores (LPDS); and the Amisano-Giacomini (2007) test; (ii) Local analysis, which evaluates specific points of the densities, interpreted here as Value-at-Risk (VaR) measures. To do so, we employ available risk management tools for VaR backtesting based on the tests of Kupiec (1995); Christoffersen (1998); and Gaglianone et al. (2011). The details of each evaluation procedure will be presented in next section together with the results of our empirical exercise.

[13] He (1997) pointed out that crossing problem occurs more frequently in multiple-variable regressions. Thus, we should not expect crossing to be an issue in our empirical exercise due to the reduced number of covariates in models 3, 4 and 5.

[14] In this sense, we only show the results for the relative PPP model, which generally presented the best results among the extended set of specifications.

3 Empirical Exercise

3.1 Data

The currency we consider is the Brazilian Real (R$) foreign exchange rate in respect to the U.S. dollar (US$). Our choice of country reflects our intention to examine exchange rate behavior for one of the largest emerging economies. The models are estimated using monthly data from January 2000 through December 2012 (156 observations), with flexible exchange rate over the considered sample.[15] The exchange rate is defined as the Brazilian Real price of a unit of U.S. dollar currency, so that an increase in the exchange rate denotes a depreciation of the Real currency.

We use data over the period January 2000-December 2004 for model estimation (training sample) and reserve the remaining data for (pseudo) out-of-sample forecasting. We construct (point and density) forecasts for horizons $h = 1, ..., 12$ months. This way, the evaluation sample for $h = 1$ is January 2005-December 2012 (96 out-of-sample forecasts), whereas for $h = 12$ we have 85 out-of-sample forecasts.

To evaluate the performance of the models, we estimate them by using both recursive estimation (increasing sample size) as well as rolling window estimation (with a fixed sample of five years = 60 monthly observations).[16] See Morales-Arias and Moura (2013) for a detailed discussion about rolling window and recursive forecasting.

Figure 1 - Exchange rate R$/US$

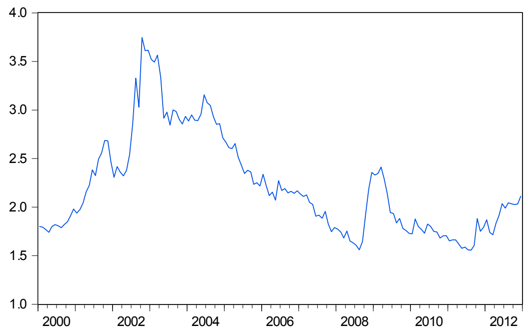


Figure 1 presents the behavior of the Brazilian FX rate along the considered sample. In the second semester of 2002, the exchange rate experienced a sharp increase due to

[15] The monthly (nominal) exchange rate is given by the sale rate (R$/US$) at the beginning of each month (Sisbacen PTAX800). The results for the end-of-month and daily-averaged rates lead to very similar conclusions. The FX-rate data is obtained from the website of the Central Bank of Brazil. For model 2, we use the BM&F's reference prices for dollar calls. For model 4, we employ survey-based (consensus) expectations from the Focus survey, organized by the Central Bank of Brazil, which collects daily information on more than 100 institutions, including commercial banks, asset-management firms, and non-financial institutions. For model 5, we also use data from the FRED dataset of the Federal Reserve Bank of St. Louis.

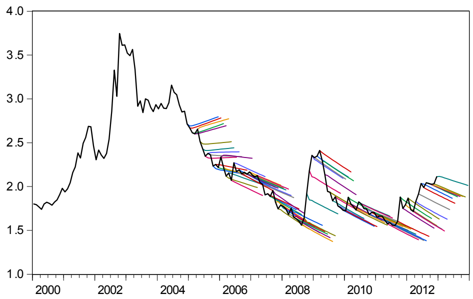
[16] Each model is initially estimated using the first 60 observations and the one-period-ahead (up to the twelve-month-ahead) point and density forecasts are generated. We, then, drop the first data point, add an additional observation at the end of the sample, re-estimate the models and generate again out-of-sample forecasts. This process is repeated along the remaining data.

(among other factors) the augment of investors uncertainty regarding the future of economic fundamentals after the presidential elections in October 2002; exhibiting a relatively consistent appreciation pattern in the remaining periods, only interrupted by the global crisis in 2008 and the recent devaluation trend initiated in 2012.

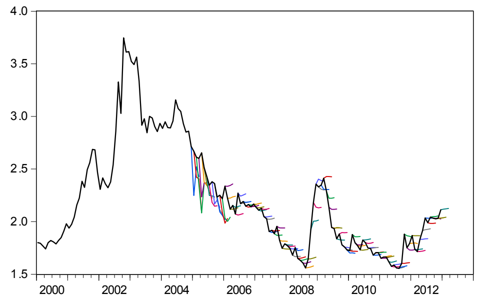
3.2 Point Forecasts

We start the model evaluation by investigating the performance of the exchange rate point forecast across the investigated models. Next figure summarizes the conditional point forecasts of selected models for $h = 1, ..., 12$ estimated along the pseudo out-of-sample exercise. Table 2 shows the respective Root Mean Squared Error (RMSE) obtained from the forecast prediction errors along the mentioned exercise. Models are labeled "a" or "b" according to the sample used in model estimation (recursive estimation or rolling window).

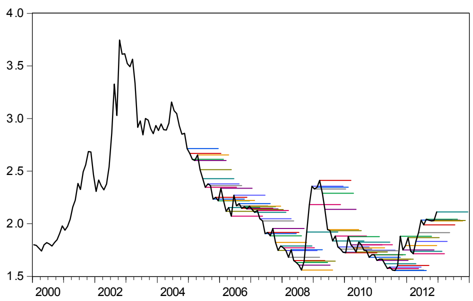
Figure 2 - Point forecasts of selected models



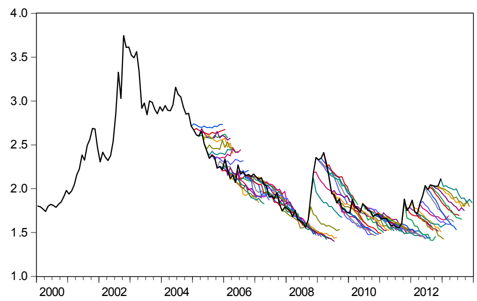
Point forecasts from Model 1, rolling window estimation



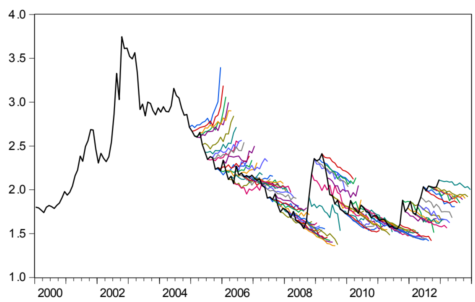
Point forecasts from Model 2



Point forecasts from Model 3, rolling window estimation



Point forecasts from Model 4, rolling window estimation



Point forecasts from Model 5, rolling window estimation

Is it possible to beat the random walk? To tackle this question we employ the Diebold-Mariano-West (1995, 1996) test of equal accuracy (see Tables 2 and 3).[17] The null hypothesis assumes equal RMSEs of two competing models. Positive test statistics indicate that the considered (alternative) model has a lower RMSE than the benchmark (random walk) model.[18]

The results based on the recursive estimation indicate that the only model which exhibits a positive DMW statistic is model 4 (for horizons ranging from 1 to 6 months). In other words, only the model which embodies survey-based expectations (model 4) is able to present a lower RMSE in comparison to the random walk.

The results for a rolling window scheme are similar. In both cases, the mentioned "gain" over the random walk is only statistically significant at horizons $h = 1$ or 2 months. Indeed, the gain for $h = 1$ is of 106% and 108%, for the recursive and rolling window estimations, respectively. For $h = 2$ these figures drop to 15% and 17%, respectively.

This result is in line with a vast literature reporting the practical difficulty on beating the naive random walk forecast in pseudo out-of-sample exercises (see Mark, 1995).[19] In our case, the results suggest that information content from survey expectations might improve short-term point forecasts.

In appendix, we also present the results of the DMW test modified by Harvey et al. (1997), which propose a hypothesis test more suitable to small sample sizes. The results are quite similar.

Table 2 - RMSE results

Recursive estimation						Rolling window				
RMSE (Root Mean Squared Error)										
Model	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
h=1	0.09	0.12	0.09	0.04	0.09	0.09	0.12	0.09	0.04	0.09
h=2	0.13	0.15	0.13	0.12	0.14	0.14	0.15	0.13	0.11	0.14
h=3	0.17	0.18	0.17	0.16	0.18	0.18	0.18	0.17	0.16	0.19
h=4	0.21		0.21	0.20	0.22	0.21		0.21	0.20	0.22
h=5	0.24		0.24	0.24	0.25	0.25		0.24	0.24	0.26
h=6	0.27		0.27	0.27	0.29	0.28		0.27	0.27	0.29
h=7	0.29		0.29	0.30	0.33	0.30		0.29	0.30	0.34
h=8	0.31		0.31	0.31	0.35	0.32		0.31	0.31	0.36
h=9	0.32		0.32	0.33	0.37	0.33		0.32	0.34	0.39
h=10	0.33		0.33	0.36	0.41	0.34		0.33	0.36	0.42
h=11	0.34		0.35	0.43	0.48	0.35		0.35	0.43	0.49
h=12	0.35		0.36	0.46	0.55	0.37		0.36	0.45	0.55

[17] A related empirical question is the following: "Are the competing models better (than the RW) in predicting just the direction of change?" We leave this point forecast comparison for future research.

[18] The table entries are Diebold-Mariano-West t-tests (p-values) of equal RMSEs, taking the random walk (model 3) as the benchmark and model $m \neq 3$ as the alternative. The variances entering the test statistics use the Newey-West estimator, with a bandwidth of 0 at the 1-month horizon and 1.5×horizon in the other cases, following Clark (2011, supplementary appendix) and Clark and McCracken (2012, p.61).

[19] Notice that given the inflation differential (between Brazil and the US) a RW with drift could possibly be even harder to beat.

Table 3 - Diebold-Mariano-West test of equal accuracy

Diebold-Mariano-West (1995, 1996): test statistic (p-value)										
	Model					Model				
h	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
1	-1.38	-1.68		3.95	-0.77	-1.02	-1.68		3.94	-0.84
	(0.17)	(0.1)		(0)	(0.44)	(0.31)	(0.1)		(0)	(0.4)
2	-0.83	-1.06		2.79	-0.88	-0.90	-1.06		2.86	-1.34
	(0.41)	(0.29)		(0.01)	(0.38)	(0.37)	(0.29)		(0.01)	(0.18)
3	-0.90	-0.50		1.28	-1.30	-0.92	-0.50		1.48	-1.14
	(0.37)	(0.62)		(0.2)	(0.2)	(0.36)	(0.62)		(0.14)	(0.26)
4	-0.80			0.36	-1.46	-0.93			0.29	-1.25
	(0.43)			(0.72)	(0.15)	(0.36)			(0.77)	(0.21)
5	-0.79			0.03	-1.22	-0.83			0.05	-1.14
	(0.43)			(0.98)	(0.22)	(0.41)			(0.96)	(0.26)
6	-0.84			0.03	-1.36	-0.79			-0.32	-1.18
	(0.4)			(0.98)	(0.18)	(0.43)			(0.75)	(0.24)
7	-0.95			-0.40	-1.21	-0.85			-0.64	-1.50
	(0.34)			(0.69)	(0.23)	(0.4)			(0.53)	(0.14)
8	-1.04			-0.19	-1.18	-0.86			-0.26	-1.12
	(0.3)			(0.85)	(0.24)	(0.39)			(0.79)	(0.26)
9	-1.03			-0.51	-1.22	-0.96			-0.53	-1.76
	(0.31)			(0.61)	(0.23)	(0.34)			(0.6)	(0.08)
10	-1.02			-0.44	-1.34	-0.95			-0.43	-1.44
	(0.31)			(0.66)	(0.19)	(0.34)			(0.66)	(0.15)
11	-0.96			-0.60	-1.39	-1.00			-0.59	-1.47
	(0.34)			(0.55)	(0.17)	(0.32)			(0.56)	(0.15)
12	-1.06			-0.59	-1.47	-1.02			-0.52	-1.36
	(0.29)			(0.56)	(0.15)	(0.31)			(0.6)	(0.18)

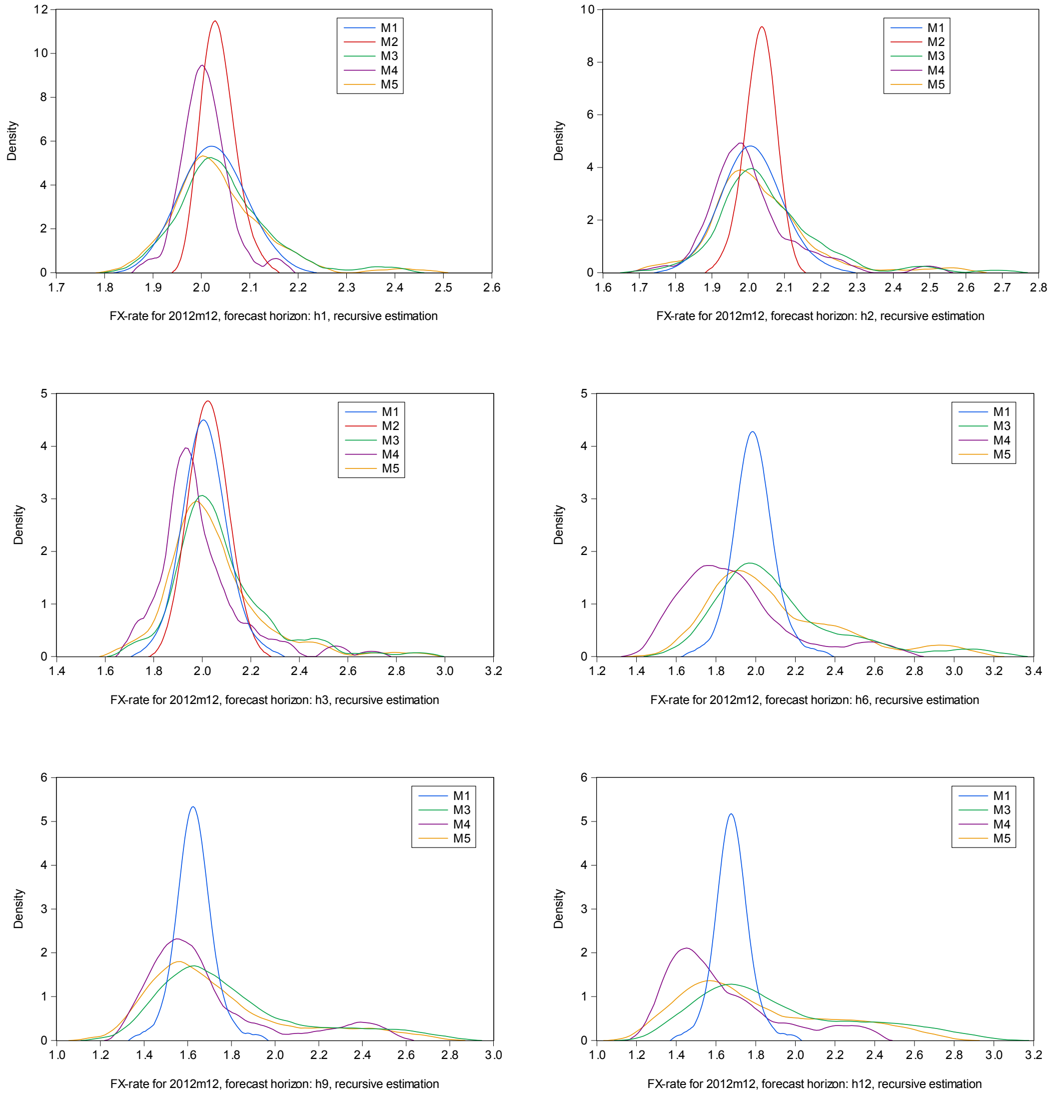
Note: Positive test statistics indicate that a model has a lower RMSE in comparison to the RW.

3.3 Density Forecasts - Global Analysis

Density forecast evaluation has become popular in the fields of time series forecasting and risk evaluation (Ko and Park, 2013) and related formal testing procedures have been developed by several recent studies.[20] We start the global density evaluation by presenting in Figure 3 the estimated conditional Probability Density Functions (PDFs) of the R\$/US\$ exchange rate at December 2012, constructed with different forecast horizons. In appendix, there are complementary graphs for given models and varying forecast horizons as well as PDFs for other selected periods.

[20] A good review of various testing methods in density forecasting is provided by Corradi and Swanson (2006).

Figure 3 - Conditional PDFs of R$/US$ at December 2012



Coverage rates

According to Clark (2011, p.336): "*In light of central bank interest in uncertainty surrounding forecasts, confidence intervals, and fan charts, a natural starting point for forecast density evaluation is interval forecasts - that is, coverage rates.*" In this sense, a necessary (but not sufficient) condition for a "good" density model is to produce a conditional density with an adequate coverage rate.[21] The objective in this section is to check whether the model departures from a given nominal coverage rate (e.g., 70%) appear to be statistically meaningful.

In practice, one needs to compute the frequency of observations of s_{t+h} which have fallen inside the forecast interval. In our case, we adopt the 70% interval band, which leads to a forecast interval based on the conditional quantiles $Q_{m,\tau}(s_{t+h} - s_t \mid \mathcal{F}_t)$ of model m, ranging from quantile level $\underline{\tau} = 0.15$ to $\overline{\tau} = 0.85$. Then, a simple statistical test verifies the equality between the frequency of observations which have fallen in the forecast interval (nominal coverage) and the true coverage. The results are presented in Table 4.

Table 4 - Coverage Rates

	Recursive estimation					Rolling window				
Forecast coverage rates: % of actual outcomes inside the 70% interval band										
	Model					Model				
h	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
1	0.70	0.71	0.66	0.78	0.69	0.71	0.71	0.68	0.69	0.69
	(0.96)	(0.96)	(0.37)	(0.06)	(0.79)	(0.86)	(0.96)	(0.63)	(0.79)	(0.79)
2	0.64	0.59	0.69	0.79	0.71	0.63	0.59	0.64	0.71	0.68
	(0.27)	(0.06)	(0.92)	(0.09)	(0.92)	(0.22)	(0.06)	(0.31)	(0.92)	(0.78)
3	0.55	0.52	0.71	0.78	0.77	0.54	0.52	0.66	0.72	0.72
	(0.01)	(0.01)	(0.82)	(0.22)	(0.24)	(0.02)	(0.01)	(0.52)	(0.73)	(0.72)
4	0.47		0.71	0.77	0.76	0.47		0.63	0.70	0.73
	(0)		(0.9)	(0.28)	(0.33)	(0)		(0.41)	(0.99)	(0.67)
5	0.34		0.76	0.74	0.73	0.45		0.57	0.67	0.71
	(0)		(0.46)	(0.63)	(0.73)	(0)		(0.18)	(0.76)	(0.94)
6	0.30		0.79	0.67	0.78	0.40		0.60	0.63	0.67
	(0)		(0.32)	(0.78)	(0.28)	(0)		(0.29)	(0.49)	(0.76)
7	0.28		0.81	0.63	0.81	0.40		0.59	0.58	0.70
	(0)		(0.27)	(0.55)	(0.18)	(0)		(0.3)	(0.28)	(1)
8	0.19		0.78	0.64	0.79	0.34		0.55	0.61	0.63
	(0)		(0.47)	(0.6)	(0.36)	(0)		(0.11)	(0.41)	(0.56)
9	0.22		0.82	0.65	0.80	0.32		0.57	0.59	0.65
	(0)		(0.21)	(0.67)	(0.28)	(0)		(0.23)	(0.35)	(0.65)
10	0.23		0.76	0.64	0.72	0.24		0.55	0.61	0.69
	(0)		(0.56)	(0.65)	(0.79)	(0)		(0.2)	(0.47)	(0.93)
11	0.23		0.84	0.65	0.69	0.24		0.57	0.63	0.63
	(0)		(0.09)	(0.69)	(0.87)	(0)		(0.3)	(0.57)	(0.54)
12	0.21		0.85	0.61	0.76	0.19		0.53	0.60	0.64
	(0)		(0.03)	(0.49)	(0.39)	(0)		(0.23)	(0.43)	(0.61)

Note: The table includes in parentheses p-values for the null of correct coverage (empirical = nominal rate of 70%), based on t-statistics using standard errors computed with the Newey-West estimator, with a bandwidth of 0 at the 1-month horizon and 1.5×horizon for other cases.

Wang and Wu (2012) point out that random walk intervals are essentially intervals from the unconditional distribution, whereas intervals based on economic models are intervals from conditional distributions. In other words, if the true model is indeed a random walk,

[21] Coverage rates reveal the difference between the probability that realizations fall into the forecasted intervals and the respective nominal coverage.

then, asymptotically, the length of both the RW and the economic model have to be exactly the same.

In our exercise, first note that the random walk[22] and the survey-based approaches (models 3 and 4) as well as the economic-driven model 5 are not rejected at a 5% confidence level in all horizons and both sampling schemes (excepting model 3 for $h = 12$). On the other hand, models 1 and 2 produced an adequate coverage rate for short horizons ($h = 1, 2$ months, in both recursive and rolling window estimations). The rolling window estimation scheme (for shorter horizons, in general) yields slightly more accurate interval forecasts (i.e., coverage rates closer to the 70% nominal rate), in line with the findings of Clark (2011, p.336).[23] As a robustness check, we also present (in appendix) the results for the 50% and 90% interval bands, which (in general) point to similar conclusions.

Probability integral transform (PIT)

The coverage rates although providing an initial approach to analyze density models can be viewed as unconditional tests, since they ignore potential cluster behavior (along the sample size) of a given percentile of the estimated density and, thus, do not take into account time dependence. We next investigate the density forecast models based on a broader measure of density calibration: the probability integral transform (PIT).

According to Clements (2005), the PIT of the realization of the variable with respect to the density forecast is given by

$$z_{t+1} = \int_{-\infty}^{y_{t+1}} \widehat{f}_{t+1,t}(u)du \equiv \widehat{F}_{t+1,t}(y_{t+1}), \tag{8}$$

where $\widehat{F}_{t+1,t}(y_{t+1})$ is the estimated probability of Y_{t+1} not exceeding the realized value y_{t+1}, and $\widehat{f}_{t+1,t}$ is our estimated density forecast of a given model m. The main idea of density forecast evaluation is that, assuming correct specification of the model, the PIT yields independent and uniformly distributed random variables. When the forecast density $\widehat{f}_{t+1,t}$ equals the true density, it follows that $z_{t+1} \sim U(0, 1)$, where $U(0, 1)$ is the uniform distribution over the interval $(0, 1)$. Clements (2005, p.104) argues that even though the actual conditional densities may be changing over time, provided the forecast densities match the actual densities at each t, then $z_t \sim U(0, 1)$ for each t, and the z_t are independently distributed from each other, such that the realized time series $\{z_t\}_{t=1}^{T}$ is an i.i.d. sample from a standard uniform distribution.

[22] Except for h=12 in a recursive estimation scheme.

[23] The referred author also argues that: "For a given model, differences in coverage across horizons likely reflect a variety of forces, making a single explanation difficult. One force is sampling error. Even if a model were correctly specified, random variation in a given data sample could cause the empirical coverage rate to differ from the nominal. Sampling error increases with the forecast horizon, due to the overlap of forecast errors for multistep horizons (effectively reducing the number of independent observations relative to the one-step horizon). Of course, an increased sampling error across horizons will translate into reduced power to detect departures from accurate coverage."

Berkowitz (2001) develops tests to evaluate the conditional density based on the normality of the normalized errors that have better power than tests based on the uniformity of the PITs. The normalized forecast error is defined as $\widetilde{z}_{t+1} \equiv \Phi^{-1}(z_{t+1})$, where z_{t+1} denotes the PIT of a one-step ahead forecast error and Φ^{-1} is the inverse of the standard normal distribution.[24] These tests have been used in recent studies such as Clements (2004), Jore et al. (2010) and Clark (2011).

Table 5 - Density test of Berkowitz (2001)

$Ho : \widetilde{z}_{t+1} \sim iid\ N(0,1)$

	Recursive estimation					Rolling window				
Berkowitz test (p-value)										
Model	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
h=1	0.04	0.49	0.22	0.00	0.82	0.21	0.49	0.17	0.00	0.73
h=2	0.38	0.34	0.01	0.00	0.23	0.54	0.34	0.07	0.05	0.32
h=3	0.36	0.39	0.00	0.00	0.00	0.78	0.39	0.01	0.00	0.01
h=4	0.00		0.00	0.00	0.00	0.57		0.04	0.00	0.00
h=5	0.00		0.00	0.00	0.00	0.45		0.00	0.00	0.00
h=6	0.26		0.00	0.00	0.00	0.89		0.00	0.00	0.00
h=7	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
h=8	0.32		0.00	0.00	0.00	0.64		0.00	0.00	0.00
h=9	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
h=10	0.00		0.00	0.00	0.00	0.00		0.06	0.00	0.00
h=11	0.00		0.00	0.00	0.00	0.00		0.00	0.10	0.00
h=12	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00

For short-term forecast horizons, such as from $h = 1$ to 3 months, models 1 and 2 present a proper density forecast, in both sampling schemes. For medium-term horizons ($h = 4, ..., 8$) there is no single model that passes the Berkowitz test in both estimation schemes. Nonetheless, it is worth mentioning the relative good performance of model 1 in the rolling window framework. For longer horizons ($h = 9, ..., 12$), again, there is no predominant model, although it is worth highlighting that the survey-based approach (model 4) is able to generate adequate forecasts for $h = 11$ when estimated in a rolling window scheme. In the extended set of economic models (results not reported, but available upon request), we noticed that the Taylor rule specification (model II) is also able to generate good density forecasts within a rolling window scheme. This result corroborates the view that the exchange rate might be "too noisy" to be predicted by economic models in the short-run[25] (directly in line with the difficulty in beating the RW model), but might respond to economic fundamentals in the long-run (see Mark, 1995).

Log predictive density score (LPDS)

Another useful indicator of the calibration of density forecasts is given by the log predictive density scores (LPDS). This approach allows us to rank the investigated models based on their log-scores (for each forecast horizon), under which a higher score implies

[24] For h=1 the test statistic (jointly) assumes independence and standard normality for $\widetilde{z}_{t+1}$ and (under the null) it converges to a $\chi^2_{(3)}$ distribution. For h>1, we adopt a modified version of the test (see Jore et al., 2010) using a two degrees-of-freedom variant (without a test for independence).

[25] Excepting, of course, the positive results for very short-run horizons based on high frequency data and microstructure approach (e.g., order flow).

a better model. The LPDS of model m and forecast horizon h is defined in the following way:

$$LPDS_{m,h} = \frac{1}{T} \sum_{t=1}^{T} \ln \left(\widehat{f}_{t+h,t}^{m} \left(y_{t+h} \right) \right) \tag{9}$$

where $\widehat{f}_{t+h,t}^{m}$ is the density of the variable of interest Y_{t+h} estimated from model m and based on the information set available at period t. The referred density is evaluated at the observed value y_{t+h} and (log) averaged along the out-of-sample observations. The LPDS results of Table 6 are summarized in Table 7, which presents the rank order of density models based on such criterion.

Table 6 - Log predictive density score (LPDS)

LPDS (Log Predictive Density Score)

Model	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
h=1	-2.21	-2.19	-1.89	-1.66	-1.87	-2.19	-2.19	-2.00	-1.57	-1.77
h=2	-2.12	-2.19	-2.08	-1.80	-2.07	-2.12	-2.19	-1.97	-1.78	-1.78
h=3	-2.12	-2.27	-2.09	-1.91	-2.19	-2.12	-2.27	-2.05	-1.76	-1.91
h=4	-2.14		-2.26	-1.93	-2.29	-2.13		-2.11	-1.85	-2.03
h=5	-2.12		-2.50	-2.16	-2.53	-2.07		-2.34	-2.11	-2.13
h=6	-2.17		-2.37	-2.31	-2.49	-2.14		-2.28	-2.12	-2.19
h=7	-2.18		-2.72	-2.19	-2.70	-2.13		-2.55	-2.16	-2.34
h=8	-2.21		-2.73	-2.28	-2.70	-2.15		-2.64	-2.19	-2.26
h=9	-2.24		-2.81	-2.43	-2.43	-2.17		-2.63	-2.25	-2.13
h=10	-2.29		-2.85	-2.36	-2.83	-2.18		-2.67	-2.15	-2.55
h=11	-2.33		-2.97	-2.40	-2.94	-2.18		-2.73	-2.30	-2.56
h=12	-2.41		-3.00	-2.38	-2.85	-2.23		-2.79	-2.22	-2.54

Note: The table entries are average values of log predictive density scores (see Adolfson, Linde, and Villani, 2005), under which a higher score implies a better model.

Table 7 - Ranking of density models according to LPDS

Rank of models based on LPDS

Model	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
h=1	5	4	3	1	2	4	5	3	1	2
h=2	4	5	3	1	2	4	5	3	1	2
h=3	3	5	2	1	4	4	5	3	1	2
h=4	2		3	1	4	4		3	1	2
h=5	1		3	2	4	1		4	2	3
h=6	1		3	2	4	2		4	1	3
h=7	1		4	2	3	1		4	2	3
h=8	1		4	2	3	1		4	2	3
h=9	1		4	2	3	2		4	3	1
h=10	1		4	2	3	2		4	1	3
h=11	1		4	2	3	1		4	2	3
h=12	2		4	1	3	2		4	1	3

Note: The best two models according to the LPDS rank ordering (i.e. higher LPDS' figures) are highlighted in yellow for each sample scheme.

Overall, the LPDS analysis indicates that models 1, 4 and 5 would be the most recommended for horizons above six months, whereas the results for shorter horizons, although not pointing to a single approach, in general, suggest the survey-based model 4.

Now, we turn to an interesting question regarding the LPDS results: Is it possible to beat the forecast from model 3 when dealing with densities? In order to answer this question, we use the Amisano-Giacomini (2007) test which compares the log score distance between two competing models. Because the theoretical setup of the test proposed by Amisano and Giacomini requires forecasts estimates with rolling samples of data, we only apply the test to the models estimated with the rolling window scheme. The null hypothesis assumes equal LPDS between model 3 and model $m \neq 3$. A negative test statistic indicates a higher LPDS of model m in comparison to the RW approach. In Table 8, the negative figures statistically significant at a 5% significance level are marked in blue.

Table 8 - Amisano-Giacomini (2007) test applied to average LPDS (rolling window)

Amisano-Giacomini (2007): test statistic (p-value)

	Model				
h	1b	2b	3b	4b	5b
1	0.19	0.19		-0.43	-0.24
	(0)	(0)		(0)	(0)
2	0.15	0.22		-0.19	-0.19
	(0)	(0)		(0)	(0)
3	0.07	0.22		-0.29	-0.14
	(0.12)	(0)		(0)	(0)
4	0.02			-0.26	-0.07
	(0.82)			(0)	(0.3)
5	-0.27			-0.23	-0.21
	(0)			(0)	(0.01)
6	-0.14			-0.15	-0.09
	(0.05)			(0.28)	(0.51)
7	-0.41			-0.38	-0.20
	(0)			(0)	(0.01)
8	-0.49			-0.45	-0.38
	(0)			(0)	(0)
9	-0.46			-0.37	-0.50
	(0)			(0)	(0)
10	-0.48			-0.52	-0.12
	(0)			(0)	(0.02)
11	-0.55			-0.43	-0.17
	(0)			(0)	(0.01)
12	-0.56			-0.57	-0.25
	(0)			(0)	(0)

Note: Null hypothesis of zero average difference in LPDS between model 3 (benchmark) and model $m \neq 3$. Similar to Clark (2011), the p-values are computed by regressions of differences in log scores (time series) on a constant, using the Newey-West estimator of the variance of the regression constant (with a bandwidth of 0 at the 1-month horizon and 1.5×horizon for other cases).

Note that the random walk density approach is easily overwhelmed in several cases. Models 1, 4 and 5 showed a relatively superior performance in respect to the naive model 3. Next table summarizes the results of the global analysis.

Table 9 - Summary of the Global Analysis

Horizon	Coverage rate	Berkowitz test	LPDS	AG test
1	1ab,2ab,3ab,4ab,5ab	1b,2ab,3ab,5ab	4ab,5ab	4b,5b
2	1ab,2ab,3ab,4ab,5ab	1ab,2ab,3b,4b,5ab	4ab,5ab	4b,5b
3	3ab,4ab,5ab	1ab,2ab	3a,4ab,5b	4b,5b
6	3ab,4ab,5ab	1ab	1ab,4ab	-
9	3ab,4ab,5ab	-	1ab,4a,5b	1b,4b,5b
12	3b,4ab,5ab	-	1ab,4ab	1b,4b,5b

Notes: Column 2 identifies the density models that presented a p-value above 0.05 in the coverage rate analysis. Column 3 exhibits models that are not rejected (5% significance level) at Berkowitz's test. Column 4 presents the best two models (within each sample scheme) according to the LPDS ranking, and column 5 shows those models that "beat" the random walk density forecast at a 5% significance level.

3.4 Density Forecasts - Local Analysis

Now, we investigate the predictive accuracy of the density models under a local analysis approach. The idea is to check for the performance of distinct parts of the conditional distribution, estimated through different approaches. A given model to generate the whole conditional density of the variable of interest might produce, for instance, an "adequate" risk measure for the left tail of the distribution (i.e., at low percentiles) but, at the same time, can generate "poor" risk measures at the central part (or even at the right tail) of the distribution. For this reason, we next analyze the density models through the lens of their respective performance along a grid of selected quantile levels $\tau = \{0.1, 0.2, ..., 0.9\}$, in order to cover the key parts of the conditional distribution.[26]

A percentile of the conditional distribution, called here by a "conditional quantile", can be viewed as a Value-at-Risk (VaR) measure (see Christoffersen et al., 2001). As point out by Wang and Wu (2012), the VaR is a prevalent risk management tool used by investors. It is essentially a one-sided forecast interval measuring downside risks. For this reason, the forecast evaluation of the selected "slices" of the distribution can also naturally be conducted by using the many statistical tests available in the risk management literature, also known as "backtests" (see Jorion (2007) and Crouhy et al. (2001) for a good review). In this paper, we use four procedures to conduct the local analysis, next described, although it should be mentioned that many more tests are currently available in the literature.[27]

(i) Local Forecast Coverage Rate: The first procedure is the forecast coverage rate ($LFCR_{m,h,\tau}$) of model m and horizon h at quantile level τ. Similar to the coverage rate discussed in the section 3.3, we now compute (for all out-of-sample observations) the percentage of outcomes below a given nominal quantile level τ. Ideally, the empirical $LFCR_{m,h,\tau}$ should be as close as possible to one minus the nominal level τ.

$$\widehat{LFCR}_{m,h,\tau} = \frac{1}{T}\sum_{t=1}^{T} H_{t+h} \tag{10}$$

where $H_{t+h} = \begin{cases} 1 \text{ ; if } y_{t+h} > \widehat{Q}_{m,\tau}(y_{t+h} \mid \mathcal{F}_t) \\ 0 \text{ ; if } y_{t+h} \leq \widehat{Q}_{m,\tau}(y_{t+h} \mid \mathcal{F}_t) \end{cases}$ and $y_{t+h} = s_{t+h} - s_t$. The statistical significance of $LFCR_{m,h,\tau} - (1-\tau)$ is checked via the Kupiec (1995) backtest.

[26] It is worth mentioning that an analysis at extreme quantile levels (e.g., $\tau = 0.995$) is possible within our framework, although would require a much higher number of observations to be used in both models' estimation and the pseudo out-of-sample exercise.

[27] Such as the nonparametric test of Crnkovic and Drachman (1997), the duration approach of Christoffersen and Pelletier (2004), the CAViaR setup of Engle and Manganelli (2004) and the Ljung-Box type-test of Berkowitz et al. (2008), among many others.

(ii) Kupiec (1995): It is a nonparametric test (also known as the unconditional coverage test) based on the proportion of violations H_{t+h}, in which the null hypothesis assumes that:

$$Ho : LFCR_{m,h,\tau} = E(H_{t+h}) = 1 - \tau \tag{11}$$

The probability of observing N violations, in which $y_{t+h} > \widehat{Q}_{m,\tau}(y_{t+h} \mid \mathcal{F}_t)$, over a sample size of T is driven by a Binomial distribution. This way, the null can be tested through a standard likelihood ratio (LR) test of the form:

$$LR_{uc} = 2\ln\left(\frac{\left(\widehat{LFCR}_{m,h,\tau}\right)^{N}(1-\widehat{LFCR}_{m,h,\tau})^{T-N}}{(1-\tau)^{N}(\tau)^{T-N}}\right), \tag{12}$$

which follows (under the null hypothesis) a chi-squared distribution with one degree of freedom.

(iii) Christoffersen (1998): The unconditional coverage test does not provide any information about the temporal dependence of observed violations. In this sense, Christoffersen (1998) extends the previous test to incorporate an evaluation of time independence of the referred violations. To do so, define T_{ij} as the number of days in which a state j occurred in one day, while it was at state i the previous day. The test statistic also depends on π_i, which is defined as the probability of observing a violation, conditional on state i the previous day. The author assumes that the H_{t+h} stochastic process follows a first order Markov sequence. This way, under the null hypothesis of independence it follows that $\pi = \pi_0 = \pi_1 = (T_{01} + T_{11})/T$, and the complementary test statistic can be constructed, as it follows.

$$LR_{ind} = 2\ln\left(\frac{(1-\pi_0)^{T_{00}}\pi_0^{T_{01}}(1-\pi_1)^{T_{10}}\pi_1^{T_{11}}}{(1-\pi)^{(T_{00}+T_{10})}\pi^{(T_{01}+T_{11})}}\right). \tag{13}$$

The conditional coverage test of Christoffersen (1998) has the following joint statistic of unconditional coverage and independence: $LR_{cc} = LR_{uc} + LR_{ind}$. The joint test statistic LR_{cc} is asymptotically distributed as $\chi^2_{(2)}$.

(iv) Value-at-Risk test based on Quantile Regression (VQR test): The previous test has a restrictive feature, since it only takes into account the autocorrelation of order 1 in the violation sequence. Moreover, it clearly ignores the magnitude of violations when comparing the observed figure of y_{t+h} with the estimated conditional quantile $\widehat{Q}_{m,\tau}(y_{t+h} \mid \mathcal{F}_t)$. To overcome these features, Gaglianone et al. (2011) proposed the VQR test in order to evaluate the predictive performance of the estimated Value-at-Risk measure $V_{t+h} \equiv \widehat{Q}_{m,\tau}(y_{t+h} \mid \mathcal{F}_t)$. The VQR test is simply a Wald test based on the following quantile regression:

$$Q_\tau\left(y_{t+h} \mid \mathcal{F}_t\right) = \alpha_0(\tau) + \alpha_1\left(\tau\right)V_{t+h}\ ;\ \tau \in (0;1) \tag{14}$$

Under the null hypothesis that V_{t+h} is an adequate conditional quantile of y_{t+h}, it follows that $V_{t+h} = Q_\tau (y_{t+h} \mid \mathcal{F}_t)$, which can be verified through the following joint coefficient test: $Ho : \alpha_0(\tau) = 0$ and $\alpha_1 (\tau) = 1$.

The results of the four mentioned procedures adopted to (locally) investigate the density models are presented in Table 10 for a one-month-ahead forecast horizon ($h = 1$). The respective results for $h = 2, 3, 6, 9$, and 12 months are presented in appendix.

Table 10 - Local coverage rates and backtests for selected percentiles

Panel A - Forecast coverage rates: % of actual outcomes below the nominal quantile level (tau)

h=1	Model					Model				
tau	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
0.1	0.09	0.10	0.15	0.01	0.11	0.08	0.10	0.15	0.07	0.08
0.2	0.18	0.21	0.28	0.08	0.24	0.18	0.21	0.27	0.14	0.20
0.3	0.30	0.32	0.43	0.15	0.34	0.27	0.32	0.35	0.23	0.31
0.4	0.45	0.40	0.51	0.24	0.46	0.39	0.40	0.53	0.31	0.40
0.5	**0.55**	**0.51**	**0.61**	**0.36**	**0.52**	**0.54**	**0.51**	**0.61**	**0.42**	**0.50**
0.6	0.65	0.59	0.74	0.51	0.68	0.61	0.59	0.68	0.51	0.59
0.7	0.75	0.76	0.79	0.64	0.77	0.71	0.76	0.77	0.64	0.75
0.8	0.82	0.85	0.83	0.77	0.80	0.81	0.85	0.82	0.78	0.79
0.9	0.91	0.89	0.93	0.92	0.93	0.88	0.89	0.91	0.90	0.89

Panel B - Kupiec (1995) test

h=1	p-value for each model					p-value for each model				
tau	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
0.1	0.84	0.89	0.16	0.00	0.64	0.58	0.89	0.16	0.35	0.58
0.2	0.57	0.84	0.06	0.00	0.34	0.57	0.84	0.10	0.10	0.96
0.3	0.96	0.63	0.01	0.00	0.36	0.53	0.63	0.25	0.12	0.79
0.4	0.34	0.93	0.03	0.00	0.25	0.77	0.93	0.01	0.08	0.93
0.5	**0.31**	**0.84**	**0.02**	**0.01**	**0.68**	**0.41**	**0.84**	**0.02**	**0.10**	**1.00**
0.6	0.36	0.90	0.00	0.08	0.12	0.77	0.90	0.12	0.08	0.90
0.7	0.28	0.19	0.04	0.18	0.12	0.86	0.19	0.12	0.18	0.28
0.8	0.57	0.17	0.40	0.48	0.96	0.76	0.17	0.57	0.65	0.84
0.9	0.84	0.64	0.35	0.58	0.35	0.43	0.64	0.84	0.89	0.64

Panel C - Christoffersen (1998) test

h=1	p-value for each model					p-value for each model				
tau	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
0.1	0.00	0.00	0.08	0.00	0.13	0.21	0.00	0.08	0.23	0.09
0.2	0.27	0.39	0.10	0.00	0.46	0.19	0.39	0.12	0.00	0.14
0.3	0.29	0.22	0.01	0.00	0.64	0.39	0.22	0.20	0.00	0.95
0.4	0.00	0.80	0.04	0.00	0.46	0.07	0.80	0.00	0.00	0.90
0.5	**0.05**	**0.90**	**0.01**	**0.00**	**0.65**	**0.31**	**0.90**	**0.01**	**0.01**	**0.72**
0.6	0.12	0.55	0.02	0.03	0.26	0.35	0.55	0.19	0.03	0.99
0.7	0.55	0.30	0.03	0.25	0.16	0.98	0.30	0.16	0.25	0.48
0.8	0.85	0.04	0.20	0.43	0.14	0.92	0.04	0.13	0.64	0.86
0.9	0.96	0.27	0.22	0.41	0.22	0.12	0.27	0.79	0.63	0.13

Panel D - VQR (2011) test

h=1	p-value for each model					p-value for each model				
tau	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
0.1	0.33	0.37	0.43	0.00	0.07	0.14	0.37	0.41	0.49	0.20
0.2	0.16	0.63	0.03	0.00	0.05	0.52	0.63	0.05	0.06	0.13
0.3	0.62	0.86	0.04	0.00	0.05	0.47	0.86	0.13	0.01	0.04
0.4	0.57	0.91	0.25	0.10	0.57	0.41	0.91	0.17	0.18	0.49
0.5	**0.23**	**0.68**	**0.12**	**0.04**	**0.05**	**0.30**	**0.68**	**0.12**	**0.19**	**0.09**
0.6	0.10	0.91	0.00	0.17	0.07	0.13	0.91	0.02	0.56	0.07
0.7	0.16	0.00	0.00	0.38	0.02	0.08	0.00	0.01	0.49	0.01
0.8	0.01	0.00	0.00	0.73	0.00	0.04	0.00	0.00	0.68	0.00
0.9	0.00	0.01	0.00	0.53	0.00	0.02	0.01	0.00	0.51	0.00

A rejection of a given model for a selected horizon and a percentile of the distribution suggests the need for local improvement on the density model, in order to eliminate (for instance) a wrong coverage rate, a clustering behavior or even a poor temporal dynamics. We summarize the local analysis results in terms of acceptance/rejection (at a 5% significance level) on the three considered backtests (Kupiec, Christoffersen, VQR).

In this sense, we aggregate the results in respect to lower quantiles, in which $\tau \in \{0.1, 0.2, 0.3\}$, or higher quantiles, where $\tau \in \{0.7, 0.8, 0.9\}$. The results are presented in Table 11. In other words, for a model to be shown in a given cell of Table 11 (e.g. $h = 1$) it requires a p-value above 0.05 in all the 9 possible cases (3 percentiles x 3 types of test).

For $h = 1$, only a few models passes (at the same time) in the three statistical tests and in all selected percentiles of the respective part of the distribution. For $h = 2$, only model 1a (recursive estimation, lower quantiles) would survive to this restrictive criterion, and for $h > 2$ not a single model would be selected. This way, for $h > 1$, the adopted criterion to select models is weakened as long as the forecast horizon increases.

Table 11 - Summary of the Local Analysis

Horizon	Criteria to Select Models	Low Quantiles ($\tau \leq 0.3$)	High Quantiles ($\tau \geq 0.7$)
1	Kupiec, Christoffersen, VQR	1b,3b,5a	4ab
2	Kupiec, Christoffersen, VQR	1a	-
2	Kupiec, Christoffersen	1a,5ab	4b,5b
3	Kupiec, Christoffersen	5a	-
3	Kupiec	3a,4b,5ab	2ab,4ab,5ab
6	Kupiec	3a,5ab	3b,4ab,5ab
9	Kupiec	3a,5ab	3ab,4a,5a
12	Kupiec	3a,5ab	3b,4a

3.5 Risk Assessment

In this section, we conduct a risk assessment exercise as a complement to the density models' local analysis. To do so, we first establish, for illustrative purposes, the following *ad-hoc* limits of both valuation ($\underline{\lambda}$) and devaluation ($\overline{\lambda}$) of the R\$/USS exchange rate. In the devaluation case, we first construct a dummy variable (D_{t+h}^{upper}) to reveal the periods which (*ex-post*) exhibited a devaluation amount equal (or greater than) the established limit ($\overline{\lambda}$):

$$D_{t+h}^{upper} = \begin{cases} 1 \quad \text{if} \quad \frac{s_{t+h}}{s_t} \geqslant \overline{\lambda} \\ 0 \quad ; \quad \text{otherwise} \end{cases} \tag{15}$$

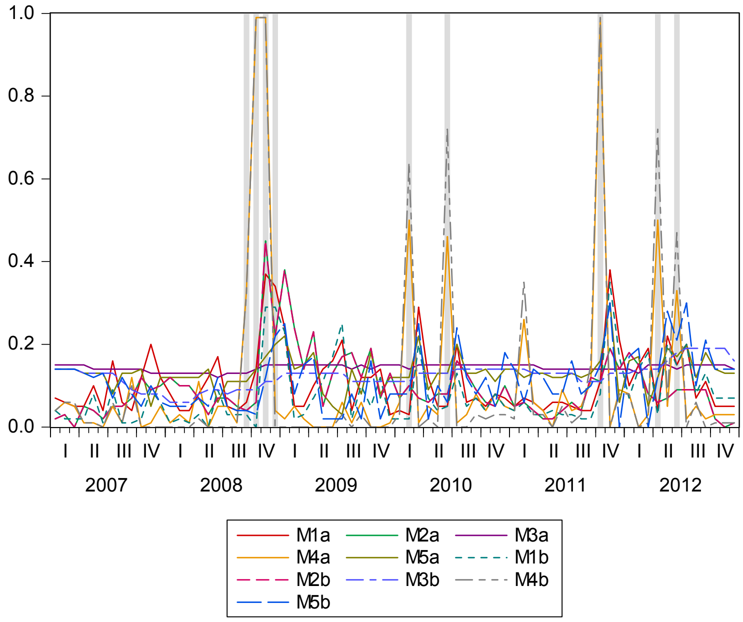
Next, for each considered model m, we search across the respective estimated density (in practice, across the grid of quantile levels $\tau = \{0.01, 0.02, ..., 0.99\}$) for the conditional quantile of s_{t+h} which matches (or is closer to) $\overline{\lambda} s_t$. In other words, we look for the quantile level $\overline{\tau}$ which corresponds to the devaluation limit and, thus, can be interpreted as the *ex-ante* conditional probability that the exchange rate will surpass ($h-$periods ahead) such limit in the future. More formally:

$$\overline{\tau}_{t+h} \equiv \Pr(s_{t+h}/s_t \geqslant \overline{\lambda} \mid \mathcal{F}_t); \text{ where } Q_{m,\overline{\tau}_{t+h}}(s_{t+h} \mid \mathcal{F}_t) = \overline{\lambda} s_t; \ \overline{\tau}_{t+h} \in [0; 1] \tag{16}$$

The next figure shows both time series $\overline{\tau}_{t+h}$ and D_{t+h}^{upper} representing (respectively) the *ex-ante* conditional probability that the FX-rate will breach the limit and the *ex-post* periods when it indeed occurred (gray vertical bars). The procedure for the downside-risk analysis (based on D_{t+h}^{lower}; $\underline{\lambda}$ and $\underline{\tau}_{t+h}$) is similar to the upside-risk analysis just described. We present the risk analysis only for $h = 1$, since it is the horizon with the greatest number of models "approved" by the three (local analysis) backtests.

Figure 4 - Conditional probabilities of (de)valuation of the R$/US$ exchange rate one-month-ahead ($h = 1$)

(a) at least 5% devaluation



(b) at least 5% valuation

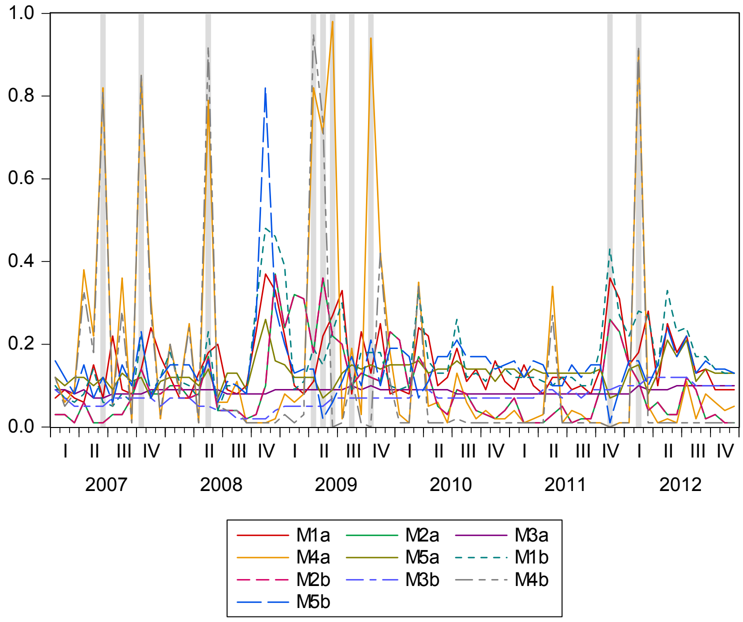
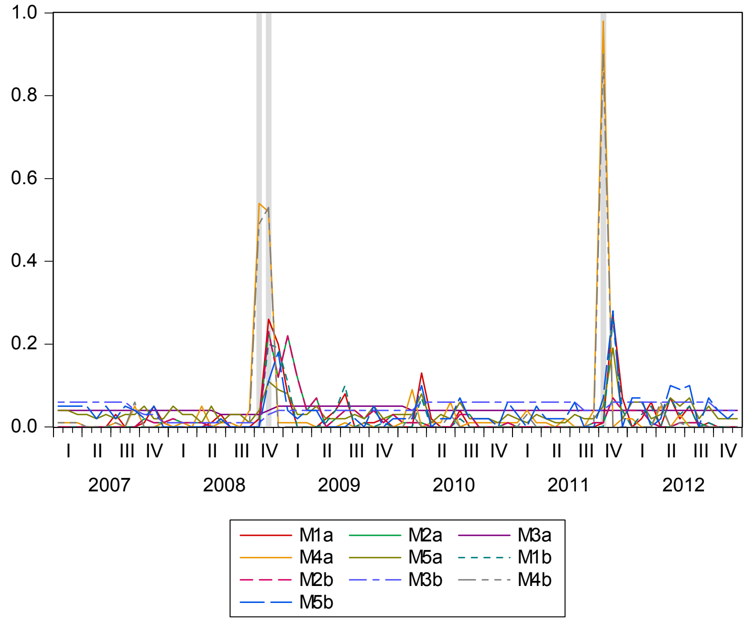
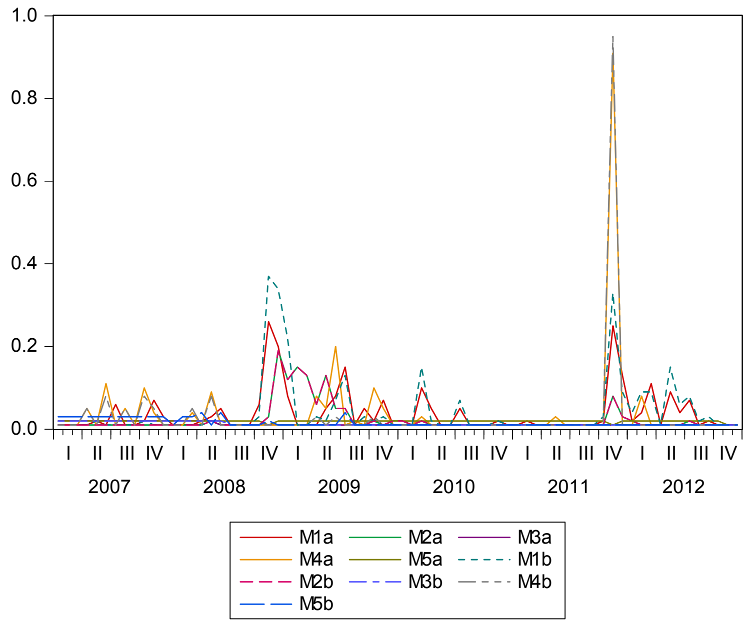


Figure 4 (cont.) - Conditional probabilities of (de)valuation of the R$/US$ exchange rate one-month-ahead ($h = 1$)

(c) at least 10% devaluation



(d) at least 10% valuation



The risk analysis reveals that the forward looking model 4 (based on survey expectations, with a bias correction device, and quantile regression) is relatively able to anticipate FX rate movements. In other words, a forecaster using model 4 would have predicted relatively well the occurrence of a FX rate devaluation equal (or larger than) 5% (or 10%) for next month. This type of tool can be useful for evaluating competing models and selecting those which historically provide more accurate predictions of a given event.

It should be viewed as a complement to the previous local analysis, mainly focused on the performance of fixed conditional quantiles (VaRs) *vis-à-vis* the actual FX rate outcomes. In this section, however, the goal was to compare pre-established FX rate movements in contrast to a set of conditional quantiles estimated at different quantile levels (i.e., the entire estimated density).

Finally, it is worth mentioning that tail risk in Brazil can be greatly affected by official interventions in the FX market, which are not properly captured by any of the investigated models based on monthly data. In this case, a setup with high-frequency data (beyond the scope of this paper) could be further explored (see Kohlscheen and Andrade, 2013).[28]

4 Conclusion

This article has examined several density forecast models for the Brazilian foreign exchange rate through the lens of a set of density forecasting evaluation and risk management tools. We follow a strand of literature which goes beyond the conditional mean analysis and focus on the entire density forecast of the FX rate. The objective is to contribute to the literature by investigating which approaches are more useful in density forecasting exchange rates.

Overall, the results point out to a general conclusion that no single model properly accounts for the entire density in all considered forecast horizons (not even the RW-based model). In fact, the choice of a density forecast model for the FX rate should depend on the part of the conditional distribution of interest as well as on the forecast horizon (short-run x long-run). The reason is that some models are more prone to produce good forecasts at high (or low) percentiles of the FX rate density; which is in line with an asymmetric response of covariates in respect to the exchange rate conditional distribution.

In other words, a given economic fundamental, for instance, which might be useless to explain conditional mean exchange rate dynamics, might be adequate to explain upside (or downside) risks of FX rate at a particular horizon. By focusing on the accuracy of the methods in predicting the likelihood of (de)valuation events we are also able to select models to be used for risk management purposes.

Our contribution is to provide a toolkit to evaluate available models according to its forecasting performance for distinct parts of the distribution and different forecast horizons, trying to bridge the gap between distinct literatures on international economics, density forecasting and financial risk analysis. To do so, we proposed a global analysis as well as a local analysis, which can reveal the suitable models for a determined goal.

Possible extensions of this research could incorporate: (i) other covariates to explain

[28] The authors investigate official interventions in the Brazilian FX market (i.e. currency swap auctions, which are focused on providing hedge to economic agents, liquidity to domestic FX market and reducing excessive market volatility) based on high-frequency data.

FX dynamics in long-run (e.g., commodity price index, as suggested by Kohlscheen, 2013); (ii) additional density models (e.g., GARCH-in-mean); (iii) density forecast combination (Hall and Mitchell, 2007; Jore et al., 2010; Kascha and Ravazzolo, 2010; Gaglianone and Lima, 2014); (iv) risk assessment based on alternative risk measures (Artzner et al., 1999); and (v) microstructure approach based on high frequency data. We leave these extensions as suggestions for future research.

References

[1] Adolfson, M., Linde, J., Villani, M., 2005. Forecasting Performance of an Open Economy Dynamic Stochastic General Equilibrium Model. Sveriges Riksbank Working Paper n.190.

[2] Amisano, G., Giacomini, R., 2007. Comparing Density Forecasts via Weighted Likelihood Ratio Tests. *Journal of Business and Economic Statistics* 25, 177-190.

[3] Artzner, P., Delbaen, F., Eber, J., Heath, D., 1999. Coherent Measures of Risk. *Mathematical Finance* 9, 203-228.

[4] Bacchetta, P., van Wincoop, E., 2006, Can Information Heterogeneity Explain the Exchange Rate Determination Puzzle? *American Economic Review* 96, 552-576.

[5] BIS - Bank for International Settlements, 2013. Triennial Central Bank Survey - Foreign exchange turnover in April 2013: preliminary global results. Available at http://www.bis.org/publ/rpfx13fx.pdf

[6] Berkowitz, J., 2001. Testing Density Forecasts, With Applications to Risk Management. *Journal of Business and Economic Statistics* 19, 465-474.

[7] Berkowitz, J., Christoffersen, P., Pelletier, D., 2008. Evaluating Value-at-Risk Models with Desk-Level Data. Mimeo. Available at: http://www4.ncsu.edu/~dpellet/papers/BCP_23Jun08.pdf

[8] Black, F., 1976. The Pricing of Commodity Contracts. *Journal of Financial Economics*, 3, 167-179.

[9] Boero, G., Marrocu, E., 2004. The Performance of SETAR Models: A Regime Conditional Evaluation of Point, Interval and Density Forecasts. *International Journal of Forecasting* 20, 305–20.

[10] Breeden, D., Litzenberger, R., 1978. Prices of state-contingent claims implicit in option prices, *Journal of Business*, 51, 621-51.

[11] Burnside, C., Eichenbaum, M., Kleshchelski, I., Rebelo, S., 2006. The Returns to Currency Speculation. NBER Working Paper 12489.

[12] Capistrán, C., Timmermann, A., 2009. Forecast Combination with Entry and Exit of Experts. *Journal of Business and Economic Statistics* 27, 428–40.

[13] Chen, Y., Tsang, K.P., 2009. What Does the Yield Curve Tell Us about Exchange Rate Predictability? Manuscript, University of Washington, and Virginia Tech.

[14] Chernozhukov, V., Fernandez-Val, I., Galichon, A., 2010. Quantile and Probability Curves without Crossing. *Econometrica* 78, 1093-1125.

[15] Cheung, Y.-W., Chinn, M.D., Pascual, A.G., 2005. Empirical exchange rate models of the nineties: are any fit to survive? *Journal of International Money and Finance* 24, 1150–1175.

[16] Christoffersen, P.F., 1998. Evaluating Interval Forecasts. *International Economic Review* 39, 841–862.

[17] Christoffersen, P.F., Hahn, J., Inoue, A., 2001. Testing and Comparing Value-at-Risk Measures. *Journal of Empirical Finance* 8, 325-342.

[18] Christoffersen, P.F., Mazzotta, S., 2005. The Accuracy of Density Forecasts from Foreign Exchange Options. *Journal of Financial Econometrics* 3, 578–605.

[19] Christoffersen, P.F., Pelletier, D., 2004. Backtesting Value-at-Risk: A Duration-Based Approach. *Journal of Financial Econometrics* 2 (1), 84-108.

[20] Clark, T.E., 2011. Real-Time Density Forecasts from Bayesian Vector Autoregressions with Stochastic Volatility. *Journal of Business and Economic Statistics* 29, 327-341.

[21] Clark, T.E., McCracken, M.W., 2012. Advances in Forecast Evaluation. Federal Reserve Bank of St. Louis, Working Paper Series 2011-025B. Available at: http://research.stlouisfed.org/wp/2011/2011-025.pdf

[22] Clements, M.P., 2004. Evaluating the Bank of England Density Forecasts of Inflation. *Economic Journal* 114, 844-866.

[23] _ _ _ _ _ _ _ _, 2005. Evaluating econometric forecasts of economic and financial variables. Palgrave Macmillan, first edition.

[24] Clews, R., Panigirtzoglou, N., Proudman, J., 2000. Recent developments in extracting information from options markets. Bank of England, Quarterly Bulletin: February 2000.

[25] Corradi, V., Swanson, N. 2006. Predictive density evaluation. In Handbook of economic forecasting, vol. 1, 197-284.

[26] Crnkovic, C., Drachman, J., 1997. Quality Control in VaR: Understanding and Applying Value-at-Risk. London: Risk Publications.

[27] Crouhy, M., Galai, D., Mark, R., 2001. Risk Management. McGraw-Hill.

[28] Della Corte, P., Sarno, L., Tsiakas, I., 2009. An Economic Evaluation of Empirical Exchange Rate Models. *Review of Financial Studies* 22, 3491–530.

[29] Diebold, F.X., Hahn, J., Tay, A.S., 1999. Multivariate Density Forecast Evaluation and Calibration in Financial Risk Management: High-Frequency Returns of Foreign Exchange. *Review of Economics and Statistics* 81, 661–73.

[30] Diebold, F.X., Mariano, R.S., 1995. Comparing Predictive Accuracy. *Journal of Business and Economic Statistics* 13, 253-263.

[31] Engel, C., Mark, N.C., West, K.D. 2012. Factor Model Forecasts of Exchange Rates. NBER Working Paper Series, Working Paper 18382.

[32] Engel, C., Wang, J., Wu, J.J., 2009. Long-Horizon Forecasts of Asset Prices when the Discount Factor is close to Unity. Globalization and Monetary Policy Institute Working Paper No. 36.

[33] Engel, C., West, K.D., 2005. Exchange rate and fundamentals. *Journal of Political Economy* 113, 485–517.

[34] Engle, R.F., Manganelli, S., 2004. CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles. *Journal of Business and Economic Statistics* 22 (4), 367-381.

[35] Fratzscher, M., Sarno, L., Zinna, G., 2012. The Scapegoat Theory of Exchange Rates: The First Tests. European Central Bank (ECB) Working Paper Series 1418.

[36] Gaglianone, W.P., Lima L.R., Linton, O., and Smith D.R., 2011. Evaluating Value-at-Risk Models via Quantile Regressions. *Journal of Business and Economic Statistics,* 29, 150-160.

[37] Gaglianone, W.P., Lima, L.R., 2012. Constructing density forecasts from quantile regressions. *Journal of Money, Credit and Banking* 44(8), 1589-1607.

[38] _ _ _ _ _ _ _ _, 2014. Constructing optimal density forecasts from point forecast combinations. *Journal of Applied Econometrics,* forthcoming.

[39] Glasserman, P., 2004. Monte Carlo Methods in Financial Engineering. Springer.

[40] Groen, J.J., Matsumoto, A., 2004. Real Exchange Rate Persistence and Systematic Monetary Policy Behavior. Bank of England Working Paper No. 231.

[41] Hall, S.G., Mitchell, J., 2007. Combining density forecasts. *International Journal of Forecasting* 23, 1-13.

[42] Harvey, D., Leybourne, S., Newbold, P. 1997. Testing the equality of prediction mean squared errors. *International Journal of Forecasting* 13(2), 281–291.

[43] He, X., 1997. Quantile curves without crossing. *The American Statistician* 51, 186-92.

[44] Hong, Y., Li, H., Zhao, F., 2007. Can the Random Walk Be Beaten in Out-of-Sample Density Forecasts: Evidence from Intraday Foreign Exchange Rates. *Journal of Econometrics* 141, 736–76.

[45] Jore, A.S., Mitchell, J., Vahey, S.P., 2010. Combining Forecast Densities from VARs with Uncertain Instabilities. *Journal of Applied Econometrics* 25(4), 621-634.

[46] Jorion, P., 2007. Value-at-risk: The new benchmark for managing financial risk. McGraw Hill, 3rd edition.

[47] Kascha C., Ravazzolo, F. 2010. Combining Inflation Density Forecasts. *Journal of Forecasting* 29, 231-250.

[48] Kilian, L., 1999. Exchange rates and monetary fundamentals: what do we learn from long-horizon regressions? *Journal of Applied Econometrics* 14, 491-510.

[49] Kilian, L., Taylor, M., 2003. Why Is It So Difficult to Beat the Random Walk Forecast of Exchange Rate? *Journal of International Economics* 60, 85-107.

[50] Ko, S., Park, S.Y. 2013. Multivariate density forecast evaluation: A modified approach. *International Journal of Forecasting* 29, 431-441.

[51] Koenker, R., 2005. Quantile Regression. Cambridge University Press, Cambridge.

[52] Koenker, R., Xiao, Z., 2002. Inference on the Quantile Regression Process. *Econometrica* 70, 1583–612.

[53] Kohlscheen, E., 2013. Long-Run Determinants of the Brazilian Real: a Closer Look at Commodities. Banco Central do Brasil. Working Paper no. 314.

[54] Kohlscheen, E., Andrade, S.C., 2013. Official Interventions through Derivatives: Affecting the Demand for Foreign Exchange. Banco Central do Brasil. Working Paper no. 317.

[55] Kupiec, P., 1995. Techniques for Verifying the Accuracy of Risk Measurement Models. *Journal of Derivatives* 3, 73–84.

[56] Marcellino, M., Stock, J., Watson, M., 2006. A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series. *Journal of Econometrics* 135, 499-526.

[57] Mark, N.C., 1995. Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability. *The American Economic Review* 85(1), 201-218.

[58] Meese, R., Rogoff, K., 1983a. Empirical exchange rate models of the seventies: Do they fit out of sample? *Journal of International Economics* 14, 3-24.

[59] _ _ _ _ _ _ _ _ _ _, 1983b. The Out-of-Sample Failure of Empirical Exchange Rate Models: Sampling Error or Misspecification. In J.A. Frenkel, (ed.) Exchange Rates and International Macroeconomics (Chicago: University of Chicago Press).

[60] Molodtsova, T., Papell, D.H. 2009 Out-of-Sample Exchange Rate Predictability with Taylor Rule Fundamentals. *Journal of International Economics* 77, 167-180.

[61] Morales-Arias, L., Moura, G.V. 2013. Adaptive forecasting of exchange rates with panel data. *International Journal of Forecasting* 29, 493-509.

[62] Newey, W.K., West, K.D., 1987. A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55(3), 703-708.

[63] Rossi, B., 2013. Exchange Rate Predictability. CEPR Discussion Papers 9575, C.E.P.R. Discussion Papers.

[64] Saliby, E., 1989. Repensando a simulação: a Amostragem Descritiva. Ed. Atlas.

[65] Sarno, L., Valente, G., 2009. Exchange Rates and Fundamentals: Footloose or Evolving Relationship? *Journal of the European Economic Association* 7, 786–830.

[66] Shimko, D., 1993. Bounds of Probability. RISK, 6, 33-37.

[67] Vincent-Humphreys, R. and Noss, J., 2012. Estimating probability distributions of future asset prices: empirical transformations from option-implied risk-neutral to real-world density functions. Working Paper 455. Bank of England.

[68] Wang, J. and Wu, J., 2012. The Taylor Rule and Forecast Intervals for Exchange Rates. *Journal of Money, Credit and Banking*, 44, 103–144.

[69] West, Kenneth D., 1996. Asymptotic Inference About Predictive Ability. *Econometrica* 64, 1067-1084.

[70] Wu, T.Y. 2008. Order Flow in the South: Anatomy of the Brazilian FX Market. Department of Economics, University of California, Santa Cruz. Mimeo.

Appendix

Table A.1 - Extended set of economic models, based on Molodtsova and Papell (2009) and Wang and Wu (2012)

Model	Label	Covariate vector $\mathbf{X}'_{m,t}$
I	Taylor rule model	$[1; \pi_t - \pi^*_t; y^{gap}_t - y^{gap*}_t; q_t]$
II	Taylor rule (PPP holds)	$[1; \pi_t - \pi^*_t; y^{gap}_t - y^{gap*}_t]$
III	Taylor rule (PPP holds and smoothing)	$[1; \pi_t - \pi^*_t; y^{gap}_t - y^{gap*}_t; i_{t-1} - i^*_{t-1}]$
IV	Taylor rule (smoothing)	$[1; \pi_t - \pi^*_t; y^{gap}_t - y^{gap*}_t; q_t; i_{t-1} - i^*_{t-1}]$
V	Absolute PPP model	$[1;q_t]$
VI	Relative PPP model	$[1; \Delta q_t]$
VII	Monetary model	$[1; s_t - ((m_t - m^*_t) - (y_t - y^*_t))]$
VIII	Monetary model (weaker version)	$[1; \Delta s_t - ((\Delta m_t - \Delta m^*_t) - (\Delta y_t - \Delta y^*_t))]$
IX	Forward premium model	$[1;i_t - i^*_t]$

Notes: $\pi_t(\pi^*_t)$ is the CPI inflation and $y^{gap}_t(y^{gap*}_t)$ is the output gap in the home (foreign) country, $i_t(i^*_t)$ is the short-term interest rate in the home (foreign) country, and $m_t(m^*_t)$ is the money supply and $y_t(y^*_t)$ is the output in the home (foreign) country.

Figure A.1 - Conditional PDFs of R$/US$ at December 2012

Recursive estimation

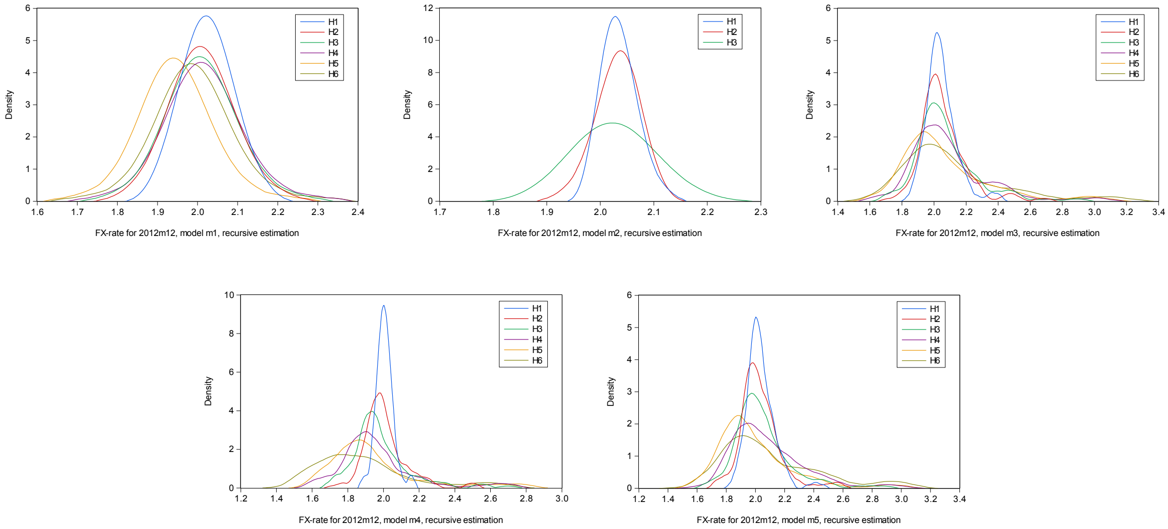


Figure A.2 - Conditional PDFs of R$/US$ at December 2012

Rolling window estimation

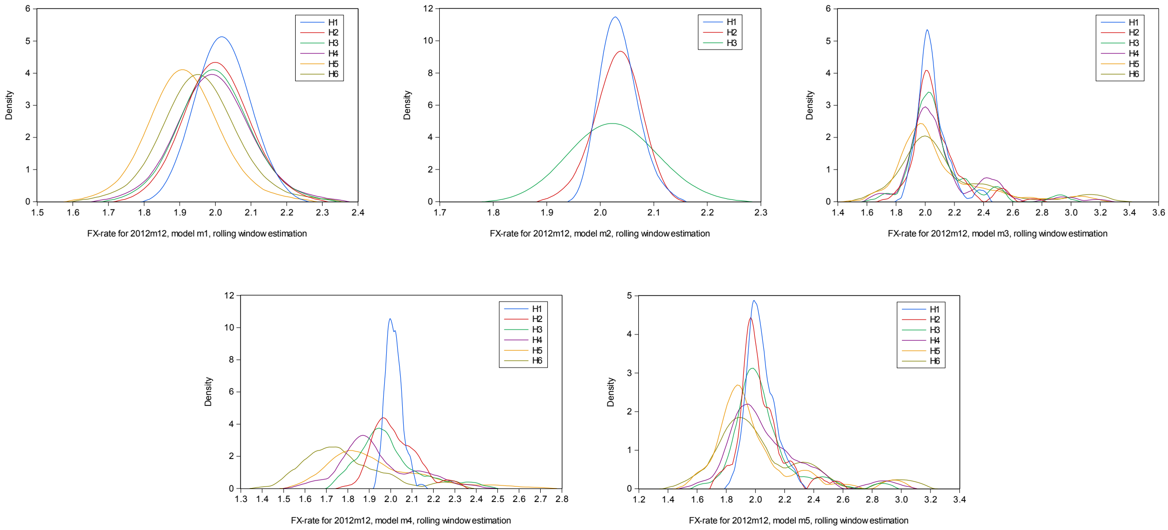


Figure A.3 - Conditional PDFs of R$/US$ at December 2009, 2010 and 2011

Recursive estimation (h=1, 3, 6, 9 and 12)

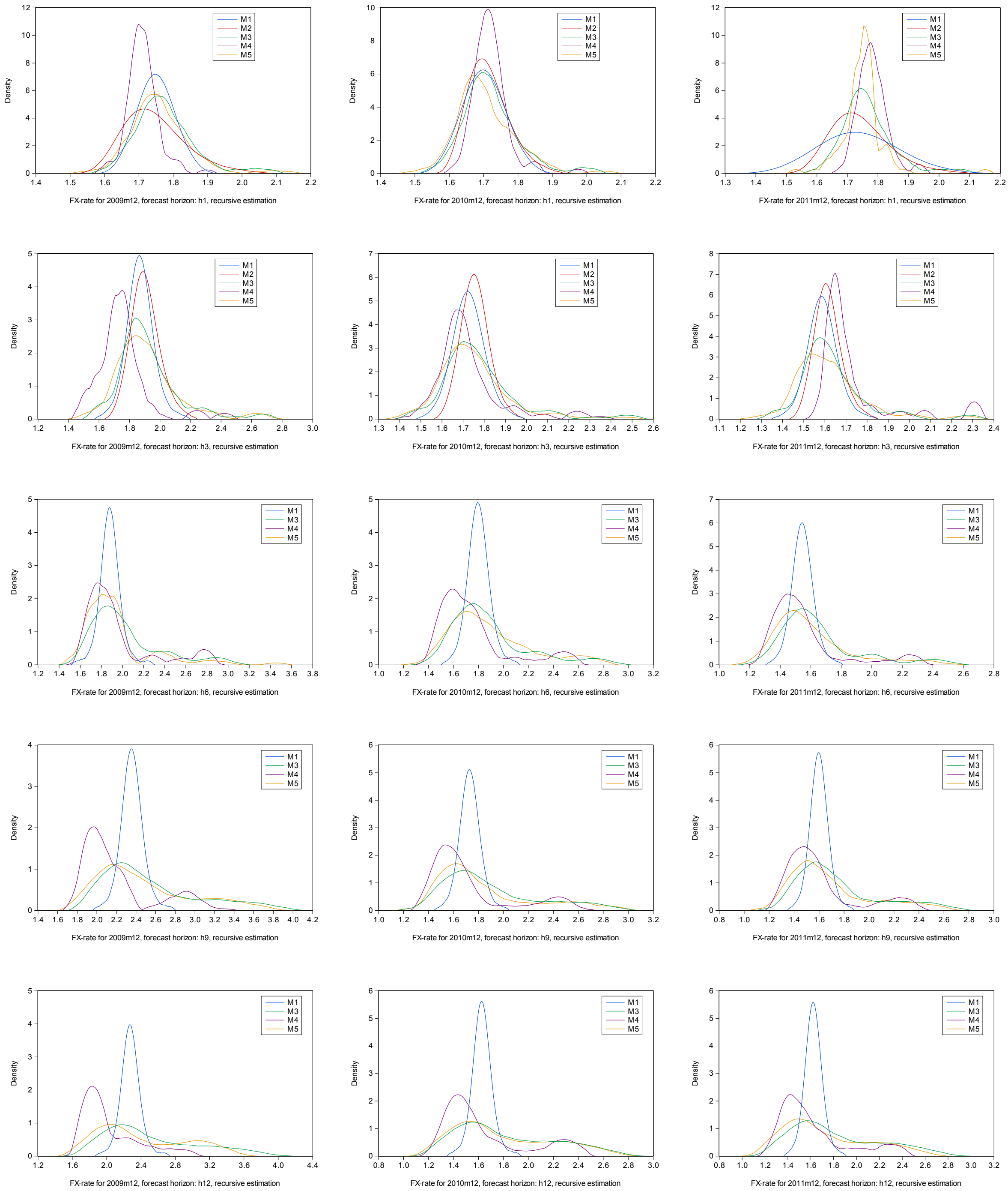


Table A.2 - Diebold-Mariano-West test modified by Harvey et al. (1997)

DMW test modified by Harvey et al. (1997): test statistic (p-value)

	Model					Model				
h	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
1	-1.37	-1.67		3.93	-0.77	-1.02	-1.67		3.92	-0.84
	(0.17)	(0.1)		(0)	(0.44)	(0.31)	(0.1)		(0)	(0.4)
2	-0.81	-1.05		2.75	-0.86	-0.89	-1.05		2.81	-1.32
	(0.42)	(0.3)		(0.01)	(0.39)	(0.38)	(0.3)		(0.01)	(0.19)
3	-0.88	-0.49		1.25	-1.27	-0.90	-0.49		1.44	-1.11
	(0.38)	(0.63)		(0.21)	(0.21)	(0.37)	(0.63)		(0.15)	(0.27)
4	-0.77			0.34	-1.40	-0.89			0.28	-1.20
	(0.44)			(0.73)	(0.16)	(0.38)			(0.78)	(0.23)
5	-0.75			0.02	-1.16	-0.79			0.05	-1.08
	(0.45)			(0.98)	(0.25)	(0.43)			(0.96)	(0.28)
6	-0.79			0.03	-1.27	-0.75			-0.30	-1.11
	(0.43)			(0.98)	(0.21)	(0.46)			(0.77)	(0.27)
7	-0.89			-0.37	-1.13	-0.79			-0.59	-1.39
	(0.38)			(0.71)	(0.26)	(0.43)			(0.56)	(0.17)
8	-0.95			-0.17	-1.08	-0.79			-0.24	-1.03
	(0.34)			(0.86)	(0.28)	(0.43)			(0.81)	(0.31)
9	-0.93			-0.46	-1.10	-0.87			-0.48	-1.59
	(0.35)			(0.65)	(0.28)	(0.39)			(0.64)	(0.12)
10	-0.91			-0.39	-1.19	-0.85			-0.39	-1.28
	(0.37)			(0.69)	(0.24)	(0.4)			(0.7)	(0.2)
11	-0.84			-0.53	-1.22	-0.88			-0.52	-1.29
	(0.4)			(0.6)	(0.23)	(0.38)			(0.61)	(0.2)
12	-0.91			-0.51	-1.27	-0.88			-0.45	-1.18
	(0.36)			(0.61)	(0.21)	(0.38)			(0.65)	(0.24)

Table A.3 - Global Coverage Rates (50%)

Panel 1 - Forecast coverage rates: % of actual outcomes inside the 50% interval band

	Model					Model				
h	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
1	0.57	0.53	0.50	0.61	0.48	0.56	0.53	0.47	0.53	0.49
	(0.15)	(0.54)	(1)	(0.02)	(0.69)	(0.31)	(0.54)	(0.54)	(0.54)	(0.84)
2	0.46	0.44	0.51	0.61	0.49	0.49	0.44	0.40	0.57	0.52
	(0.48)	(0.37)	(0.93)	(0.04)	(0.92)	(0.92)	(0.37)	(0.08)	(0.21)	(0.76)
3	0.38	0.39	0.52	0.51	0.55	0.36	0.39	0.50	0.44	0.48
	(0.04)	(0.08)	(0.71)	(0.88)	(0.38)	(0.02)	(0.08)	(1)	(0.29)	(0.76)
4	0.30		0.59	0.57	0.61	0.34		0.52	0.55	0.53
	(0)		(0.24)	(0.42)	(0.14)	(0.02)		(0.83)	(0.56)	(0.75)
5	0.20		0.47	0.54	0.62	0.29		0.34	0.53	0.49
	(0)		(0.75)	(0.68)	(0.2)	(0)		(0.07)	(0.74)	(0.91)
6	0.23		0.58	0.52	0.58	0.26		0.41	0.52	0.55
	(0)		(0.4)	(0.9)	(0.35)	(0)		(0.32)	(0.89)	(0.64)
7	0.14		0.51	0.50	0.53	0.27		0.37	0.50	0.51
	(0)		(0.92)	(1)	(0.75)	(0.01)		(0.18)	(1)	(0.92)
8	0.13		0.48	0.51	0.52	0.21		0.31	0.47	0.48
	(0)		(0.85)	(0.96)	(0.87)	(0)		(0.04)	(0.83)	(0.88)
9	0.13		0.51	0.50	0.52	0.20		0.28	0.48	0.45
	(0)		(0.92)	(1)	(0.84)	(0)		(0.03)	(0.85)	(0.68)
10	0.14		0.54	0.52	0.51	0.18		0.37	0.46	0.44
	(0)		(0.72)	(0.88)	(0.96)	(0)		(0.22)	(0.73)	(0.57)
11	0.15		0.56	0.49	0.52	0.12		0.35	0.49	0.43
	(0)		(0.63)	(0.93)	(0.85)	(0)		(0.19)	(0.93)	(0.54)
12	0.13		0.56	0.52	0.47	0.13		0.34	0.51	0.39
	(0)		(0.55)	(0.89)	(0.82)	(0)		(0.14)	(0.96)	(0.3)

Note: The table includes in parentheses p-values for the null of correct coverage (empirical = nominal rate of 50%), based on t-statistics using standard errors computed with the Newey-West estimator, with a bandwidth of 0 at the 1-month horizon and 1.5×horizon for other horizons.

Table A.4 - Global Coverage Rates (90%)

Panel 2 - Forecast coverage rates: % of actual outcomes inside the 90% interval band

	Model					Model				
h	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
1	0.90	0.88	0.90	0.97	0.91	0.90	0.88	0.85	0.86	0.88
	(0.89)	(0.46)	(0.89)	(0)	(0.83)	(0.89)	(0.46)	(0.21)	(0.32)	(0.46)
2	0.87	0.78	0.89	0.95	0.89	0.85	0.78	0.88	0.89	0.87
	(0.49)	(0.02)	(0.89)	(0.08)	(0.89)	(0.27)	(0.02)	(0.69)	(0.9)	(0.54)
3	0.77	0.67	0.93	0.91	0.93	0.79	0.67	0.87	0.85	0.88
	(0.02)	(0)	(0.56)	(0.67)	(0.51)	(0.03)	(0)	(0.6)	(0.33)	(0.75)
4	0.72		0.91	0.96	0.91	0.72		0.82	0.84	0.85
	(0.01)		(0.77)	(0.06)	(0.76)	(0.01)		(0.25)	(0.3)	(0.45)
5	0.62		0.90	0.92	0.89	0.63		0.85	0.80	0.87
	(0)		(0.97)	(0.54)	(0.88)	(0.01)		(0.55)	(0.23)	(0.7)
6	0.48		0.88	0.89	0.89	0.56		0.81	0.80	0.85
	(0)		(0.79)	(0.85)	(0.89)	(0)		(0.36)	(0.27)	(0.57)
7	0.44		0.90	0.89	0.90	0.52		0.81	0.78	0.84
	(0)		(1)	(0.85)	(1)	(0)		(0.4)	(0.21)	(0.58)
8	0.39		0.89	0.90	0.88	0.46		0.81	0.79	0.84
	(0)		(0.86)	(0.98)	(0.75)	(0)		(0.39)	(0.23)	(0.58)
9	0.36		0.91	0.85	0.95	0.44		0.76	0.77	0.85
	(0)		(0.87)	(0.51)	(0.17)	(0)		(0.25)	(0.21)	(0.63)
10	0.32		0.93	0.83	0.95	0.40		0.74	0.78	0.85
	(0)		(0.51)	(0.38)	(0.18)	(0)		(0.21)	(0.27)	(0.61)
11	0.30		0.92	0.87	0.93	0.37		0.76	0.81	0.84
	(0)		(0.74)	(0.67)	(0.53)	(0)		(0.21)	(0.37)	(0.54)
12	0.28		0.91	0.82	0.94	0.38		0.74	0.78	0.85
	(0)		(0.93)	(0.38)	(0.29)	(0)		(0.12)	(0.26)	(0.59)

Note: The table includes in parentheses p-values for the null of correct coverage (empirical = nominal rate of 90%), based on t-statistics using standard errors computed with the Newey-West estimator, with a bandwidth of 0 at the 1-month horizon and 1.5×horizon for other horizons.

Table A.5 - Local coverage rates and backtests for selected percentiles

Panel A - Forecast coverage rates: % of actual outcomes below the nominal quantile level (tau)

h=2	Model						Model				
tau	1a	2a	3a	4a	5a		1b	2b	3b	4b	5b
0.1	0.13	0.19	0.14	0.02	0.09		0.09	0.19	0.18	0.05	0.09
0.2	0.23	0.29	0.27	0.12	0.23		0.19	0.29	0.34	0.16	0.19
0.3	0.35	0.39	0.43	0.21	0.34		0.27	0.39	0.49	0.18	0.28
0.4	0.52	0.47	0.53	0.34	0.47		0.40	0.47	0.57	0.36	0.39
0.5	**0.61**	**0.56**	**0.65**	**0.49**	**0.65**		**0.54**	**0.56**	**0.65**	**0.49**	**0.49**
0.6	0.68	0.66	0.73	0.57	0.71		0.62	0.66	0.72	0.56	0.66
0.7	0.73	0.74	0.79	0.63	0.76		0.71	0.74	0.81	0.69	0.71
0.8	0.77	0.81	0.87	0.80	0.84		0.74	0.81	0.88	0.77	0.80
0.9	0.89	0.88	0.94	0.92	0.89		0.83	0.88	0.91	0.88	0.88

Panel B - Kupiec (1995) test

h=2	p-value for each model						p-value for each model				
tau	1a	2a	3a	4a	5a		1b	2b	3b	4b	5b
0.1	0.41	0.01	0.25	0.00	0.86		0.86	0.01	0.02	0.09	0.86
0.2	0.45	0.03	0.08	0.03	0.45		0.80	0.03	0.00	0.29	0.80
0.3	0.32	0.06	0.01	0.05	0.44		0.57	0.06	0.00	0.01	0.74
0.4	0.02	0.15	0.01	0.20	0.15		1.00	0.15	0.00	0.40	0.83
0.5	**0.03**	**0.26**	**0.00**	**0.92**	**0.00**		**0.47**	**0.26**	**0.00**	**0.92**	**0.92**
0.6	0.09	0.20	0.01	0.53	0.03		0.67	0.20	0.02	0.40	0.20
0.7	0.57	0.43	0.05	0.15	0.21		0.91	0.43	0.01	0.91	0.91
0.8	0.45	0.80	0.06	1.00	0.29		0.14	0.80	0.03	0.45	1.00
0.9	0.87	0.62	0.20	0.60	0.87		0.04	0.62	0.86	0.62	0.62

Panel C - Christoffersen (1998) test

h=2	p-value for each model						p-value for each model				
tau	1a	2a	3a	4a	5a		1b	2b	3b	4b	5b
0.1	0.20	0.02	0.13	0.00	0.20		0.20	0.02	0.03	0.03	0.79
0.2	0.36	0.01	0.12	0.03	0.42		0.45	0.01	0.00	0.19	0.08
0.3	0.32	0.00	0.00	0.04	0.25		0.04	0.00	0.00	0.02	0.48
0.4	0.01	0.00	0.00	0.03	0.01		0.05	0.00	0.00	0.02	0.05
0.5	**0.03**	**0.00**	**0.00**	**0.06**	**0.00**		**0.05**	**0.00**	**0.00**	**0.06**	**0.02**
0.6	0.12	0.03	0.00	0.00	0.02		0.15	0.03	0.01	0.02	0.07
0.7	0.04	0.02	0.00	0.00	0.03		0.19	0.02	0.00	0.14	0.19
0.8	0.00	0.00	0.00	0.04	0.00		0.01	0.00	0.00	0.42	0.05
0.9	0.03	0.01	0.07	0.21	0.03		0.00	0.01	0.01	0.58	0.58

Panel D - VQR (2011) test

h=2	p-value for each model						p-value for each model				
tau	1a	2a	3a	4a	5a		1b	2b	3b	4b	5b
0.1	0.06	0.01	0.05	0.00	0.01		0.01	0.01	0.07	0.28	0.00
0.2	0.12	0.14	0.11	0.00	0.04		0.02	0.14	0.04	0.16	0.05
0.3	0.11	0.38	0.22	0.20	0.17		0.13	0.38	0.02	0.58	0.17
0.4	0.05	0.59	0.15	0.16	0.09		0.37	0.59	0.00	0.30	0.25
0.5	**0.07**	**0.64**	**0.00**	**0.04**	**0.01**		**0.13**	**0.64**	**0.00**	**0.06**	**0.06**
0.6	0.02	0.43	0.01	0.12	0.00		0.03	0.43	0.00	0.35	0.00
0.7	0.00	0.00	0.00	0.08	0.00		0.00	0.00	0.00	0.14	0.00
0.8	0.00	0.00	0.00	0.02	0.00		0.00	0.00	0.00	0.10	0.00
0.9	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00

Table A.6 - Local coverage rates and backtests for selected percentiles

Panel A - Forecast coverage rates: % of actual outcomes below the nominal quantile level (tau)

h=3			Model					Model		
tau	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
0.1	0.21	0.26	0.15	0.05	0.07	0.17	0.26	0.19	0.09	0.07
0.2	0.36	0.32	0.23	0.12	0.20	0.27	0.32	0.29	0.14	0.17
0.3	0.41	0.44	0.39	0.21	0.32	0.36	0.44	0.44	0.26	0.26
0.4	0.54	0.50	0.56	0.32	0.45	0.41	0.50	0.53	0.32	0.38
0.5	**0.62**	**0.55**	**0.69**	**0.40**	**0.59**	**0.51**	**0.55**	**0.69**	**0.49**	**0.47**
0.6	0.68	0.66	0.76	0.57	0.68	0.60	0.66	0.74	0.60	0.62
0.7	0.71	0.76	0.83	0.64	0.78	0.67	0.76	0.81	0.64	0.69
0.8	0.79	0.81	0.89	0.76	0.86	0.70	0.81	0.86	0.73	0.76
0.9	0.83	0.85	0.94	0.89	0.91	0.81	0.85	0.91	0.86	0.88

Panel B - Kupiec (1995) test

h=3			p-value for each model					p-value for each model		
tau	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
0.1	0.00	0.00	0.14	0.10	0.39	0.04	0.00	0.01	0.62	0.39
0.2	0.00	0.01	0.42	0.03	0.96	0.12	0.01	0.04	0.12	0.46
0.3	0.02	0.01	0.05	0.06	0.69	0.20	0.01	0.01	0.34	0.34
0.4	0.01	0.05	0.00	0.10	0.36	0.77	0.05	0.01	0.10	0.74
0.5	**0.02**	**0.30**	**0.00**	**0.06**	**0.10**	**0.84**	**0.30**	**0.00**	**0.84**	**0.54**
0.6	0.10	0.23	0.00	0.61	0.10	0.93	0.23	0.00	0.93	0.74
0.7	0.79	0.23	0.00	0.20	0.10	0.53	0.23	0.02	0.20	0.86
0.8	0.76	0.84	0.01	0.29	0.12	0.02	0.84	0.12	0.12	0.29
0.9	0.04	0.14	0.21	0.84	0.62	0.01	0.14	0.62	0.24	0.59

Panel C - Christoffersen (1998) test

h=3			p-value for each model					p-value for each model		
tau	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
0.1	0.00	0.00	0.14	0.22	0.25	0.06	0.00	0.01	0.06	0.25
0.2	0.00	0.00	0.33	0.01	0.05	0.00	0.00	0.02	0.01	0.10
0.3	0.00	0.00	0.00	0.01	0.11	0.02	0.00	0.00	0.02	0.00
0.4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
0.5	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.01**
0.6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.18
0.7	0.00	0.01	0.00	0.00	0.00	0.02	0.01	0.00	0.00	0.00
0.8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.9	0.00	0.00	0.07	0.04	0.00	0.00	0.00	0.00	0.00	0.00

Panel D - VQR (2011) test

h=3			p-value for each model					p-value for each model		
tau	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
0.1	0.01	0.00	0.07	0.00	0.01	0.00	0.00	0.08	0.03	0.00
0.2	0.00	0.01	0.15	0.00	0.05	0.03	0.01	0.10	0.04	0.05
0.3	0.01	0.11	0.11	0.10	0.01	0.01	0.11	0.05	0.32	0.01
0.4	0.00	0.46	0.00	0.06	0.01	0.01	0.46	0.00	0.19	0.01
0.5	**0.00**	**0.73**	**0.00**	**0.02**	**0.00**	**0.00**	**0.73**	**0.00**	**0.03**	**0.00**
0.6	0.00	0.58	0.00	0.02	0.00	0.00	0.58	0.00	0.02	0.00
0.7	0.00	0.02	0.00	0.01	0.00	0.00	0.02	0.00	0.00	0.00
0.8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table A.7 - Local coverage rates and backtests for selected percentiles

Panel A - Forecast coverage rates: % of actual outcomes below the nominal quantile level (tau)

h=6	Model					Model				
tau	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
0.1	0.37		0.09	0.07	0.09	0.29		0.15	0.10	0.10
0.2	0.46		0.18	0.14	0.15	0.32		0.29	0.14	0.14
0.3	0.53		0.35	0.15	0.33	0.36		0.45	0.18	0.26
0.4	0.55		0.58	0.21	0.46	0.45		0.58	0.25	0.36
0.5	**0.64**		**0.69**	**0.45**	**0.66**	**0.49**		**0.69**	**0.47**	**0.51**
0.6	0.67		0.74	0.53	0.71	0.54		0.74	0.51	0.60
0.7	0.69		0.80	0.67	0.77	0.62		0.77	0.64	0.70
0.8	0.73		0.85	0.71	0.85	0.66		0.81	0.73	0.77
0.9	0.76		0.95	0.93	0.95	0.71		0.90	0.85	0.86

Panel B - Kupiec (1995) test

h=6	p-value for each model					p-value for each model				
tau	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
0.1	0.00		0.70	0.25	0.70	0.00		0.11	0.97	0.97
0.2	0.00		0.56	0.16	0.26	0.01		0.05	0.16	0.16
0.3	0.00		0.29	0.00	0.54	0.20		0.00	0.01	0.44
0.4	0.00		0.00	0.00	0.23	0.33		0.00	0.00	0.46
0.5	**0.01**		**0.00**	**0.35**	**0.00**	**0.92**		**0.00**	**0.60**	**0.92**
0.6	0.17		0.01	0.16	0.02	0.23		0.01	0.07	0.93
0.7	0.87		0.03	0.54	0.14	0.09		0.14	0.20	0.95
0.8	0.09		0.26	0.05	0.26	0.00		0.75	0.09	0.47
0.9	0.00		0.12	0.25	0.12	0.00		0.97	0.11	0.20

Panel C - Christoffersen (1998) test

h=6	p-value for each model					p-value for each model				
tau	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
0.1	0.00		0.00	0.25	0.00	0.00		0.00	0.01	0.00
0.2	0.00		0.02	0.00	0.00	0.00		0.00	0.00	0.00
0.3	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.4	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.5	**0.00**		**0.00**	**0.00**	**0.00**	**0.00**		**0.00**	**0.00**	**0.00**
0.6	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.7	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.8	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.9	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00

Panel D - VQR (2011) test

h=6	p-value for each model					p-value for each model				
tau	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
0.1	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.2	0.00		0.06	0.00	0.01	0.00		0.01	0.00	0.01
0.3	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.4	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.5	**0.00**		**0.00**	**0.00**	**0.00**	**0.00**		**0.00**	**0.00**	**0.00**
0.6	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.7	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.8	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.9	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00

Table A.8 - Local coverage rates and backtests for selected percentiles

Panel A - Forecast coverage rates: % of actual outcomes below the nominal quantile level (tau)

h=9	Model					Model				
tau	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
0.1	0.44		0.09	0.11	0.06	0.28		0.22	0.13	0.06
0.2	0.50		0.23	0.16	0.20	0.35		0.39	0.16	0.15
0.3	0.57		0.36	0.17	0.28	0.39		0.51	0.18	0.23
0.4	0.60		0.61	0.23	0.52	0.43		0.65	0.22	0.41
0.5	**0.63**		**0.67**	**0.38**	**0.64**	**0.52**		**0.67**	**0.38**	**0.50**
0.6	0.65		0.73	0.57	0.69	0.57		0.73	0.53	0.61
0.7	0.66		0.78	0.64	0.75	0.58		0.73	0.59	0.65
0.8	0.66		0.86	0.73	0.82	0.61		0.80	0.67	0.68
0.9	0.72		0.94	0.92	0.93	0.64		0.91	0.86	0.86

Panel B - Kupiec (1995) test

h=9	p-value for each model					p-value for each model				
tau	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
0.1	0.00		0.77	0.68	0.14	0.00		0.00	0.45	0.14
0.2	0.00		0.53	0.32	0.92	0.00		0.00	0.32	0.20
0.3	0.00		0.20	0.01	0.74	0.08		0.00	0.01	0.13
0.4	0.00		0.00	0.00	0.02	0.54		0.00	0.00	0.86
0.5	**0.02**		**0.00**	**0.02**	**0.01**	**0.67**		**0.00**	**0.02**	**1.00**
0.6	0.36		0.01	0.54	0.07	0.54		0.01	0.21	0.79
0.7	0.41		0.08	0.20	0.30	0.02		0.57	0.03	0.29
0.8	0.00		0.12	0.10	0.67	0.00		0.92	0.00	0.01
0.9	0.00		0.14	0.51	0.29	0.00		0.77	0.28	0.28

Panel C - Christoffersen (1998) test

h=9	p-value for each model					p-value for each model				
tau	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
0.1	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.2	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.3	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.4	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.5	**0.00**		**0.00**	**0.00**	**0.00**	**0.00**		**0.00**	**0.00**	**0.00**
0.6	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.7	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.8	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.9	0.00		0.02	0.02	0.00	0.00		0.00	0.00	0.00

Panel D - VQR (2011) test

h=9	p-value for each model					p-value for each model				
tau	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
0.1	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.2	0.00		0.06	0.00	0.00	0.00		0.00	0.00	0.00
0.3	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.4	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.5	**0.00**		**0.00**	**0.00**	**0.00**	**0.00**		**0.00**	**0.00**	**0.00**
0.6	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.7	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.8	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.9	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00

Table A.9 - Local coverage rates and backtests for selected percentiles

Panel A - Forecast coverage rates: % of actual outcomes below the nominal quantile level (tau)

h=12	Model					Model				
tau	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
0.1	0.45		0.11	0.13	0.13	0.35		0.25	0.12	0.11
0.2	0.52		0.25	0.14	0.25	0.41		0.38	0.14	0.16
0.3	0.56		0.34	0.18	0.36	0.42		0.41	0.18	0.36
0.4	0.56		0.54	0.26	0.51	0.44		0.55	0.26	0.44
0.5	**0.61**		**0.67**	**0.28**	**0.65**	**0.47**		**0.67**	**0.27**	**0.51**
0.6	0.64		0.72	0.41	0.68	0.52		0.68	0.38	0.58
0.7	0.66		0.80	0.65	0.75	0.54		0.72	0.55	0.62
0.8	0.67		0.92	0.74	0.86	0.55		0.75	0.69	0.67
0.9	0.68		1.00	0.89	0.99	0.60		0.89	0.81	0.81

Panel B - Kupiec (1995) test

h=12	p-value for each model					p-value for each model				
tau	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
0.1	0.00		0.86	0.38	0.38	0.00		0.00	0.60	0.86
0.2	0.00		0.29	0.16	0.29	0.00		0.00	0.16	0.40
0.3	0.00		0.41	0.01	0.20	0.02		0.03	0.01	0.20
0.4	0.00		0.01	0.01	0.05	0.51		0.00	0.01	0.51
0.5	**0.04**		**0.00**	**0.00**	**0.01**	**0.59**		**0.00**	**0.00**	**0.91**
0.6	0.50		0.02	0.00	0.12	0.12		0.12	0.00	0.66
0.7	0.41		0.04	0.29	0.28	0.00		0.72	0.00	0.13
0.8	0.01		0.00	0.19	0.16	0.00		0.29	0.02	0.01
0.9	0.00		0.00	0.86	0.00	0.00		0.86	0.01	0.01

Panel C - Christoffersen (1998) test

h=12	p-value for each model					p-value for each model				
tau	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
0.1	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.2	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.3	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.4	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.5	**0.00**		**0.00**	**0.00**	**0.00**	**0.00**		**0.00**	**0.00**	**0.00**
0.6	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.7	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.8	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.9	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00

Panel D - VQR (2011) test

h=12	p-value for each model					p-value for each model				
tau	1a	2a	3a	4a	5a	1b	2b	3b	4b	5b
0.1	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.2	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.3	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.4	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.5	**0.00**		**0.00**	**0.00**	**0.00**	**0.00**		**0.00**	**0.00**	**0.00**
0.6	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.7	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.8	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00
0.9	0.00		0.00	0.00	0.00	0.00		0.00	0.00	0.00