

2011 | 06

Working Paper

Research Department

Forecasting the intraday market price of money

Andrea Monticini and Francesco Ravazzolo

Working papers fra Norges Bank, fra 1992/1 til 2009/2 kan bestilles over e-post:

servicesenter@norges-bank.no

Fra 1999 og fremover er publikasjonene tilgjengelig på www.norges-bank.no

Working papers inneholder forskningsarbeider og utredninger som vanligvis ikke har fått sin endelige form. Hensikten er blant annet at forfatteren kan motta kommentarer fra kolleger og andre interesserte. Synspunkter og konklusjoner i arbeidene står for forfatterens regning.

Working papers from Norges Bank, from 1992/1 to 2009/2 can be ordered by e-mail:

servicesenter@norges-bank.no

Working papers from 1999 onwards are available on www.norges-bank.no

Norges Bank's working papers present research projects and reports (not usually in their final form) and are intended inter alia to enable the author to benefit from the comments of colleagues and other interested parties. Views and conclusions expressed in working papers are the responsibility of the authors alone.

ISSN 1502-8143 (online)

ISBN 978-82-7553-828-8 (online)

Forecasting the intraday market price of money*

Andrea Monticini[†] Francesco Ravazzolo[‡]
Universita Cattolica - Milano Norges Bank

June 6, 2011

Abstract

Market efficiency hypothesis suggests a zero level for the intraday interest rate. However, a liquidity crisis introduces frictions related to news, which can cause an upward jump of the intraday rate. This paper documents that these dynamics can be partially predicted during turbulent times. A long memory approach outperforms random walk and autoregressive benchmarks in terms of point and density forecasting. The gains are particular high when the full distribution is predicted and probabilistic assessments of future movements of the interest rate derived by the model can be used as a policy tool for central banks to plan supplementary market operations during turbulent times. Adding exogenous variables to proxy funding liquidity and counterparty risks does not improve forecast accuracy and the predictability seems to derive from the econometric properties of the series more than from news available to financial markets in realtime.

Keywords: interbank market, intraday interest rate, forecasting, density forecasting, policy tools.

JEL codes: C22, C53, E4, E5.

*We thank Paolo Angelini, Christophe Beuve, Fabiola Ravazzolo, Simon van Norden and seminar participants at Norges Bank, University of Padova and University of Venice for helpful comments. The views expressed in this paper are our own and do not necessarily reflect those of Norges Bank.

[†]andrea.monticini@unicatt.it

[‡]francesco.ravazzolo@norges-bank.no

1 Introduction

An explicit market for intraday loans does not exist. However, we can observe an intraday interest rate by the spread between the interest rates on two overnight loans delivered at different times within the same day (provided they are repaid at the same time next day). Furfine (2001), Baglioni and Monticini (2008), Baglioni and Monticini (2010a) and Jurgilas and Zikes (2011) find empirical evidence for the existence of such a market in the US, in the European and in the UK markets. That market is partially unexplored and rich in aspects worth to analyzing: efficiency, the microstructure, arbitrage opportunities and so on. The market efficiency hypothesis suggests a zero level for the intraday interest rate and therefore a flat intraday pattern for the rate. As discussed in Baglioni and Monticini (2010a), there are at least two reasons to support this view. The first one is related to the role of the policymakers. A positive intraday rate might induce individual banks to delay payments, imposing a negative externality on the banking system, see Angelini (1998), Bech and Garratt (2003), Mills and Nesmith (2008), Martin and McAndrews (2008), and consequently increasing the operational risk in the payment systems, see e.g. FED (2006) and FED (2007). The second one refers to the role of money as a medium of exchange. The intraday rate is just a transaction cost to settle debt which should be minimized, see Zhou (2000). Moreover, a zero level for the intraday rate provides an insurance for consumers against liquidity shocks (see e.g. Martin (2004) and Bhattacharya et al. (2007)). Central banks' daily market operations support these arguments and indeed central banks often provide free daylight credit to the banking system. For example, the Eurosystem does not charge any fee on daylight overdrafts, and cash settlements must be cleared late in the afternoon and not early.

Baglioni and Monticini (2008) show that thanks to central bank interventions the intraday markets function fairly well in normal times with interest rates close to zero. However, liquidity crises change the functioning of the markets enormously. Baglioni and Monticini (2010a) find that the ability of central banks to reduce the market price of intraday liquidity partially vanishes during crises and Baglioni and Monticini (2010b) build up a simple model to explain why. In normal times, the only friction in action is related to settlement procedures and to the cost of central bank intraday credit (see the above references and VanHoose (1991)). A liquidity crisis introduces a second component related to the chance of an upward jump of the intraday rate within the day due to some news (e.g., liquidity problems for some players in the market). Furthermore, Brunetti et al. (2010) find that central bank interventions during the recent crisis introduced uncertainty and pushed up the intraday money market rate further than (negative) economic news. Durré and Nardelli (2008)

show that intraday money market rate have been more sensitive to fine-tuning operations in recent years and Brunetti et al. (2010) claim that central banks either did not fully grasp the crowding effect, meaning commercial banks replace money market liquidity with central bank liquidity so that market conditions did not improve, see Heider et al. (2009), or consistently underestimated funding liquidity demand.

Using a database from the e-Mid market similar to Baglioni and Monticini (2010b), we document that positive intraday rates are often observed in the European market from January 2007 to April 2009, when our database stops. The average spread is around 0.05 basis points over our sample period, but it can grow up to 1 basis point. Moreover, we show that the dynamics of the series over our sample period are not random, but both in-sample and out-of-sample predictability seems to exist. In particular, our results find that a long memory approach, represented by an ARFIMA(0,d,0) model where d is the order of integration, provides superior fit-measures and statistically outperforms, in terms of point and density forecasting, random walk and autoregressive benchmarks during periods of high volatilities. This evidence seems to contradict Brunetti et al. (2010) who do not find mean reversion. Their linear specifications might not capture high persistence and nonstationarity modelled by our ARFIMA model. Moreover, our more recent sample where the intraday interest rates reduce in the final part of the sample and lower frequency data could explain the difference. Our results also indicate that intraday interest rates behave somewhat differently than longer maturity interest rates for which predictability is often not found and a random walk model is very difficult to beat, see Ang and Piazzesi (2003), Diebold and Li (2006) and de Pooter et al. (2010). Finally, adding exogenous variables which could proxy funding liquidity and counterparty risks in financial markets as the spread between the three-month Euribor and the three-month Eonia swap rates does not improve forecast accuracy, suggesting that predictability seems to derive from the econometric properties of the series more than from economic news available in real-time to the market.

We believe that our findings are very important for at least to two players in the intraday market. First, market investors could plan their liquidity management more accurately and possibly exploit opportunities related to several consecutive days of (possible predictable) positive intraday rates. For example, they could raise more funding the day before at a cheap rate and fulfill their obligations in advance. The average level of the rate seems low, but considering that the length of the investment in our exercises is just a few hours and the predictability seems to exist over several consecutive days, such strategies might generate large profits. Secondly, central banks

could plan supplementary interventions to keep intraday rates close to zero when forecasts indicate severe deviations from the zero level. Our forecast will be available every evening for the value of the rate the next day, therefore central banks will have sufficient time to plan their actions. We construct probabilistic assessments for future rates to exceed some ad-hoc target values. High probabilities of large spreads could be interpreted as an indication of market dysfunction related to, e.g., liquidity hoarding, underfunding or counterparty risk, and central banks could decide to intervene via *ad-hoc* fine-tuning operations. This measure would refer to general market conditions and it would not be linked to specific market participants. Therefore, its disclosure could help to reduce asymmetric information as market participants would know that the central bank would be willing to intervene to keep rates close to zero, discouraging high rates. But the measure does not require the disclosure of whether some individual banks have large problems, which could cause deviations from the zero level, thus avoiding stigma issues. We show that the probabilities given by our model track the realized interest rates more accurately and react faster to shocks than those given by no predictability (random walk) or mean reversion (ARMA) benchmarks.

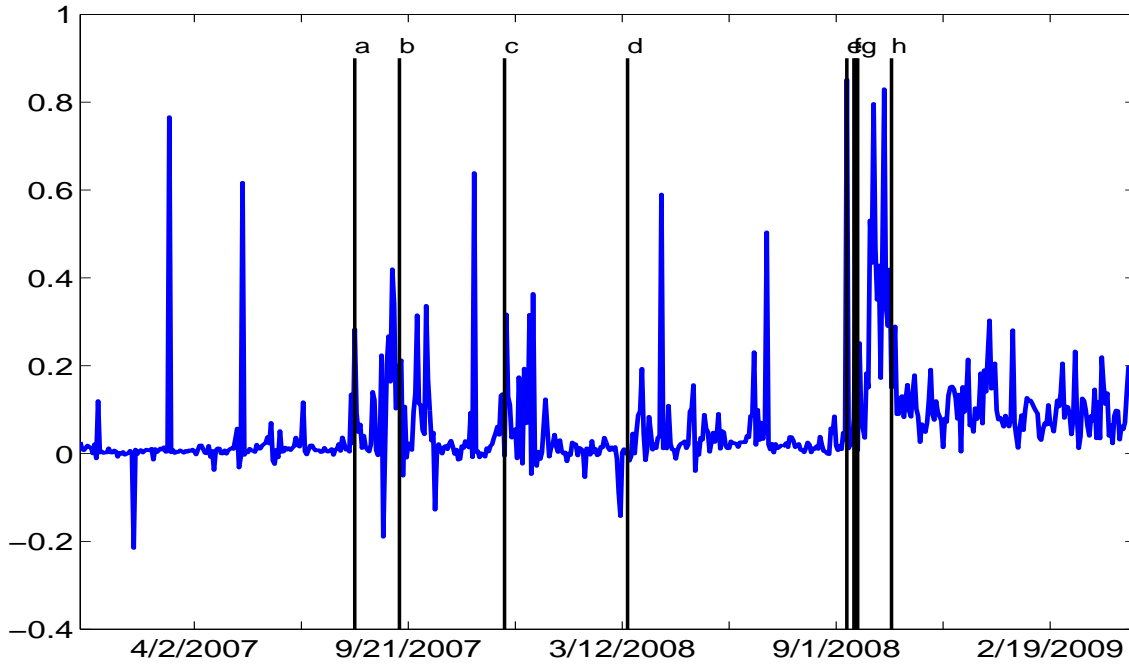
The paper is organized as follow. Section 2 describes the data set, while section 3 introduces long memory models for the intraday rate and compares them with the short memory ARMA processes, our benchmarks. Section 4 describes forecasting results and section 5 discusses how they could be used by central banks during turbulent periods. Section 6 concludes.

2 Data

Our data set includes all overnight (ON) trades taking place on the e-MID interbank market over the period January 2nd 2007 to April 30th 2009 for a total of 593 working days (see Figure 1).

The e-MID represents the only readily available source of micro data on interbank transactions in the euro area, as most other transactions of the species are conducted in the over-the-counter (OTC) market. The e-MID is located in Italy, but it is quite representative of the money market of the whole euro area (more than 200 counterparties from all over Europe have access to the system). As banks can arbitrage between the e-MID and the OTC market, the ON interest rate in the e-MID market is closely related to the Eonia rate, which is the euro ON index computed daily by the European Central Bank and published by the European Banking Federation. Furthermore, e-MID interest rates reflect actual transactions, and therefore they do not suffer from the potential distortions affecting offered rates, such as Euribor rates.

Figure 1: Data



Note: The figure shows the intraday interest rate in basis points. Timeline legend: a - 8/9/2007, BNP Paribas redemptions on three investment funds; b - 9/14/2007, liquidity support for Northern Rock; c - 12/6/2007, bank writedowns (UBS, Lehman); d - 3/17/2008, collapse of Bear Stearns; e - 9/15/2008, Lehman bankruptcy; f - 9/16/2008, loan to AIG; g - 9/17/2008, money market funds trouble; h - 10/14/2008, US TARP announced; 10/15/2008, ECB extraordinary measures.

We use our data to build up a daily time series of the intraday interest rate as in Baglioni and Monticini (2010b). For each business day in our sample, we compute the average interest rate for the ON trades taking place between 9 a.m. and 1 p.m.: this is the morning rate (R1). In the same fashion, we compute the average rate of the ON trades taking place between 2 p.m. and 6 p.m.: this is the afternoon rate (R2). The difference between R1 and R2 is the intraday interest rate. The average value of the series is 0.055 basis point, with a standard deviation of 0.110 and it is positively skewed (the skewness is equal to 2.277).¹

As theoretically explained in Baglioni and Monticini (2010b), the difference between R1 and R2 is due to both funding liquidity (hereafter we refer to funding liquidity as ‘liquidity’) and

¹We have also computed the ON series using different hours, such as the difference between the average rate of the opening value computed over the first hour of trading and the average closing value over the last hour, and differences are marginal.

counterparty risk. Forecasting the intraday rate requires the use of a proxy to take into account these risks. An obvious candidate is the spread between the three-month Euribor and the three-month Eonia swap rate (thereafter referred to Eonia-Euribor). This is a well known indicator, often used in the analysis of the liquidity crisis. It reflects both the liquidity and the counterparty risks perceived by the participants in the money market; at the same time, it is not affected by changes in interest rate expectations. Both the Euribor and the Eonia swap rate are calculated at 11 a.m., using the information provided by a panel of primary European banks². To explain why this spread is a good risk indicator, suppose this morning the spread is larger than yesterday: this may be taken as an indication that the liquidity risk and the counterparty risk perceived by market participants have gone up, presumably reflecting the release of some negative news. For this reason, we take the daily change of the spread (ω_1) as an indicator of movements of the ON rate within the day.³

3 Model

As a preliminary check, we test the intraday interbank market spread (y_t) time-series for a unit root. The unit root hypothesis is rejected (GLS-ADF (Elliott et al. (1996))) in favour of stationary ARMA models, although with autoregressive coefficients close to unity. Then we test for long memory, and the Lo's RS test (Lo (1991)) (3.062 with p-value < 0.005) provides evidence in favour of long memory. Therefore we focus on ARMA and ARFIMA models.

ARMA The first model is a traditional time series approach to modelling interest rates, the autoregressive moving average (ARMA) model (see, for example, Hamilton (1994)). The ARMA(p, q) model implies that the current value of the investigated process (say, the intraday interbank market spread y_t) is expressed linearly in terms of its past p values (autoregressive part) and in terms of the q previous values of the process ϵ_t (moving average part) and a possible set of k deterministic terms and l stochastic variables as predictors. We specify the ARMA as:

$$\phi(L)y_t = \theta_0 + \sum_{i=1}^k \psi_i x_{i,t} + \sum_{j=1}^l \varphi_j w_{j,t} + \theta(L)\epsilon_t, \quad (1)$$

²They are provided by the European Banking Federation. See <http://www.euribor.org/> for detailed information and for daily data.

³We also considered a generic Credit Default Swap Index, but predictive results are inferior to the results using the Eonia-Euribor swap rate. Results are available upon request.

where θ_0 is a constant term, $\phi(L)$ and $\theta(L)$ are the autoregressive and moving average polynomials in the lag operator L respectively, defined as:

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p, \quad (2)$$

$$\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q, \quad (3)$$

and ϵ_t is an independent and identically distributed (iid) noise process with zero mean and finite variance σ^2 ; $x_t = (x_{1,t}, x_{2,t}, \dots, x_{k,t})'$ is the $(k \times 1)$ vector of deterministic terms (dummies) at time t and $\psi = (\psi_1, \psi_2, \dots, \psi_k)'$ is a $(k \times 1)$ vector of coefficients; $w_t = (w_{1,t}, w_{2,t}, \dots, w_{l,t})'$ is the $(l \times 1)$ vector of stochastic predictors for time t , and $\varphi = (\varphi_1, \varphi_2, \dots, \varphi_l)'$ is a $(l \times 1)$ vector of coefficients.⁴

Depository institutions in the Eurosystem ought to meet some reserve requirements. The reserve requirements are the amount of funds that a depository institution ought to hold in reserve against specified deposit liabilities. The requirement has to be satisfied on average over a maintenance period. See Durré and Nardelli (2008) for a description of the current Eurosystem monetary policy operational framework. There is empirical evidence, as in Angelini (2000) for the previous Eurosystem monetary policy operational framework and Durré and Nardelli (2008) for the current one, that on "settlement" days, i.e. once a month, at the end of the reserve requirement maintenance period, the intraday volatility of the interbank rate is substantially higher than during the rest of the month. For this reason, we introduce a 0, 1 dummy variable. The dummy variable takes one on settlement days and zero otherwise. We apply two ARMA specifications: in the first model $k = 1$ ($l = 0$), which corresponds to the dummy variable. This model will be referred to the remainder of the manuscript as ARMA. In the second specification $k = 1$ and $l = 1$, referring to the use of the liquidity and counterparty risk variable defined by the Eonia-Euribor spread as stochastic predictors. The value of the spread at time t is used to predict the intraday rate at time $t + 1$. We refer to it as ARMAX.⁵

⁴An ARMA model collapses to a random walk (RW), model for absence of predictability, by assuming $\phi_1 = 1$, $\phi_2 = \dots = \phi_p = 0$ and $\theta_1 = \theta_2 = \dots = \theta_q = 0$. We find that point and density forecasting accuracies of the RW are statistically inferior to those of the ARMA models so we do not report them in section 4.2. Alternative, we use a nonparametric method based on the assumption of no predictability in section 5 as the benchmark to underline the presence of predictability. This method is similar to a RW in mean model.

⁵We have also tested ARIMA models, but results for this class of models are substantially inferior to those for the ARMA models and not reported.

3.1 ARFIMA

Long memory behaviors found in section 3 can be modelled in a more rigorous way than by using p lags in ARMA models. We propose the following ARFIMA model:

$$\phi(L)(1-L)^d y_t = \theta_0 + \sum_{i=1}^k \psi_i x_{i,t} + \sum_{j=1}^l \varphi_j w_{j,t} + \theta(L)\epsilon_t, \quad (4)$$

where we recall $\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p$, $\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q$ and ϵ_t is an independent and identically distributed (iid) noise process with zero mean and finite variance σ^2 . The parameter d specifies the order of integration. We require $0 < d < 1$, $\sum_{i=1}^p |\phi_i| < 1$, $\sum_{i=1}^q |\theta_i| < 1$. We implemented two different estimation methods for d over the full sample, a variant of the log-periodogram regression proposed by Geweke and Porter-Hudak (1983) ($\hat{d} = 0.326$ (0.044)), and the maximum likelihood (ML) estimator implemented by Sowell (1992) ($\hat{d} = 0.279$ (0.028)). Similar results were obtained in each case. We apply two ARFIMAs depending on whether the explanatory stochastic variable is used (referred as ARFIMAX) or not (referred as ARFIMA).

Bhardwaj and Swanson (2006) discuss how ARFIMA processes perform better in (point) forecasting exercises when the data sample is small. Moreover, spurious long memory behaviors arise in many contexts, such as when there are (stochastic) structural breaks in linear and non-linear models, regime switching models, and when forming models using variables that are non-linear transformations of underlying “short memory” variables, see for example Diebold and Inoue (2001), Engle and Smith (1999) and Bhardwaj and Swanson (2006).

3.2 Estimation

Figure 1 shows that the intraday interbank interest rate follows a quite erratic pattern with several possible breaks. We know ex-post what happened in our sample and we can explain breaks by specific events. However, agents did not have this information in realtime and there was high uncertainty in the market already from the second half of 2007. We estimate models over rolling windows of 65 days (roughly 3 months) to mitigate this uncertainty in an efficient but simple way.⁶

We estimate our models using maximum likelihood estimators in Ox, see Doornik and Ooms (2006). For computing 1-step ahead density forecasts, we use a conditional normal approximation as

⁶If the data sample was longer, we would advocate that a more formal analysis of breaks should be used. The shortness of the sample and the chaotic years we are investigating suggest to remove systematically “old” data even if this involves a loss of estimation efficiency.

in Huurman et al. (2010). Assuming that the past errors and coefficients are known, the conditional expectation corresponds to the point forecast of each individual model. The forecast variance is computed by approximating the forecast error variance with the in-sample estimate of the error variance σ^2 . The predictive density given by any of the models in the suite is then

$$f_{t+1,i}(y_{t+1}) \sim N(\mu_{t+1,i}, \sigma_{t+1,i}^2), \quad (5)$$

where $\mu_{t+1,i}$ is the point forecast for model i , $\sigma_{t+1,i}^2$ is the variance forecast for model i , and $i = 1, \dots, 4$ as above made at time t for $t + 1$.

4 Results

We start our study with an in-sample analysis. In-sample evidence can be interpreted as an ex-post analysis of the relevance of long memory properties in the data. Inoue and Kilian (2004) suggest that in-sample tests are likely to have greater power than out-of-sample tests. They examine the question of in-sample versus out-of-sample testing of predictability, motivated by the finding that positive in-sample evidence of predictability is often not associated with out-of-sample predictability. They argue that the claim made by Ashley et al. (1980), i.e. that in-sample inference without out-of-sample verification is likely to be spurious, with an out-of-sample approach inherently involving less overfitting, is not compelling since there is ample opportunity for the researcher to data mine in a simulated out-of-sample study, and because data snooping adjustments can be made to both tests. However, the evidence presented by Inoue and Kilian (2004) refers to point forecasting, and as the models under consideration are chosen by data-drive model selection procedures, the “impossibility” theorem in Leeb and Potscher (2005) might imply that the true distribution of the test statistics might be unknowable even by standard simulation methods. Therefore, we view the results we obtain as a natural complement to the set of mixed and conflicting results reported by leading scholars in the literature and refer to the argument of Welch and Goyal (2008) that out-of-sample tests provide “useful diagnostic” information about the underlying dynamic relationship.

First, we use in-sample analysis to choose the lag orders of our models. The best ARMA model in terms of AIC and BIC is the ARMA(1,0) (that is, $\phi_1 = 1$, $\phi_2 = \dots = \phi_p = 0$ and $\theta_1 = \theta_2 = \dots = \theta_q$), the best ARFIMA model is the ARFIMA(0,d,0). The evidence is robust to full sample data such as for the initial 65-day estimation window.

Second, Figure 2 presents in-sample evidence via Akaike Information Criterion (AIC) and

Bayesian Information Criterion on the predictability of alternative models with respect to the ARMA benchmark. Alternative models provide lower values of the criteria around unstable periods such as August 2007, Bear Stearns crisis in March 2008 or the collapse of Lehman Brothers in September 2008. ARMAX and ARFIMAX seem to provide the best fitting, even if differences between them and the ARFIMA are often small. When the market is calmer such as the period before August 2007, the benchmark seems to be marginally superior in terms of AIC and BIC.

4.1 Forecast evaluation

To shed light on the predictive power of individual models, we use a number of evaluation statistics for point and density forecasts previously proposed in literature. Our ARFIMA models, of order (0,d,0), are not nested in the ARMA benchmark, of order (1,0). We compare point forecasts in terms of mean absolute prediction error (MAPE) and root mean square prediction errors (RMSPE). We test whether mean square prediction errors are statistically different as in Diebold and Mariano (1995) (hereafter DM).⁷ Moreover, following Welch and Goyal (2008), we graphically analyze what we call the Cumulative Squared Prediction Error Difference (CSPED):

$$CSPED_{t,i} = \sum_{s=N}^t \widehat{d}_{s+1,i}, \quad t = N, \dots, T - 1, \quad (6)$$

where $\widehat{d}_{t+1,i} = e_{t+1,ARMA} - e_{t+1,i}$, $i = \text{ARMAX, ARFIMA and ARFIMAX}$ and $e_{t+1,i}$ is the MSPE for model i . Increases in $CSPED_{t,h}$ indicate that the alternative to the benchmark (ARMA model) predicts better at out-of-sample observation t .

We evaluate the predictive densities using tests of absolute forecast accuracy. Like Diebold et al. (1998), we utilize the probability integral transforms, *PITS*, of the realization of the variable with respect to the forecast densities. A forecast density is preferred if the density is correctly calibrated, regardless of the forecasters' loss function. We follow (among others) Jore et al. (2010) and use a battery of (one-shot) tests of absolute forecast accuracy, relative to the 'true' but unobserved density.

The *PITS* are:

$$z_{y_{t+1}} = \int_{-\infty}^{y_{t+1}} f_{t+1}(u) du.$$

⁷The ARMA and ARMAX are nested models, therefore the DM is not valid. We have applied the Clark and West (2007) for this comparison. Results, available upon request, are qualitatively similar.

The *PITS* should be both uniformly distributed and independently and identically distributed if the forecast densities are correctly calibrated. Hence, calibration evaluation requires the application of tests for goodness-of-fit and independence.

Our goodness-of-fit tests include the two versions of the Likelihood Ratio test proposed by Berkowitz (2001): the LR2 Berkowitz test is a two degrees of freedom variant, with a test for independence, where under the alternative $z_{y_{t+1}}$, $t = N, \dots, T - 1$, follows a white noise process; the LR3 Berkowitz test is a three degrees of freedom variant, with a test for independence, where under the alternative $z_{y_{t+1}}$ follows an AR(1) process. We also test directly for independence of the *PITS* using a Ljung-Box (LB) test, based on autocorrelation coefficients up to four. A well calibrated density should give high probability values for all three of these tests—implying the null hypothesis of no calibration failure cannot be rejected.

Turning to our analysis of relative predictive accuracy, we consider a Kullback Leibler Information Criterion (KLIC)-based test, utilizing the expected difference in the Logarithmic Scores of the candidate forecast densities; see for example Kitamura (2002), Mitchell and Hall (2005), Amisano and Giacomini (2007), Huurman et al. (2010) and Caporin and Pres (2010). Geweke and Amisano (2010) and Mitchell and Wallis (2010) discuss the value of information-based methods for evaluating forecast densities that are well calibrated on the basis of *PITS* tests. The KLIC chooses the model which on average gives higher probability to events that have actually occurred. Specifically, the KLIC distance between the true density f_{t+1} of a random variable y_{t+1} and some candidate density $f_{t+1,i}$ obtained from model i is defined as

$$\begin{aligned} \text{KLIC}_{t+1} &= \int f_{t+1}(y_{t+1}) \ln \frac{f_{t+1}(y_{t+1})}{f_{t+1,i}(y_{t+1})} dy_{t+1}, \\ &= E[\ln f_{t+1}(y_{t+1}) - \ln f_{t+1,i}(y_{t+1})]. \end{aligned} \quad (7)$$

Under some regularity conditions, a consistent estimate can be obtained from the average of the sample information, y_{N+1}, \dots, y_T , on f_{t+1} and $f_{t+1,i}$:

$$\overline{\text{KLIC}} = \frac{1}{n} \sum_{t=N}^{T-1} [\ln f_{t+1}(y_{t+1}) - \ln f_{t+1,i}(y_{t+1})]. \quad (8)$$

Even though we do not know the true density, we can still compare multiple densities, $f_{t+1,i}$. For the comparison of two competing models, it is sufficient to consider only the latter term in the

above sum,

$$LS_i = -\frac{1}{n} \sum_{t=N}^{T-1} \ln f_{t+1,i}(y_{t+1}), \quad (9)$$

for all i and to choose the model for which the expression in (9) is minimal, or as we report in our tables, the opposite of the expression in (9) is maximal. Differences in KLIC can be statistically tested. We apply a test of equal accuracy of two density forecasts similar to Mitchell and Hall (2005) and Amisano and Giacomini (2007). Suppose there are two 1-step ahead density forecasts, $f_{t+1,1}(y_{t+1})$ and $f_{t+1,2}(y_{t+1})$, and consider the loss differential

$$d_{t+1} = \ln f_{t+1,1}(y_{t+1}) - \ln f_{t+1,2}(y_{t+1}).$$

We apply the following WALD test:

$$GW = n(n^{-1} \sum_{t=N}^{T-1} h_t d_{t+1})' \widehat{\Sigma}_{t+1} (n^{-1} \sum_{t=N}^{T-1} h_t d_{t+1}), \quad (10)$$

where $h_t = (1, d_t)'$, and $\widehat{\Sigma}_{t+1}$ is the HAC estimator for the variance of $(h_t d_{t+1})$. Under the null of equal predictability, $GW \sim \chi_2^2$.

Analogous to our use of the CSPED for graphically examining relative MSPEs over time, and following Kascha and Ravazzolo (2010), we define the Cumulative Log Score Difference (*CLSD*):

$$CLSD_{t,i} = - \sum_{s=N}^t d_{s+1,i}, \quad (11)$$

where $d_{t+1} = \ln f_{t+1,i}(y_{t+1}) - \ln f_{t+1,ARMA}(y_{t+1})$. If $CLSD_{t,i}$ increases at observation t , this indicates that the alternative to the benchmark has a higher log score.

4.2 Forecast results

We construct 1-step ahead point and density forecasts for the intraday interbank rate spread over the sample period from 2 April 2007 to 30 April 2009, for a total of 529 observations.

Table 1 reports point forecast results. We think there are two clear conclusions. First, the ARFIMA model has the higher predictability: it produces both the smallest MAPE and RMSPE and the difference with respect to the benchmark ARMA model is statistically significant in terms of the DM test. Therefore, adding the fractionally integrated parameter has a substantial predictive

power, improving the accuracy of forecasts.

Second, extending ARMA and ARFIMA models with the Eonia-Euribor spread as explanatory variable deteriorates forecasting performance. Both models provide less accurate forecasts than the benchmark, reverting in-sample results. Baglioni and Monticini (2010b) discuss how the intraday rate is affected by the likelihood of a liquidity dry-up in financial markets (possible proxied by our spread). The spread at time $t + 1$ is, however, not available when forecasts are made at time t (contrary to in-sample analysis) and the variable seems to introduce forecasting errors. A better model to forecast the Eonia-Euribor spread might improve results.

The CSPED graph in Figure 3 can give further intuitions on the results. ARMAX and ARFIMAX perform particularly poorly in August 2007 at beginning of the financial crisis and after the Lehman Brothers collapse on September 15th 2008. The two models seem to lag movements in the intraday bank rate. The ARFIMA is the only model with always positive CSPED, few important increases, e.g. at the time of Lehman Brothers collapse and following days, and a (very marginally) positive trend over the full sample.

Evaluations of point forecast accuracy are only relevant for highly restricted loss functions. More generally, complete probability distributions over outcomes provide information which is helpful for making economic decisions. Therefore, we turn our analysis to density forecasting. Table 2 reports full sample results. We first focus on absolute accuracy. The null hypotheses of correct calibration for the three tests we apply is not rejected in almost all the cases for the ARMA and ARFIMA models. Only the null of the LB test for the ARMA benchmark model is rejected at 5% level, but not at 10%. *PITS* for ARMA and ARFIMA models in Figure 4 span the $[0,1]$ interval quite well, even if they are more concentrated on 0.3-0.6 percentiles. Following the discussion in Mitchell and Wallis (2010), we apply log score measures and test on them to discriminate between ARFIMA and ARMA models. Similar to point forecast evidence, ARFIMA gives the highest score and it is the only model which outperforms the benchmark in terms of the log score test. Inserting the Eonia-Euribor spread also in this case deteriorates forecast accuracy. Moreover, the CLSD of the ARFIMA model in Figure 3 is the only one that always increases over the full sample.

The fan chart in Figure 5 shows that there is a substantial increase in uncertainty after the Lehman Brothers collapse, which reduces only in 2009 when September 2008 data are discarded from our 3-month moving window estimation sample. The ARFIMA model cannot anticipate such unexpected shocks, but it copes relatively better with volatile observations just after them, giving

higher log scores than alternative models, as Figure 3 shows.⁸ The initial months of the financial crisis in August and September 2007 and days after Bear Stearns acquisitions are characterized by a different pattern: mean forecasts move largely, but variances relative less compared to September 2009.⁹

5 Probabilistic assessments for policy analysis

Economic agents in the intraday interbank market might be interested in computing probabilities that the interest rate exceeds a potential threshold target value. For example, banks could decide to postpone borrowing of money to the afternoon or not to borrow at all on those days, see Brousseau and Manzanares (2005), when the spread is forecasted to be large. But, above all, we think such assessments could be used as policy tools in crisis periods or in view of a possible exit strategy.¹⁰ Brunetti et al. (2010) finds that central bank interventions succeed in reducing uncertainty in the intraday interbank markets during normal times, but the effect is negative during crisis periods. In a study using a very detailed e-Mid database from 2 January 2006 to 1 April 2008, they show that European Central Bank (ECB) intervention in 2007 and 2008 resulted in crowding out trades that may not have otherwise occurred as commercial banks replaced money market liquidity with liquidity from the central bank, and consequently price changes, volumes, trades and volatility did not revert to mean values as the asymmetry information was not mitigated.¹¹ Brunetti et al. (2010) argue that the ECB either did not fully grasp the crowding effect or consistently underfunded the money market and suggest that *ad-hoc* operations such as interbank loan guarantees or direct asset purchases (Blanchard (2009)), should be added to standard policy

⁸The Eonia-Euribor spread increases volatility, often resulting in less precise predictive densities (fan charts for these models are available upon request). It would be interesting to investigate whether information in the intraday market can improve forecasts for the Eonia and Euribor spreads. We leave this question for future research.

⁹Looking to the full sample series we could identify two types of shocks: transitory shocks such as in August and September 2007 and September 2008 with very high volatility, and a permanent shock such as after 15 October 2008, when ECB switched from the variable rate tender format to a fixed rate full allotment policy see Abbassi and Linzert (2011), where the mean of the series shifts upward and volatility is high but less than for the previous shock. Our 65-day moving window assumption is useful for the first type of shocks. Furthermore, we find as in Bhardwaj and Swanson (2006) that ARFIMA models produce accurate forecasts when there are several stochastic and unknown structural breaks. However, a regime switching model with heteroscedasticity to cope with different types of shocks could also be applied, in particular if the sample is extended allowing for all regimes to be observed in the in-sample analysis. We leave it for further research.

¹⁰We define the exist strategy as the operation of the central bank to remove exceptional accommodative policy stance, including temporarily liquidity facilities.

¹¹Our analysis, however, shows that by using lower frequency aggregate data and extending the sample to 2009 mean reversion is found, even if there is high persistence and nonstationarity.

tools.¹² Therefore, high probabilities of large spreads given by our model could be interpreted as an indication of market dysfunction, allowing the ECB a longer period to decide whether to intervene and how to implement possible fine-tuning operations.¹³

Another property of our model-based measure would be that it refers to general market conditions and is not necessarily limited to specific market participants. This will allow the ECB to disclose it, following the suggestions in Brunetti et al. (2010) of greater transparency to reduce asymmetric information. Market participants would know part of the central bank's preference, but the disclosure of which individuals have larger problems would not be required, letting market selection discriminate between them without fear of adverse selection problems.

We compute different probabilities from the predictive densities in the previous section. We propose two definitions of shock: 1- next day $t + 1$ spread is higher than 0.1 basis points; 2- next day $t + 1$ spread is higher than 0.2 basis points. We use our best model, the ARFIMA, to compute these probabilities for our out-of-sample period from 2 April 2007 to 30 April 2009. We compare it to a nonparametric approach with forecasts of no changes in the future. This means that the 1-step ahead predictive density has a mean equal to the sample mean over the in-sample period (the last 65 days) and standard deviations equal to the sample standard deviation over the in-sample period. Figure (6) shows these probabilities.

Not surprisingly, the model-based probabilities increase substantially after August 2007, March 2008 and the Lehman Brothers collapse in September 2008, as we discussed in the previous section. In those periods, probabilities of exceeding 0.1 basis points are higher than 50%. Probabilities decrease only when big jumps associated with those events and the following turbulent days are discarded from the in-sample period. Non parametric measures are slower to react in both cases. They are also highly persistent, which results in high probabilities even in the first part of 2009 when the market was calmer than in September and October 2008 and there were only 3 observations

¹²McAndrews et al. (2008) and Fleming et al. (2010) find, however, that the Feds Term Auction Facility and Term Securities Lending Facility reduced LIBOR-OIS (London Interbank Offer Rate-Overnight Indexed Swap) spreads and repo spreads between Treasury and less liquid collateral, respectively. Federal reserve operations to reduce both funding liquidity and counterparty risks compared to the ECB focus mainly restricted to funding liquidity could explain different findings. Furthermore, Figure 1 shows that the spread has partially reduced when extending Brunetti et al. (2010) sample to 2009. Finally, Angelini et al. (2009) argue that positive high interest rates in the e-MID for maturities longer than 1 week were mainly driven by aggregate factors, notably risk aversion and accounting practices, and not by funding liquidity, capital shortage and central bank interventions.

¹³Inference on the in-sample estimator for the constant θ_0 in equation (4) being different than zero is an alternative way to test market dysfunction. However, huge shocks in the series result in high variance for the estimator, making the inference less reliable. When the level of the rate is large and positive for several consecutive days, estimation precision improves and the constant θ_0 becomes statically different than zero. The information does not seem very timely. Plot of the p-values over the OOS sample is available upon request.

higher than 0.2 basis points. We find similar evidence for the initial part of the OOS sample.

6 Conclusion

The findings presented in this paper point to the conclusion that the dynamic of the intraday interest rate during high volatility periods such as financial crises has a remarkable characteristic: a highly persistent, nonstationary process that nonetheless reverts to the mean. More precisely, we can say that the dependence is typically explained by a one-parameter time series model, the ARFIMA(0, d, 0). We provide evidence in favour of the ARFIMA model compared to random walk and ARMA benchmarks both in terms of in-sample predictability and out-of-sample point and density predictability. Statistical gains are substantially higher when focusing on the full distribution.

As an extra analysis, we propose to compute probabilistic assessments from our density forecasts for future shocks of intraday money market rate as a policy tool in times of crisis or exit strategies. Brunetti et al. (2010) argue that central bank interventions had negative results in money markets during 2007 and the initial part of 2008, supporting the more general analysis in Taylor and Williams (2009) and Sundaresan and Wang (2009) indicating that central bank facilities were not effective to counter the recent crises. Crowding effects dominate intervention news effects during crises, in contrast to calmer times. Several authors, however, do not support this view. We show that our model-based probabilities are more accurate than a nonparametric benchmark based on data over the previous three months both in tracking high jumps after dramatic shocks such as the Lehman Brothers collapse, but also in indicating the market has reverted to less volatile conditions. Therefore, our probabilistic assessments could signal the full magnitude of crowding (or other negative issues causing high rates) one day in advance giving market participants more time to plan their liquidity management and giving the central bank the opportunity to intervene as necessary in order to reduce market uncertainty, to set correct expectations and possibly to reduce the need for and size of central bank interventions, see Demiralp and Jordà (2002).

Our analysis is unfortunately agnostic on the source of the problems in the intraday money market. But we show that exogenous variables to proxy liquidity and counterparty risks deteriorates forecast accuracy. The predictability could derive from frictions in the market associated with market participant behaviors (such as unwillingness to borrow, stigma issues, central bank interventions as a source of news), which are difficult to model and above all to forecast.

References

- Abbassi, P., Linzert, T., 2011. The effectiveness of monetary policy in steering money market rates during the recent financial crisis. Working Paper Series 1328, European Central Bank.
- Amisano, G., Giacomini, R., 2007. Comparing density forecasts via weighted likelihood ratio tests. *Journal of Business & Economic Statistics* 25 (2), 177–190.
- Ang, A., Piazzesi, M., 2003. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics* 50, 745–787.
- Angelini, P., 1998. An analysis of competitive externalities in gross settlement systems. *Journal of Banking and Finance* 22, 1–18.
- Angelini, P., 2000. Are banks risk averse? intraday timing of operations in the interbank market. *Journal of Money, Credit and Banking* 32 (1), 54–73.
- Angelini, P., Nobili, A., Picillo, M. C., Oct. 2009. The interbank market after august 2007: what has changed, and why? *Temi di discussione (Economic working papers) 731*, Bank of Italy, Economic Research Department.
- Ashley, R., Granger, C. W. J., Schmalensee, R., July 1980. Advertising and aggregate consumption: An analysis of causality. *Econometrica* 48 (5), 1149–67.
- Baglioni, A., Monticini, A., October 2008. The intraday price of money: Evidence from the e-mid interbank market. *Journal of Money, Credit and Banking* 40 (7), 1533–1540.
- Baglioni, A., Monticini, A., 2010a. The intraday interest rate under a liquidity crisis: The case of august 2007. *Economics Letters* 107 (2), 198–200.
- Baglioni, A., Monticini, A., 2010b. Why does the interest rate decline over the day? evidence from the liquidity crisis. *Disce - quaderni dell'istituto di economia e finanza, Università Cattolica del Sacro Cuore, Dipartimenti e Istituti di Scienze Economiche (DISCE)*.
- Bech, M., Garratt, R., 2003. The intraday liquidity management game. *Journal of Economic Theory* 109, 198–219.
- Berkowitz, J., 2001. Testing density forecasts, with applications to risk management. *Journal of Business & Economic Statistics* 19 (4), 465–74.

- Bhardwaj, G., Swanson, N. R., 2006. An empirical investigation of the usefulness of arfima models for predicting macroeconomic and financial time series. *Journal of Econometrics* 131 (1-2), 539–578.
- Bhattacharya, J., Haslag, J., Martin, A., 2007. Why does overnight liquidity cost more than intraday liquidity? Tech. Rep. Staff Report no.281, Federal Reserve Bank of New York.
- Blanchard, O. J., Apr. 2009. The crisis: Basic mechanisms and appropriate policies. IMF Working Papers 09/80, International Monetary Fund.
- Brousseau, V., Manzanares, A., Feb. 2005. A look at intraday frictions in the euro area overnight deposit market. Working Paper Series 439, European Central Bank.
- Brunetti, C., di Filippo, M., Harris, J., 2010. Effect of central bank intervention on the interbank market during the subprime crisis. *Review of Financial Studies* forthcoming.
- Caporin, M., Pres, J., 2010. Modelling and forecasting wind speed intensity for weather risk management. *Computational Statistics and Data Analysis* forthcoming.
- Clark, T., West, K., 2007. Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics* 138(1), 291–311.
- de Pooter, M., Ravazzolo, F., van Dijk, D., Mar. 2010. Term structure forecasting using macro factors and forecast combination. Working Paper 2010/01, Norges Bank.
- Demiralp, S., Jordà, O., 2002. The announcement effect: evidence from open market desk data. *Economic Policy Review* 8 (May), 29–48.
- Diebold, F., Mariano, R., 1995. Comparing predictive accuracy. *Journal of Business and Economic Statistics* 13, 253–263.
- Diebold, F. X., Gunther, T., Tay, A. S., 1998. Evaluating density forecasts with applications to finance and management. *International Economic Review* 39, 863–883.
- Diebold, F. X., Inoue, A., November 2001. Long memory and regime switching. *Journal of Econometrics* 105 (1), 131–159.
- Diebold, F. X., Li, C., 2006. Forecasting the term structure of government bond yields. *Journal of Econometrics* 130, 337–364.

- Doornik, J., Ooms, M., 2006. A package for estimating, forecasting and simulating arfima models: Arfima package 1.04 for ox.
- Durré, A., Nardelli, S., 2008. Volatility in the euro area money market: effects from the monetary policy operational framework. *International Journal of Finance & Economics* 13 (4), 307–322.
- Elliott, G., Rothenberg, T. J., Stock, J. H., July 1996. Efficient tests for an autoregressive unit root. *Econometrica* 64 (4), 813–36.
- Engle, R. F., Smith, A. D., November 1999. Stochastic permanent breaks. *The Review of Economics and Statistics* 81 (4), 553–574.
- FED, 2006. Consultation paper on intraday liquidity management and payment system risk policy. Board of Governors of the Federal Reserve System.
- FED, 2007. Comments on the consultation paper. Federal Reserve Bank of Chicago.
- Fleming, M., Hrungr, W., Keane, F., May 2010. Repo market effects of the term securities lending facility. *American Economic Review* 100 (2), 591–96.
- Furfine, C. H., 2001. Banks as monitors of other banks: Evidence from the overnight federal funds market. *The Journal of Business* 74 (1), 33–57.
- Geweke, J., Amisano, G., April 2010. Comparing and evaluating bayesian predictive distributions of asset returns. *International Journal of Forecasting* 26 (2), 216–230.
- Geweke, J., Porter-Hudak, S., 1983. The estimation and application of long memory time series models. *Journal of Time Series Analysis* 4 (4), 221–238.
- Hamilton, J., 1994. *Time Series Analysis*. Princeton University Press.
- Heider, F., Hoerova, M., Holthausen, C., Dec. 2009. Liquidity hoarding and interbank market spreads: The role of counterparty risk. Working Paper Series 1126, European Central Bank.
- Huurman, C., Ravazzolo, F., Zhou, C., 2010. The power of weather. *Computational Statistics and Data Analysis* forthcoming.
- Inoue, A., Kilian, L., 2004. In-sample or out-of-sample tests of predictability: Which one should we use? *Econometric Reviews* 23 (4), 371–402.

- Jore, A., Mitchell, J., Vahey, S. P., 2010. Combining forecast densities from vars with uncertain instabilities. *Journal of Applied Econometrics* 25 (4), 621–634.
- Jurgilas, M., Zikes, F., 2011. Implicit intraday interest rate in the uk unsecured overnight loan market, unpublished manuscript.
- Kascha, C., Ravazzolo, F., 2010. Combining inflation density forecasts. *Journal of Forecasting* 29 (1-2), 231–250.
- Kitamura, Y., 2002. Econometric comparisons of conditional models. Discussion paper,, University of Pennsylvania.
- Leeb, H., Potscher, B. M., February 2005. Model selection and inference: Facts and fiction. *Econometric Theory* 21 (01), 21–59.
- Lo, A. W., September 1991. Long-term memory in stock market prices. *Econometrica* 59 (5), 1279–313.
- Martin, A., 2004. Optimal pricing of intraday liquidity. *Journal of Monetary Economics* 51, 401–424.
- Martin, A., McAndrews, J., 2008. Liquidity-saving mechanisms. *Journal of Monetary Economics* 55, 544–557.
- McAndrews, J., Sarkar, A., Wang, Z., 2008. The effect of the term auction facility on the london inter-bank offered rate. Staff Reports 335, Federal Reserve Bank of New York.
- Mills, D., Nesmith, T., 2008. Risk and concentration in payment and securities settlement systems. *Journal of Monetary Economics* 55, 542–553.
- Mitchell, J., Hall, S. G., 2005. Evaluating, comparing and combining density forecasts using the KLIC with an application to the Bank of England and NIESER “fan” charts of inflation. *Oxford Bulletin of Economics and Statistics* 67, 995–1033.
- Mitchell, J., Wallis, K. F., 2010. Evaluating density forecasts: Forecast combinations, model mixtures, calibration and sharpness. *Journal of Applied Econometrics* forthcoming.
- Sowell, F., 1992. Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *Journal of Econometrics* 53 (1-3), 165–188.

- Sundaresan, S., Wang, Z., March 2009. Y2k options and the liquidity premium in treasury markets. *Review of Financial Studies* 22 (3), 1021–1056.
- Taylor, J. B., Williams, J. C., January 2009. A black swan in the money market. *American Economic Journal: Macroeconomics* 1 (1), 58–83.
- VanHoose, D., 1991. Bank behavior, interest rate determination, and monetary policy in a financial system with an intraday federal funds market. *Journal of Banking and Finance* 15, 343–365.
- Welch, I., Goyal, A., 2008. A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies* 21 (4), 253–303.
- Zhou, R., 2000. Understanding intraday credit in large-value payment systems,. In: *Economic Perspectives*. Vol. 3. Federal Reserve Bank of Chicago, pp. 29–44.

Table 1: Point forecasting

	MAPE	RMSPE	DM
ARMA	0.055	0.104	
ARMAX	0.058	0.109	0.774
ARFIMA	0.053	0.100	0.006
ARFIMAX	0.056	0.106	0.607

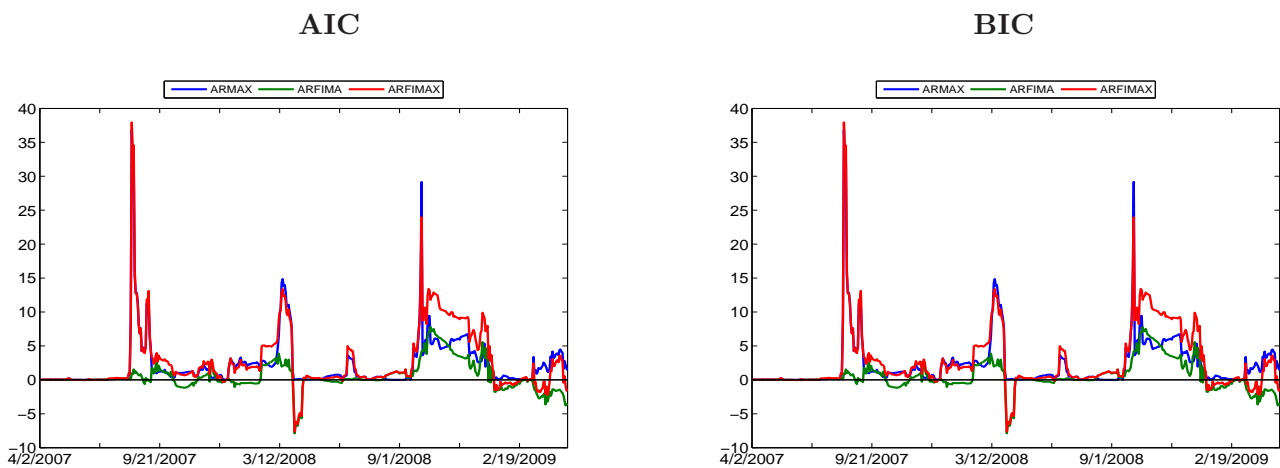
Note: Table reports results for tests of equal out-of-sample point forecast accuracy for models of European overnight interbank rate over the out-of-sample period from 2 January 2008 to 30 April 2009 for 1-step ahead forecasting horizon. The models were estimated using moving windows of 65 days. The models are: ARMA (the benchmark) and ARFIMA and their versions extended with the variable Eonia-Euribor spread, ARMAX and ARFIMAX respectively. The column MAPE reports the mean absolute prediction error. The column RMSPE gives the root mean square prediction errors. The column DM reports p -values for the Diebold and Mariano (1995) test for equal forecast accuracy. Bold numbers indicate that the null of the DM test of equal density predictive accuracy relative to the ARMA benchmark is rejected at 5% significance level. See section 3 for explanation of notation used for names of models.

Table 2: Density forecasting

	LR2	LR3	LB	LS	LS_test
ARMA	362	0.410	0.016	0.446	
ARMAX	0.002	0.004	0.058	0.334	0.000
ARFIMA	0.593	0.378	0.399	0.506	0.000
ARFIMAX	0.009	0.011	0.445	0.385	0.001

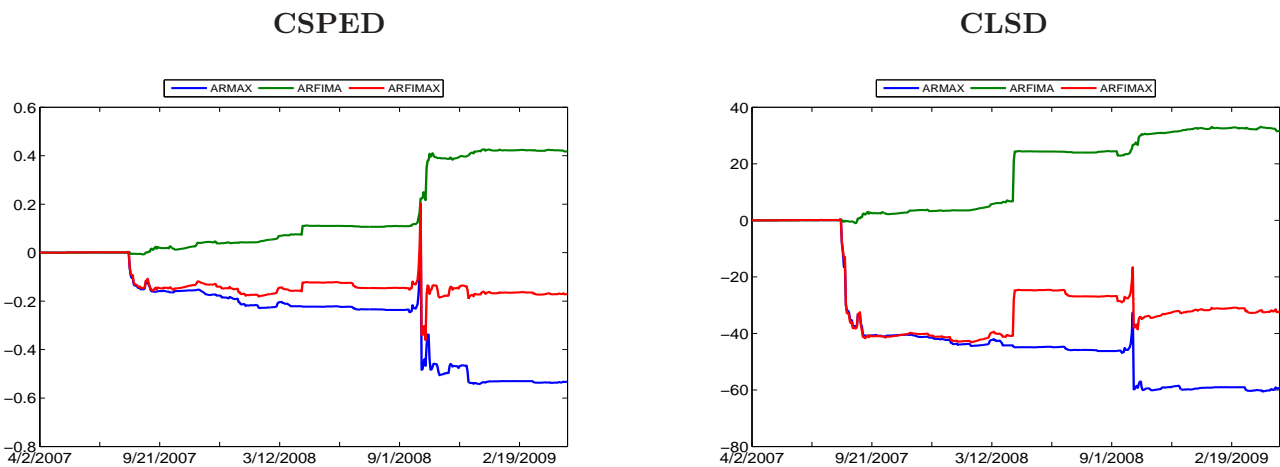
Note: The column LR2 is the Likelihood Ratio p -value of the test of zero mean and unit variance of the inverse normal cumulative distribution function transformed $PITS$, with a maintained assumption of normality for transformed $PITS$ proposed by Berkowitz (2001). The column LR3 is the version of the same test where independence is also tested. LB is the p -value from a Ljung-Box test for independence of the $PITS$. A bold number indicates that the null hypothesis of a correctly specified model cannot be rejected at 5% significance level for LR2, LR3 and LB. LS is the average Logarithmic Score over the evaluation period. The number with highest mean LS is reported in bold. The column LS_test is the p -value for the test of equal predictive density accuracy. Bold numbers indicate that the null of the test of equal density predictive accuracy relative to the ARMA benchmark is rejected at 5% significance level.

Figure 2



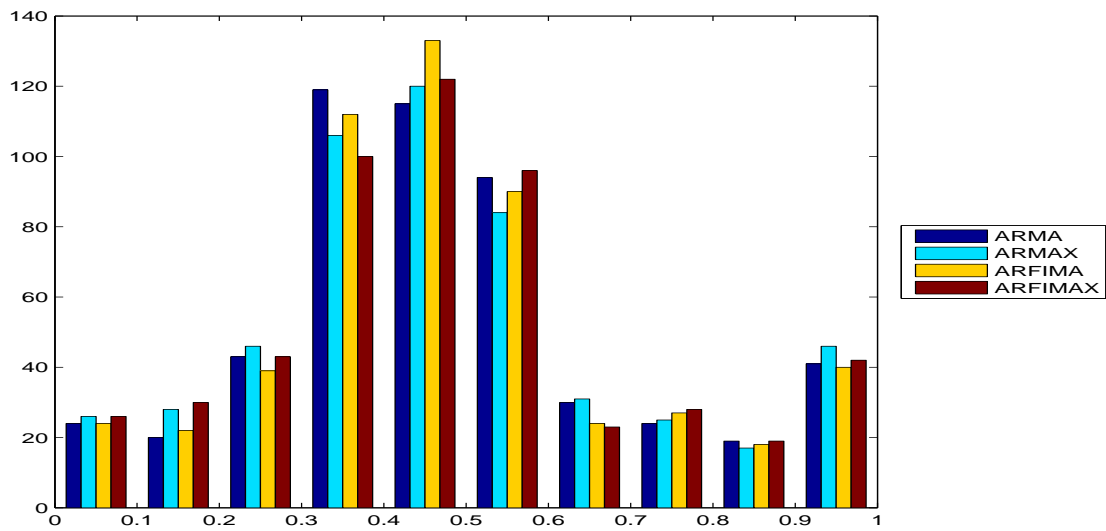
Notes: The graphs show differences in AIC ($AIC(\text{benchmark}) - AIC(\text{alternative})$) and BIC ($BIC(\text{benchmark}) - BIC(\text{alternative})$) across fixed length 65-day moving estimation windows; if the benchmark model generates the better fit, then the AIC and BIC differences are negative. The blue, green and red lines show comparisons between, respectively, the ARMA and the ARMAX models, the ARMA and the ARFIMA models and the ARMA and the ARFIMAX models.

Figure 3



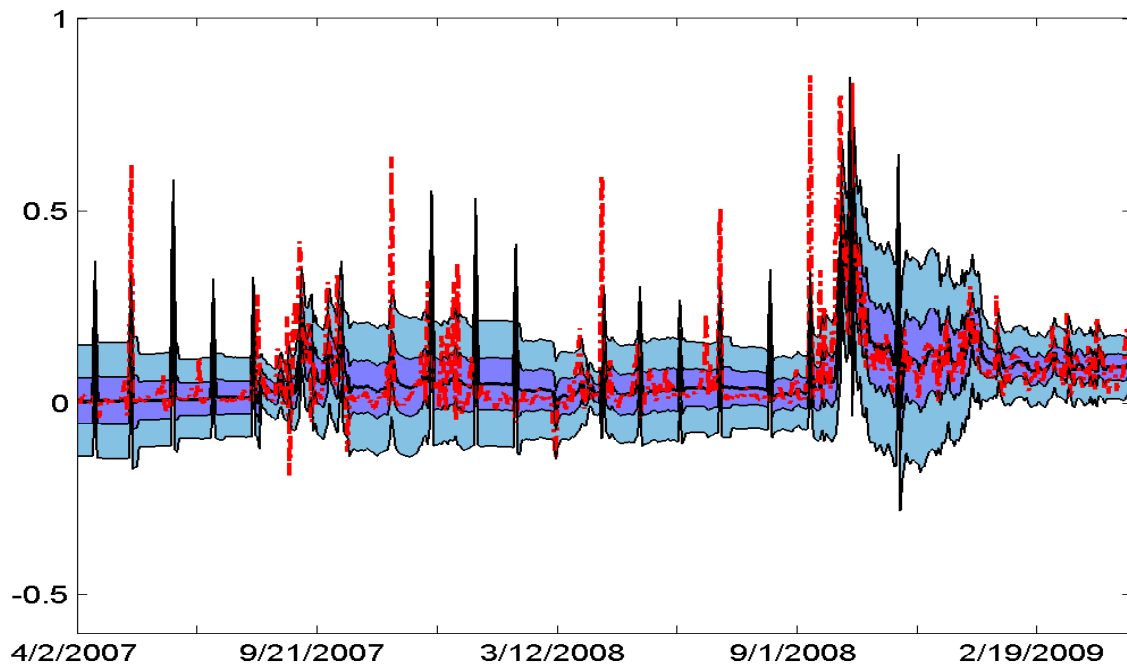
Note: The figures show cumulative square prediction error differences (CSPED) and cumulative log score differences (CLSD) of various models versus the ARMA benchmark to predict the intraday interbank rate.

Figure 4: *PITS* histograms



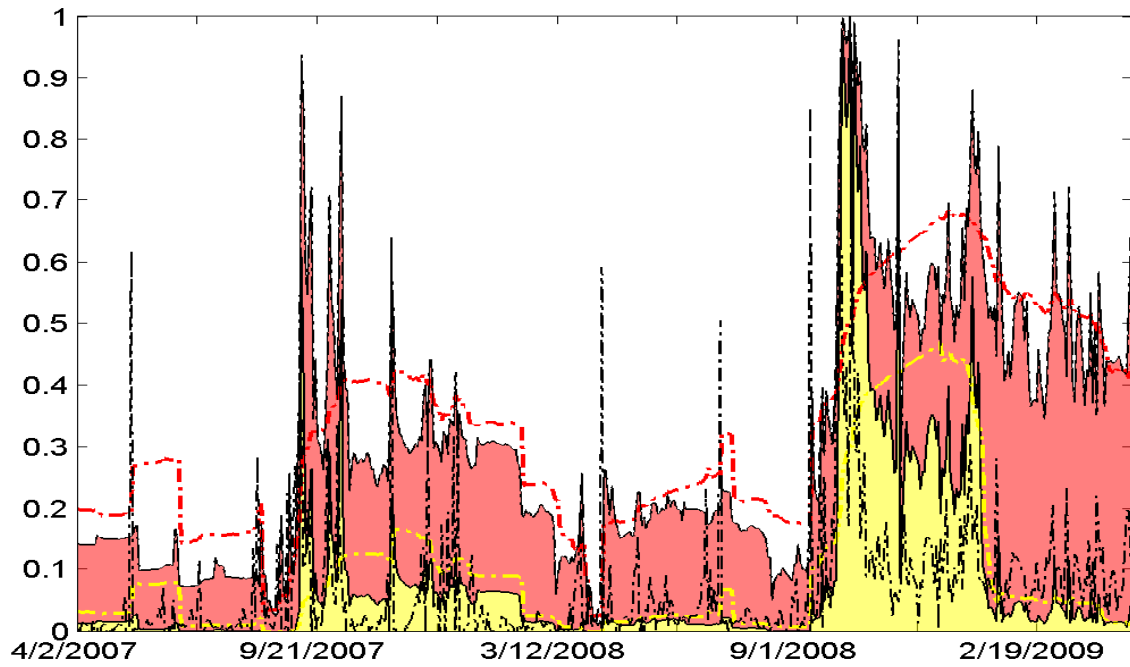
Note: The histogram shown are the decile counts of the *PITS* transforms for different models.

Figure 5: Fan charts



Note: The figure shows the fan chart given by the ARFIMA model with 5th, 25th, 50th, 75th and 95th percentiles and the European overnight interbank rate (red dashed line).

Figure 6: Probability of large spread



Note: The area plots the probability that the intraday interbank rate spread is greater than 0.1 basis points (red color), and intraday interbank rate spread is greater than 0.2 basis points (yellow color) computed using the ARFIMA model the day before. The dotted red line is the probability that the interbank rate spread is greater than 0.1 basis points, and the dotted yellow line is the probability that the interbank rate spread is greater than 0.2 basis points computed using a nonparametric method. The black dotted line shows the intraday interbank rate (only positive values).