Regime Switches in Swedish Interest Rates

Erlandsson, Ulf

2002

Regime Switches in Swedish Interest Rates

March 4, 2005

Abstract

This paper examines the forecasting properties of a Markov regime-switching model applied to Swedish interest rate volatility. A Monte Carlo testing procedure is used to arrive at a three state specification that is able to capture the high degree of leptokurtosis in the data without additional modelling of conditional heteroskedasticity. The final specification is shown to possess good forecasting properties both in general and for specific samples and horizons, something that the benchmark processes are unable to achieve.

Keywords: Regime switching, forecasting, volatility

JEL classification: C22; C52; E43

Introduction

Models of regime dependence have proven to be adept at capturing the often non-standard moments of financial and macroeconomic data. In particular, the Hamilton (1989) model has been applied in the analysis of interest rates, foreign exchange, stock indices and business cycles in numerous studies (for a survey, see Hamilton and Raj, 2002). One reason for this is to account for excess kurtosis prevalent in some series even after accounting for standard conditional heteroskedasticity effects, e.g. as is done in the GARCH family of models. Rather than define observations very far from the mean of the series as 'outliers' and remove them from the analysis, one can allow them to co-exist with more standard parts of the sample by placing them in an additional regime. For practical purposes, such as Value-at-Risk calculations, correctly depicting probabilities for rare and extreme events is of great importance.

This paper evaluates the Markov regime switching model with regard to particularly badly behaved data. The data under study is the Swedish 90 day interbank rate during the 1990s. Similar data for a set of other European countries has been studied in Dahlquist and Gray (2000). Engsted and Nyholm (2000) research regime shifts in the Danish term structure in a multivariate context. Our analysis corroborates these papers' findinds in terms of modelling the attacks on the European Exchange Rate Mechanism (ERM); multiple regimes are needed to capture the full dynamics of interest rates both in normal and volatile periods. To validate the model for practical purposes, we conduct several forecasting experiments. Diagnostic tests on the forecasts indicate that the chosen regime-switching model offers good predictive capabilities relative to the competing models.

Model and Setup

As suggested by Dewachter (1996) and Gray (1996) for similar data, we will consider the following mean-reverting data generating process:

$$\Delta r_t = \alpha_{S_t} + \beta_{S_t} r_{t-1} + \epsilon_t \tag{1}$$

where $E\left[\epsilon_t|\Omega_{t-1}\right] = 0$, $E\left[\epsilon_t^2|\Omega_{t-1}\right] = \sigma_{t,S_t}^2$, and Ω_t denotes the information set available at time t. The variable $S_t = 1, 2, ...N$ is a state variable following a first order Markov chain. The conditional variance follows $\sigma_{t,S_t}^2 = \omega_{S_t,1} + \omega_{S_t,2}\sqrt{r_{t-1}}$ in line with Cox, Ingersoll and Ross (1985). This process allows for mean-reversion, which occurs when the speed-of-adjustment variable $\beta < 0$, which in turn reverts to the long-run mean $\overline{r} = -\frac{\alpha}{\beta}$.

Using maximum likelihood methodology, the estimation procedure reduces to a optimization problem with iterative calculations of regime probabilities, as described in Hamilton (1994). The unconditional (ergodic) state probabilities are used as starting values for filtered probabilities at $t = 1$. Since the model is highly non-linear, numerous local maxima can be expected. To increase the probability of finding the global maximum of the likelihood function, we re-estimate the model 100 ($N = 2$), 200 ($N = 3$) and 500 ($N = 4$) times respectively with randomized vectors of starting values.[1]

Although laborious, the problem with specifying a Markov switching model does not lie in the estimation process. Instead, problems arise when deciding on the number of states to use. Standard testing procedures suggest just calculating a likelihood ratio statistic since the $N + 1$ state model nests the N state model. Unfortunately, this is not valid, since under the null, some of the parameters will not be identified. A number of solutions, eg. Hansen (1992) and Garcia (1998) involving grid searches have been proposed, but are difficult to implement in practice when the null hypothesis is more

[1]For level equation parameters, randomized values are drawn from a $N(\sum x/T, \sigma_x^2)$ distribution, where x is the independent variable for which we want a parameter value (for the intercept, we let $x = \Delta r$); for variance parameters using $\sqrt{\frac{\chi_1^2}{T-1}\sigma_{\Delta_r}^2}$ where χ_1^2 is a draw for a χ^2 distribution with one degree of freedom. For the transition parameters we use draws from a uniform distribution in the range [0.5 1] for values along the principal diagonal.

than one state. To reach a conclusion on the exact number of states, we instead employ the testing procedure suggested by Rydén, Teräsvirta and Åsbrink (1998), and further developed in Cheung and Erlandsson (2005). It involves generating a large number of simulated time series under the null hypothesis, and estimating the model under the null and the alternative to be able to compute a simulated likelihood ratio. The large set then represents the empirical distribution functions of the test statistic. To obtain a probability value for the empirical likelihood ratio, i.e. the one obtained when estimating the null and alternative using the real data set, one then simply calculates it by using the empirical distribution.

Empirical Specification

Our data set consists of weekly observations of the Stockholm Interbank Offered Rate (STIBOR) 90 day lending rate beginning in January, 1987, and ending in May, 2000, for a total of 690 observations. The series has been calculated using aggregated daily observations. Several observations in late 1992 show very large movements occuring on a daily basis, making inference based on just a single day's closing rate likely to be misleading for movement over the whole week. Figure 1 depicts the series in levels.

Swedish interest rates underwent several significant events during the time period. ≪Figure 1.
Most prominent was the 500% rate level in late 1992, when the financial markets anticipated a floating of the Swedish krona. In our series, this is reflected by a maximum of 35%[2]and several days of changes in the rate of more than 5% nominally. This attack on the SEK was the last of four, which also hit several other European countries during the period 1990-1992.

After 1992, rates showed a downward trend with a slight peak in the latter part of 1994, probably caused by the combination of forthcoming general elections and a runaway government deficit. The period of macroeconomic stabilization after 1994 was tranquil compared to previous years. All in all, rates went down from an average of

[2]This 'low' rate was set at the discretion of *Riksbanken* to avoid a complete credit crunch; see Dennis (1998). After averaging out, the series' maximum is 28.83%.

13% to 4% with an even more dramatic reduction of average volatility.

Some summary statistics reflect these disparities in the dynamics of the STIBOR rate. Kurtosis exceeds 100, and a high degree of leptokurtosis and non-normality continues to persist in the standardized residuals of a GARCH-model, which also has parameter estimates implying explosive volatility. Estimating the GARCH model with an underlying Student t distribution yields non-explosive estimates, but the degree of freedom parameter of 2.77 implies a non-finite fourth moment. Hence, the standard framework for working with volatility appears inadequate for this data.

Leads on the proper number of regimes can be obtained by estimation of (1) for $N = 1, 2, 3, 4$.[3] The most general model with four states encompasses the regime structure revealed in lower state models, and the increases in log likelihood values are seemingly large as N increases. Whether this holds in the statistical sense or not is evaluated with the Rydén et al. (1998) procedure. The evidence in Table 1 confirms the use of a regime-switching structure to model the data. Given the simulated empirical distributions, we are unable to reject three states in favor of four, so our final specification will consequently be a three state model. To corroborate the results on the test of one state under the null and two states under the alternative hypothesis, we have also calculated the Hansen (1992) supremum likelihood ratio statistic which yields a value of 5.97 and a probability value of < 0.01. Hence, our results of at least two states in the data does not seem to depend upon the specific testing methodology.

The final specification presented in Table 2 has been reached through a general to specific testing approach where several possible exogenous variables were introduced.[4] ≪Tables 1&2.
As predicted by the Cox et al. model, the lag of the nominal STIBOR 90 rate contains important information both for the level and variance equations. Also, the spread between the STIBOR 30 and 90 day possesses explanatory power in the level equation. This dependency on the term structure is in line with the results of Engsted and Nyholm (2000).

[3]The estimates and technical details are available in the working paper version of the paper.

[4]The following variables, in different transformations of levels and differences, have been tested: DM/Euro/SEK exchange rates, German Tagesgeld rates (equivalent to STIBOR) and STIBOR 30.

The dynamics of the regimes are in stark contrast to each other, which is further illustrated in Figure 2. A high-volatility regime, with an average standard error of 1.86,[5] occurs during the speculative attacks in the early 90s. It is short lived with a half-life of 2.22 weeks and reverts back to a long-run level of 10.41, although this estimate is imprecise and insignificant. The second regime has an average standard error of 0.22, implying a 95% confidence band for the new week's interest rate movement of ±0.4 percentage points. The half-life is 4.88 weeks and the significant long-run mean is 6.63 percent. ≪Figure 2.

Turning to the third final regime, we observe a calm state with neither much volatility nor rapid mean-reversion. The average standard error within the state is 0.063 and the long-run mean is at a relatively high 10.8 percent. This seems inconsistent considering the frequent occurrence of the regime in the latter part of the sample, where average interest rates are much lower. Since the reversion speed is so slow, this should not be considered as great a disparity after all. Movements towards the long-run mean, because of reversion, are minuscule compared to the magnitude of the random shocks entering the system.

With this regime-switching structure in place, we are able to reduce auto-correlation in the squared standardized residuals to non-significant level, as indicated in Table 2. A Jarque-Bera test of normality of the standardized residuals cannot be rejected (p-value 0.825), indicating that kurtosis has been reduced to insignificant levels.[6] In contrast, the same specification but with only two states fails to reduce excess kurtosis to insignificant levels.

[5]Average volatility and long-run means for the states have been calculated using probability weighted averages of the squared interest rate and spread series.

[6]In the spirit of Hamilton and Susmel (1994), we have also estimated the model allowing for ARCH effects with significant results. The resulting model's parameters do however imply explosive volatility within the high volatility state, and conditional normality is rejected for the residuals. Since this may be an effect of over-fitting to the most extreme observations, and in order to retain parsimony, we choose to continue without ARCH effects.

Forecasting

To validate the goodness-of-fit of the specified model, we conduct a number of forecasting experiments where its predictive power is compared with that of a number of benchmark processes. Dacco and Satchell (1999) study the forecasting performance of Markov switching models and conclude that already minor mis-specifications result in quickly deteriorating predictive performance. This is not surprising considering the high degree of parameterization and risk of overfitting inherent in the model.

The first observation to be made in this context is that the true volatility is not observed but has to be approximated. We have opted for the realized volatility measure proposed by Akgiray (1989):

$$\sigma_{o,t}^2 = \frac{\sum_{j \in I_t} \left(\hat{r}_j - \hat{r}_{j-1}\right)^2}{\sharp I_t} \tag{2}$$

where I_t denotes the set in the series of daily observations $\hat{r}$ (from Thursday to Wednesday) that correspond to the weekly observation at date t. Hence, we use higher frequency information to deduce realized volatility. One advantage of this approach is that it minimizes the probability for having weekly observations of zero volatility compared to the more standard squared difference measure $\sigma_{o,t}^2 = \Delta r_t^2$. If the Wednesday rate is the same as the previous Thursday's, the latter measure will indicate zero volatility even if there have been large movements in the days in between.[7] Figure 1 depicts the Akgiray realized volatility measure for this data set.

Calculating forecasts from a Markov switching model . Following the notation of Hamilton (1994), we collect (filtered) probabilities of states occurring at time t in the vector $\xi_{t|\Omega_t}$. We can then find the h step ahead forecast of the states simply as $\xi_{t+h|\Omega_t} = \mathbf{P}^h \xi_{t|\Omega_t}$.[8]. In this context, one can then compute a probability weighed sum

[7] To check the robustness of the results, we have also used the more standard Δr_t^2 volatility measure. Results are similar to those presented here.

[8] We note that in our case, the forecasts actually exhibit path dependence since r_t appears both in level and volatility equations. Hence, every path that the process can follow in between t and $t+h$ could be evaluated individually to produce a forecast. However, we choose the more standard procedure of assuming $r_{t+i|\Omega_t} = r_t$ for all $i \leq h$

of the individual state forecasts such that $\sigma^2_{f,t+h|\Omega_t} = \sum_{n=1}^{N} \xi^k_{t+h|\Omega_t} \cdot \sigma^2_{t+1,n}$, where $\sigma^2_{t+1,n}$ is the conditional volatility forecast from equation (1).

The first benchmark model is a random walk in volatility, denoted RW, with forecasting equations $\sigma^2_{f,t+1|\Omega_t} = \sigma^2_{o,t}$, where the subscript f denotes forecast. We will also consider an integrated GARCH model, denoted IG, with forecasts $\sigma^2_{f,t+1|\Omega_t} = \hat{h}_{t+1}$ where $\hat{h}_{t+1} = \nu + \gamma\hat{\epsilon}^2_t + \delta h_t$.[9] The residuals follow $\hat{\epsilon}_t = \Delta r_t - (\alpha + \beta_1 r_{t-1} + \beta_2 d_{t-1})$ where d_t is the interest rate differential at time t. Based on the same equations we also compute a constant variance benchmark, denoted CV, imposing the constraint $\gamma = \delta = 0$ when estimating the model.

We have considered four different samples, three different forecasting horizons (one, four and eight weeks ahead) and studied the end point forecasts. Out-of-sample forecasts have been calculated using parameters from the estimation of all information up to the start of the forecasts and fixed for the whole forecasting period. These estimates are quite stable. The different samples have been chosen to represent a number of different situations. In-sample forecasts are produced to see how the model fares overall. To study the situation where a practitioner has the model and data pre-dating the start of the spectacular volatility prior to the flotation of the krona, we let one experiment be based on data up till 1992, and then forecast for the period 1992-2000. In order to assess the performance of the model in periods where the high volatility regime does not occur, we let one set of forecasts range over 1994-2000. The last sample, 1997-2000, is considered to evaluate performance in periods of very small intra-weekly interest rate movements.

Some broad tendencies based on the results in Tables 3 and 4 can be distinguished. First of all, in none of the cases do we reject the general goodness-of-fit test of the regime-switching (MS) model's forecasted volatility regressed on realized volatility, as indicated in the bottom panel and row of Table 3. Contrary to this, both the random walk and the IGARCH processes are rejected for some subsamples and fore-

≪Table 3&4.

[9] The standard GARCH model with the explosive behavior in volatility yields poor forecasts compared to the IGARCH specification, as is the case for the non-explosive Student t specification (results available upon request).

cast horizons. Second and unsurprisingly, the constant variance model fares badly in comparison to the dynamic alternatives. There is one exception though, for the 1992 sample it does well both in the mean squared sense and in terms of mean absolute values. The inference based on the MSE measure must be taken with a grain of salt though, because it is very sensitive to the outcome of individual observations.

Relating the performance of the MS model to that of the RW model, we find strong evidence in favor of the former. In terms of MAE, the MS model is never significantly outperformed by the RW model at the 5% level (probability values are to be found in the bottom row, second column in the individual panels in Table 4.) For horizons of four and eight weeks, the MS process always produces lower errors than the RW. At the one week horizon, we find some superior performance of the RW for the samples where the speculative attacks are included, albeit the differences are insignificant at the 5% level and using the MAEs, suggesting that the higher frequency information in the RW model can be valuable in the short-run.

For the relative comparison of the IGARCH and MS models, the conclusions are the opposite in terms of in which samples the relative performance of the MS model is superior. By all measures, the IG model has very large problems capturing the volatile periods of the data, but does much better in the later, less volatile samples. For the in-sample and 1992 experiment the MS model outperforms the IG in all instances, albeit not significantly for the four and eight week forecasts in the in-sample forecasts. For the later samples, the IG outperforms the MS significantly in some instances (1994, one week; 1997, four weeks) at the 5% level, but is outperformed in others. It is natural that the MS fares worse in the periods where one or more states do not occur, since as long as the probabilities of going there are non-zero, forecasts will be worse than their single state counterparts.

In the light of these results, although the MS model forecasts compare favorably in most experiments in absolute terms, its main advantage lies in consistently being able to perform well unconditional on the forecasting horizon, the forecasted time-period and the events that occur during it. To verify this claim, we calculate the average magnitude of the loss measure in excess of the best individual alternative, accumulated over all the samples and horizons. The results in Table 5 further underlines the consistency in the MS model. It produces forecast that are on average 7.4%/15.8% worse than the best forecast for any individual sample and horizon. With corresponding figures of 28.2%/50.6%, although it is the second best option, the RW is far behind. Admittedly, the averages could be sensitive to single experiments where performance is very bad. Our calculations of median values does however not alter the results.

≪Table 5.

Conclusion

In our analysis, we have garnered strong support for the use of a three state Markov regime-switching model for analyzing and forecasting Swedish interest rate volatility. In a framework similar to Cox, Ingersoll and Ross (1985), allowing for three different states, we find no need for additional modelling of conditional heteroskedasticity and cannot reject that the model's standardized residuals are normally distributed. In contrast, alternative specification efforts using GARCH effects yield explosive estimates unless we depart from the assumption of conditional normality.

Via its inferred state probabilities, the model replicates the speculative attacks on the Swedish krona in the early 1990s letting one regime quite precisely reflect these high volatility events. The inferred long-run means range approximately between 6.6 and 10.8% but with very low speed of reversion in the latter part of the sample. Parameter estimates and state probabilities consequently appear consistent with qualitative descriptions of the data. The forecasting properties of the model are attractive compared to a number of alternative models. Although some of the benchmarks may outperform it for specific samples and horizons, the MS model is always among the best alternatives for all forecasting experiments. Its forecasts are on average 7.4%/15.8% worse, in terms of MAE/MSE, than the best forecast in a specific forecasting experiment. This indicates that for cases where it is not superior in performance, it still does not lag far behind the better alternative, which is an important aspect for practical applications.

References

Akgiray, V. (1989). Conditional Heteroskedasticity in Time Series of Stock Returns: Evidence and Forecasts. *Journal of Business*, 62:55–81.

Cheung, Y.-W. and Erlandsson, U. G. (2004). Exchange rates and Markov switching dynamics. Forthcoming in Journal of Business and Economic Statistics.

Cox, J. C., Ingersoll, Jr., J. E., and Ross, S. A. (1985). A Theory of the Term Structure of Interest Rates. *Econometrica*, 53:385–407.

Dacco, R. and Satchell, S. (1999). Why Do Regime-Switching Models Forecast So Badly. *Journal of Forecasting*, 18:1–16.

Dahlquist, M. and Gray, S. F. (2000). Regime-switching and interest rates in the European monetary system. *Journal of International Economics*, 50:399–419.

Dennis, B. (1998). *500%*. Bokförlaget DN, Stockholm.

Dewachter, H. (1996). Modelling Interest Rate Volatility: Regime-Switches and Level Links. *Weltwirtschaftliches Archiv*, 132:236–257.

Engsted, T. and Nyholm, K. (2000). Regime shifts in the Danish term structure of interest rates. *Empirical Economics*, 25:1–13.

Garcia, R. (1998). Asymptotic Null Distribution of the Likelihood Ratio Test in Markov Switching Models. *International Economic Review*, 39(3):763–788.

Gray, S. F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics*, 42:27–62.

Hamilton, J. and Raj, B. (2002). New directions in business cycle research and financial analysis. *Empirical Economics*, 27:149–162.

Hamilton, J. D. (1989). A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica*, 57:357–384.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press, Princeton.

Hamilton, J. D. and Susmel, R. (1994). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics*, 64:307–333.

Hansen, B. E. (1992). The Likelihood Ratio Test under Non-Standard Conditions: Testing the Markov Switching Model of GNP. *Journal of Applied Econometrics*, 7:S61–S82.

Harvey, D., Leybourne, S., and Newbold, P. (1998). Test for Forecast Encompassing. *Journal of Business and Economic Statistics*, 16:254–259.

Rydén, T., Teräsvirta, T., and Åsbrink, S. (1998). Stylized facts of daily return series and the hidden Markov model. *Journal of Applied Econometrics*, 13:217–244.

H_0	H_{alt}	Max.	Median	Mean	Empirical	prob.
$N = 1$	$N = 2$	1019.21	364.91	400.63	1619.98	0.02
$N = 2$	$N = 3$	1133.67	24.43	65.72	270.86	0.04
$N = 3$	$N = 4$	1092.06	29.03	66.50	67.94	0.16

Table 1: Test for Markov switching dynamics. "Max", "Median" and "Mean" refers to summary statistics of the simulated likelihood ratio statistics; "Empirical" refers to the likelihood ratio test statistic based on the likelihood values from the estimation of N and $N+1$ state models on the empirical data set. "prob." is the fraction of simulated statistics exceeding the empirical ones, interpreted as the probability value to reject H_0 in favor of H_{alt}. The number of simulations M has been set to 250.

Parameters			Transition matrix			
	Value	prob.	P_{ij}	1	2	3
α_1	-7.4100	0.3025	1	0.7320	*0.2146*	*0.0534*
$\beta_{1,1}$	0.6407	0.2347	2	0.0208	0.8677	0.1115
$\beta_{1,2}$	-0.5617	0.0532	3	0	0.0574	0.9426
ω_1	0.9381	0.0682				
α_2	0.1618	0.0037				
$\beta_{2,1}$	-0.0245	0.0000				
$\beta_{2,2}$	-0.2558	0.0000				
ω_2	0.0162	0.0000				
α_3	-0.0324	0.0000				
$\beta_{3,1}$	0.0030	0.0014				
$\beta_{3,2}$	-0.0581	0.0010				
ω_3	0.0015	0.0000				
LL	477.3525					
ARCH	4.2850	0.3688				

Table 2: Estimates of the final specifications $\Delta r_t = \alpha_{S_t} + \beta_{S_t,1} r_{t-1} + \beta_{S_t,2} d_{t-1} + \epsilon_t$, where $\epsilon_t \sim \mathrm{N}\left(0, \omega_{S_t}\sqrt{r_{t-1}}\right)$. ARCH refers to Engle's Lagrange Multiplier statistic for ARCH effects in the standardized residuals to the order of 4. In the transition matrix, italics indicate insignificance at the 10% level and significance at the 20% level. The remaining coefficient are significant at the 1% level. Probabilitiy values have been calculated using a Wald test based on a heteroskedasticity consistent variance-covariance matrix.

Sample		1987-2000			1992-2000			1994-2000			1997-2000		
Steps ahead		1	4	8	1	4	8	1	4	8	1	4	8
MAE	CV	0.2212	0.2212	0.2212	0.2566	0.2566	0.2566	0.1393	0.1393	0.1393	0.1524	0.1524	0.1524
	RW	0.1620	0.2495	0.2575	0.1623	0.2906	0.2999	0.0692	0.0795	0.0779	0.0435	0.0488	0.0499
	IG	0.2980	0.3207	0.3158	0.3676	0.4235	0.4535	0.0573	0.0617	0.0779	0.0365	0.0338	0.0394
	MS	0.1884	0.1818	0.1701	0.1974	0.2070	0.2078	0.0685	0.0604	0.0635	0.0379	0.0403	0.0425
MSE	CV	0.6555	0.6555	0.6555	0.9895	0.9895	0.9895	0.0259	0.0259	0.0259	0.0249	0.0249	0.0249
	RW	0.3528	1.0188	1.1340	0.5031	1.5406	1.7125	0.0142	0.0216	0.0225	0.0043	0.0055	0.0055
	IG	2.9840	3.6695	3.2397	1.4339	1.6915	1.7125	0.0135	0.0128	0.0141	0.0037	0.0030	0.0029
	MS	0.4581	0.5969	0.6504	0.9879	1.0230	1.0264	0.0157	0.0132	0.0132	0.0037	0.0037	0.0033
F	RW	0.0773	0.0383	0.0351	0.0998	0.0477	0.0436	0.0782	0.0537	0.0518	0.0757	0.0615	0.0609
	IG	0.0150	0.0141	0.0148	0.0500	0.0439	0.0433	0.1016	0.2787	0.1683	0.1071	0.2204	0.3239
	MS	0.3984	0.1990	0.1701	0.0657	0.1443	0.1614	0.0782	0.1591	0.6726	0.1013	0.1085	0.1503

Table 3: Forecast evaluation measures for the different samples and horizons. Notation: CV is the constant variance model, RW is a random walk in volatility, IG is an IGARCH process and MS is the Markov regime-switching process according to the final specification in Table 2. The loss measures for the h step ahead forecast (subscripted f) for a sample $t = 1...T$ are Mean Squared Error MSE $= T^{-1}\sum_{t=1}^{T}(\sigma_{o,t} - \sigma_{f,t|\Omega_{t-h}})^2$ and Mean Absolute Error MAE $= T^{-1}\sum_{t=1}^{T}|\sigma_{o,t} - \sigma_{f,t|\Omega_{t-h}}|$, where subscript o denotes observed volatility. F-test probability values for rejection of the joint hypothesis $\theta_1 = 0$, $\theta_2 = 1$ in the regression $\sigma_{o,t+h} = \theta_1 + \theta_2\sigma_{f,t+h|\Omega_t}$ are reported in the bottom panel of the table.

In-sample (1987-2000)

Steps	1				4				8			
Model	CV	RW	IG	MS	CV	RW	IG	MS	CV	RW	IG	MS
CV	.	0.9950	0.0723	0.9941	.	0.2430	0.2138	0.9998	.	0.1821	0.2320	1.0000
RW	0.0050	.	0.0082	0.0889	0.7570	.	0.2168	0.9422	0.8179	.	0.2536	0.9858
IG	0.9277	0.9918	.	0.9754	0.7862	0.7832	.	0.8631	0.7680	0.7464	.	0.8681
MS	0.0059	0.9111	0.0246	.	0.0002	0.0578	0.1369	.	0.0000	0.0142	0.1319	.

Out-of-sample (1992-2000)

Model	CV	RW	IG	MS	CV	RW	IG	MS	CV	RW	IG	MS
CV	.	0.9961	0.0018	1.0000	.	0.2958	0.0268	1.0000	.	0.2424	0.0692	0.9994
RW	0.0039	.	0.0000	0.1592	0.7042	.	0.0213	0.9199	0.7576	.	0.0579	0.9600
IG	0.9982	1.0000	.	1.0000	0.9732	0.9787	.	0.9957	0.9308	0.9421	.	0.9731
MS	0.0000	0.8408	0.0000	.	0.0000	0.0801	0.0043	.	0.0006	0.0400	0.0269	.

Out-of-sample (1994-2000)

Model	CV	RW	IG	MS	CV	RW	IG	MS	CV	RW	IG	MS
CV	.	1.0000	1.0000	1.0000	.	1.0000	1.0000	1.0000	.	1.0000	1.0000	1.0000
RW	0.0000	.	0.9980	0.5633	0.0000	.	0.9934	0.9982	0.0000	.	0.5010	0.9618
IG	0.0000	0.0020	.	0.0000	0.0000	0.0066	.	0.7025	0.0000	0.4990	.	0.9999
MS	0.0000	0.4367	1.0000	.	0.0000	0.0018	0.2975	.	0.0000	0.0382	0.0001	.

Out-of-sample (1997-2000)

Model	CV	RW	IG	MS	CV	RW	IG	MS	CV	RW	IG	MS
CV	.	1.0000	1.0000	1.0000	.	1.0000	1.0000	1.0000	.	1.0000	1.0000	1.0000
RW	0.0000	.	0.9824	0.9532	0.0000	.	1.0000	0.9954	0.0000	.	0.9861	0.8985
IG	0.0000	0.0176	.	0.0880	0.0000	0.0000	.	0.0012	0.0000	0.0139	.	0.0331
MS	0.0000	0.0468	0.9120	.	0.0000	0.0046	0.9988	.	0.0000	0.1015	0.9669	.

Table 4: Significance of forecast improvements. Probability values for the modified Diebold-Mariano statistic as in Harvey, Leybourne and Newbold (1998) testing the null hypothesis $MAE_r = MAE_c$ versus the alternative $MAE_r < MAE_c$ where r and c denote rows and columns respectively. Note that a p-value in excess of 0.5 means that $MAE_r > MAE_c$.

		$Mean(W)$	$Median$(W)
MAE	CV	128.5	88.7
	RW	28.2	29.1
	IG	51.7	49.5
	MS	7.4	1.9
MSE	CV	212.1	94.0
	RW	50.6	69.6
	IG	166.2	38.9
	MS	15.8	3.6

Table 5: Average and median decrease in forecasting performance versus the best alternative, in percentages, calculated over all samples and horizons. Denoting the loss-measure as LM, it has been calculated as $W = \sum_{c=1}^{C} \frac{LM_c - LM_{best}}{LM_{best}}$, where c refers to columns in Table 3, and LM_{best} denotes the lowest loss measure in the corresponding column c.

Figure 1: Stockholm Interbank Offered Rate (percent, left axis) and the Akgiray volatility measure (standard deviations of Δr_t, right axis).

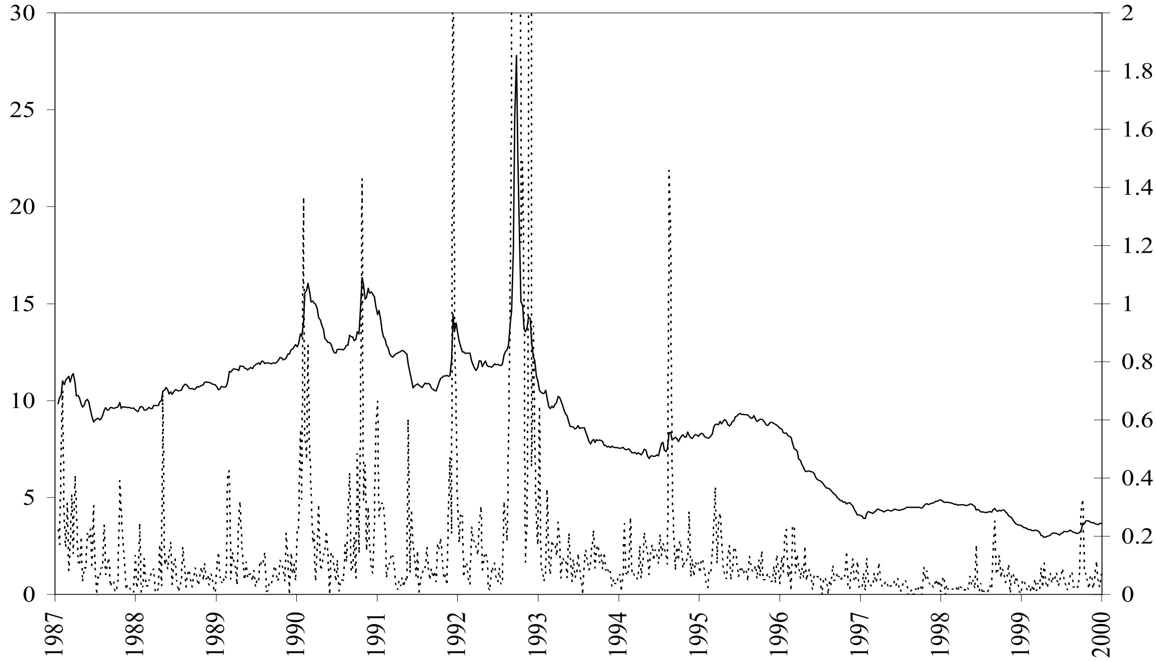

Figure 2: Smoothed state probabilities ordered from state 1 (top row, highest average volatility) to 3 (bottom row, lowest average volatility).