SYRTO

SYstemic Risk TOmography
Signals, Measurements, Transmission Channels, and Policy Interventions

Volatility prediction based on scheduled macroeconomic announcements

Athanassios Petralias, Petros Dellaportas

SYRTO WORKING PAPER SERIES
Working paper n. 13 | 2014

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement n° 320270.

Volatility prediction based on scheduled macroeconomic announcements

Athanassios Petralias[a], Petros Dellaportas[a,*]

[a] *Department of Statistics, Athens University of Economics and Business Patission 76, Athens 10434, Greece*

Abstract

We investigate the impact of scheduled macroeconomic announcements on the volatility of exchange rates by introducing a flexible model formulation. For each macroeconomic index we estimate cutoff points in the surprise component of the announcement that specify the degree the volatility process is affected. This degree is quantified by jumps of unknown size that occur before and at the time of the announcement and then die out exponentially with unknown rate. We make inferences using a population Markov chain Monte Carlo reversible jump algorithm and illustrate our methodology by predicting exchange rates volatility using fifteen U.S. macroeconomic announcements.

Keywords: Bayesian methods, Exchange rates, GARCH models, Model selection, Nonlinear time series, Nonparametric methods, Population Markov chain Monte Carlo, Reversible jump, Thresholds, Volatility forecasting

MSC 2010: 62C10, 62J10, 62P20, 62F15, 62G05

---

*Correspondence to: Prof. Petros Dellaportas, Department of Statistics, Athens University of Economics and Business, Patission 76, Athens 10434, Greece. Tel: +30 2108203567, Fax: +30 2108203565.

Email addresses: petralia@aueb.gr (Athanassios Petralias), petros@aueb.gr (Petros Dellaportas)

1. Introduction

Scheduled announcements of macroeconomic indices have been found to affect decisively the volatility of stocks and exchange rates in a wide range of studies. In this study we propose a rich model formulation that is capable of exploring how announcement surprises, defined as the absolute percentage differences of the realised minus the Bloomberg consensus reported values, affect the volatility process of exchange rates.

We propose a model formulation based on two ingredients. The first is a GARCH-type process that captures the usual stylised facts such as heavy tails and volatility clustering. The second is a nonparametric, threshold-based process that captures the impact of announcement surprises. Due to the predictive nature of the GARCH process and the fact that the announcements are scheduled, our model specification may serve as a volatility forecasting tool.

The non-linear structure possesses the following characteristics. First, it assumes that for each macroeconomic index there is a different pre-announcement effect, quantified by a jump of unknown size in the volatility process. Second, for each index the degree of announcement surprise has a different impact on the volatility process. To accommodate this, we define regions in the announcement surprise data that are specified by threshold points. Depending on both the index and the region of the announcement surprise, a volatility jump of unknown size that decays exponentially with unknown rate is added to the volatility process. The number of regions and their corresponding threshold points are unknown and different for each index.

To make the most of our flexible model framework, we need a powerful inference tool to explore a large number of models. We adopt Bayesian inference and construct a population reversible jump Markov chain Monte Carlo (MCMC) algorithm in order to efficiently sample from both the model and the parameter space. The resulting samples of posterior summaries of interest can serve as forecasting tools through model averaging.

The methodology proposed here is illustrated in an empirical study in which we use data of three exchange rates of the U.S. dollar and fifteen U.S. macroeconomic announcements. The results indicate that our proposed model formulation provides more accurate forecasts than typical conditional volatility models, and this is so even if we enrich the existing models appropriately so that macroeconomic announcements are incorporated. Furthermore, we identify the macroeconomic announcements that mostly affect exchange rates volatility, and provide insights on the way each announcement affects volatility.

The rest of the paper proceeds as follows. In Section 2 there is a short review of studies investigating the effect of news announcements on volatility. In Section 3 we present a new class of models for volatility based on macroeconomic announcements. Section 4 describes the details of the MCMC implementation whereas Section 5 presents the data and the results. Finally, Section 6 concludes with a short discussion.

2. Announcements and volatility

The impact of macroeconomic announcements on the asset and foreign exchange (Forex) markets volatility has been the subject of extensive research in past years. A large number of related studies is based on GARCH type models (Engle, 1982; Bollerslev, 1986), where additional explanatory variables are used to capture the effect of news announcements. A common formulation for these models is given as

$$y_t \mid \mathcal{F}_{t-1} = \mu_t + \epsilon_t, \quad \epsilon_t \sim N(0, G_t H_t) \tag{1}$$

where y_t, $t = 1, \ldots, T$, denotes the exchange rate return, $\mathcal{F}_t$ denotes the information set up to time t, μ_t is the expected value of the mean process at time t, H_t is a function of some explanatory variables related to news announcements affecting the volatility process and G_t is a GARCH-type process.

The main difference among these studies stands in the selection of explanatory variables affecting H_t. In fact, it is not unusual to differ not only on the type of announcements but also in the temporal effect on volatility, for example adopting formulations specifying before and after announcement effects. In this context, a large number of studies use indicator variables to account for macroeconomic announcements (Jones et al. (1998); Bomfim (2000); Kim et al. (2004); Bauwens et al. (2005); Nikkinen et al. (2006)). In DeGennaro & Shrieves (1997) news announcements are captured by the number of headlines in each news category on the Reuters money news-alerts, while Hautsch & Hess (2002) and Brenner et al. (2009) use explanatory variables that capture surprises on headline figures.

Significant empirical results are also drawn from studies analysing the effect of news announcements directly on the absolute returns (Ederington & Lee (1993); Mitchell & Mulherin (1994)); on the implied volatility (Ederington & Lee (1996); Nikkinen &

Sahlström (2001)); or on the order flows stemming from daily and intraday trades (Evans & Lyons (2005, 2008)). Flexible Fourier forms have been also adopted by incorporating ARCH, calendar and macroeconomic announcement effects (Andersen & Bollerslev (1998b); Bollerslev et al. (2000); Laakkonen (2004)). Andersen et al. (2007a) by measuring separately the continuous sample path variation and the discontinuous jump part of the quadratic variation process, observe that significant jumps in the realized volatility process tend to coincide with the release of macroeconomic indices. In statistics literature, significant jump patterns in daily volatility similar to those advocated in this paper have been uncovered in, for example, Barndorff-Nielsen & Shephard (2001) and Roberts et al. (2004).

Although the empirical results from the above studies are in some cases contradictory, there are several points of agreement. First, all relevant studies agree that news announcements are affecting decisively the volatility of exchange rates and stock markets, when analysed on either a daily or an intraday basis. Furthermore, all studies underline that the effect on the volatility is related to the type of announcement and the content of news, rather than the very act of releasing information. From all the types of news considered, most studies find that scheduled macroeconomic announcements are the most important, and their effect turns out to be stronger when compared to ARCH or calendar effects (DeGennaro & Shrieves (1997); Andersen & Bollerslev (1998b); Bollerslev et al. (2000); Bauwens et al. (2005)). The documented volatility autocorrelation and the day-of-the-week volatility patterns seem to depend strongly on the news generating process and the timing of major macroeconomic announcements (see Ederington & Lee (1993) for intraday evidence and Ederington & Lee (1996); Jones et al. (1998) for daily data). In addition, there is evidence that while the surprising factor plays a key role, even news that come out as expected seem to affect the volatility process (Laakkonen (2004)).

Studies seem to diverge with respect to the direction with which pre-announcement effects affect volatility. Nikkinen & Sahlström (2001); Hautsch & Hess (2002); Bauwens et al. (2005) found that volatility increases before news announcements. Contradictory results providing evidence that the volatility decreases before news announcements appeared in DeGennaro & Shrieves (1997); Bomfim (2000); Brenner et al. (2009).

There are also different results concerning the duration of the effect of macroeconomic announcements. Most of the studies using intraday data find that the effect of macroeconomic announcements on the intraday volatility is short lived, lasting from several minutes to one or two hours. On the other hand Evans & Lyons (2005) find that currency markets are still absorbing news for several days after an announcement while Brenner et al. (2009) estimate that volatility shifts do not offset each other over a three-day window around the announcement. In a related context, Evans & Lyons (2008) find that arrival of macro news accounts for the 36% of the total daily variance, while they point that macro news contribute far more to price variation in the long term than previously thought.

3. The Proposed Models

3.1. A Flexible Threshold-GARCH model

We propose the following variation of the general model (1). Assume there are K macroeconomic indices affecting the volatility process, and at time periods denoted by t_i^* there are announcements of index i, $i = 1, \ldots, K$. Then, we propose the model

$$
\begin{aligned}
y_t \mid \mathcal{F}_{t-1} &= \mu + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma^2 G_t H_t) && (2)\\
G_t &= (1 - \alpha_1 - \alpha_2) + \alpha_1 \frac{\epsilon_{t-1}^2}{\sigma^2 H_{t-1}} + \alpha_2 G_{t-1} && (3)\\
H_t &= 1 + \sum_{i=1}^{K} \left(\sum_{j=1}^{J_i} I_{ij}(t_i^*) \gamma_{ij} \exp\left(-r_{ij}(t - t_i^*)\right) + s_i \mathbf{1}\left(t = t_i^* - 1\right) \right) && (4)\\
I_{ij}(t_i^*) &= \mathbf{1}\left(c_{ij} \leq Z_{it_i^*} < c_{i,j+1}\right). && (5)
\end{aligned}
$$

The positive scalar σ^2 is considered as the global static error variance whereas the GARCH process G_t has the form used in the Spline-GARCH model of Engle & Rangel (2008) with $E(G_t) = 1$ and positive parameters $0 < \alpha_1 + \alpha_2 < 1$. The threshold process $H_t > 0$ describes the way that announcements affect the volatility. In particular, for each index $i = 1, \ldots, K$ there are J_i regions. $I_{ij}(t_i^*)$ is a region-based indicator variable indicating that region j, $j = 1, \ldots, J_i$, affects volatility through a jump of size $\gamma_{ij} > -1$ that occurred at time $t_i^* \leq t$, and through an exponential decay with rate $r_{ij} > 0$. The announcement surprises of index i at time t_i^* are expressed through a variable $Z_{it_i^*}$ and the degree of surprise is determined by threshold points c_{ij}, where $c_{i,J_i} := \infty$. Finally, our model specification is completed by allowing pre-announcement volatility jumps $s_i > -1$ occurring one time period before the announcements.

We call the above model formulation Flexible Threshold-GARCH model and we emphasize its rich flexibility: there are many models with varying number of parameters specified by the number of regions J_i and the number of indices K that affect volatility. Call m such a model, M the set of all models and use superscript m to denote the corresponding parameters within each model. Then, for each $m \in M$, the parameter vector is

$$\theta_m = \left(\mu^m, \sigma^m, \alpha_1^m, \alpha_2^m, K^m, J_i^m, \gamma_{ij}^m, s_i^m, r_{ij}^m, c_{ij}^m, \quad i = 1, \ldots, K^m, \quad j = 1, \ldots, J_i^m\right).$$

Therefore, our target under a Bayesian inference setup is to estimate the posterior probability of each model m and the posterior densities of θ_m within each model m.

An interesting feature of the above formulation is the way H_t is modelled in (4). Instead of using a functional form such as an exponential function that ensures positivity of H_t, we prefer to directly use a formulation which results to an immediate sensible interpretation of the parameters γ_{ij}, r_{ij} and s_i with respect to the changes occurred to σ^2. A related additive formulation was suggested in Dellaportas et al. (2007). As a side effect, the prior densities of γ_{ij} and s_i are restricted to the regions that make H_t positive.

The effect of the specification (4) in the volatility process is depicted in Figure 1. There are macroeconomic announcements at time points 3 and 10 of index 1, with different degrees of surprise, and at time 7 of index 2. Assume that pre-announcement jumps do not exist. The left panel depicts the three components that are added to produce the process H_t at the right panel. Note that jump I_{12} corresponds to a decrease in the volatility following the release of a macroeconomic index, since we allow positive and negative jumps (γ_{ij}) occurring in the volatility.

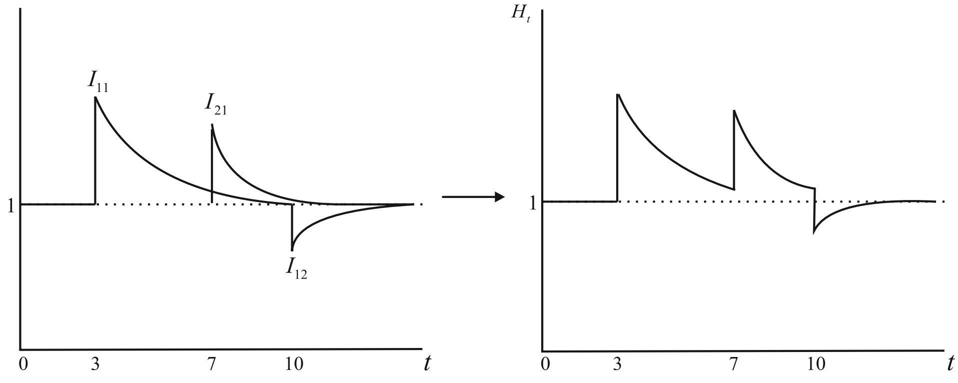

Figure 1: The effect of explanatory variables in the threshold process.

3.2. A Spline-GARCH model

Another way to capture, in a non-linear fashion, the volatility shifts caused by news announcements is to use a modification of the Spline-GARCH model of Engle & Rangel (2008). We propose an amendment of this model by applying the exponential spline function to the time instances in which announcements occur and not to equally spaced time intervals as in Engle & Rangel (2008). The number of spline knots, denoted by K, is assumed to be unknown. Hence, our proposed Spline-GARCH model is based on (2) and (3), but (4) and (5) are replaced by

$$H_t = \exp\left(w_0 t + \sum_{i=1}^{K} \left[w_i \left((t - t_i^*)_+ \right)^2 + \gamma_i Z_{it_i^*} + s_i \mathbf{1} \left(t = t_i^* - 1 \right) \right] \right), \tag{6}$$

where $(t - t_i^*)_+ = (t - t_i^*)$ if $t > t_i^*$ and zero otherwise. The parameter vector in this model, for each model $m \in M$, is

$$\theta_m = \left(\mu^m, \sigma^m, \alpha_1^m, \alpha_2^m, w_0^m, K^m, w_i^m, \gamma_i^m, s_i^m, \quad i = 1, \ldots, K^m \right).$$

In this way when a variable $Z_{it_i^*}$ is included in the model a new knot is generated, causing shifts in the variance process. By performing model selection on the number of variables included in the model we automatically determine the number of knots K^m in this quadratic spline. Thus, the more variables included in the model the more frequent shifts take place, while their sharpness depends on w_i. As in Engle & Rangel (2008), the values of explanatory variables $Z_{it_i^*}$ and the pre-announcement effect affect exponentially the volatility process.

4. Inference

4.1. Priors

Our prior specification is non-informative, in the sense that vague proper priors are proposed, but special care is taken to avoid issues of Lindley's paradox (Bartlett, 1957) and model identifiability. For example, as r_{ij} increases so that the exponent in (4) approaches zero, H_t as well as the likelihood function remain constant. Thus, the integrability of the posterior density conditional on any given model requires the use of proper priors. We performed the following (necessary) Bayesian data analysis exercise. For the models used in the empirical analysis, we increased the standard deviations of all prior densities by a factor of ten. We then examined the robustness of our results to the prior dispersion by inspecting posterior model probabilities and posterior summaries of interest such as volatility predictions. Although the posterior model probabilities do change slightly, the predictive variance densities based on model averaging remain identical. This certifies that our prior densities are vague enough not to affect the posterior summaries of interest. The details of these experiments can be found in Section 5.3.

The prior probabilities of the discrete densities on the model space and on threshold points are set to be discrete uniforms over all possible models and distinct observed values of announcement surprises respectively. Within each model, prior densities for the parameters of the Flexible Threshold-GARCH model are chosen to be as follows. We first apply transformations to the real line, $g_{ij} = \log(\gamma_{ij} + 1)$, $\varsigma_i = \log(s_i + 1)$ and $\rho_{ij} = \log(r_{ij})$, $i = 1, \ldots, K$, $j = 1, \ldots, J_i$. We then place non-informative priors $\mu \sim N(0, 1)$, $\sigma^2 \sim IG(10^{-5}, 10^{-5})$, $\alpha_1, \alpha_2 \sim U(0, 1)$ with $\alpha_1 + \alpha_2 < 1$, $g_{ij} \sim N(0, 0.4^2)$, $\varsigma_i \sim N(0, 0.4^2)$, $\rho_{ij} \sim N(0, 2^2)$. Recall from Section 3.1 that these prior densities are constrained so that $H_t > 0$. For the Spline-GARCH model we use $w_i \sim N(0, 0.003^2)$, $\gamma_i \sim N(0, 3^2)$, $s_i \sim N(0, 0.5^2)$.

4.2. The population reversible jump MCMC algorithm

In problems with complex multi-modal distributions standard vanilla MCMC samplers may fail to efficiently move around the support of the target distribution. A way to deal with these problems is to adopt Population-based MCMC methods (Geyer (1991); Gilks et al. (1994); Liu et al. (2000); Liu (2001); Liang & Wong (2001); Jasra et al. (2007a)). In this setting, MCMC operates by embedding the target density into a sequence of $\ell = 1, \ldots, L$ independent distributions obtained by simulating L parallel chains, whilst allowing the chains to interact via various moves. Jasra et al. (2007b) proposed an extension of these methods to transdimensional parameter spaces and we adopt ideas from this paper to develop the proposed inference algorithm.

Our population reversible jump algorithm includes the following basic moves. An exchange move is used to swap information between two adjacent (in terms of temperature) chains by exchanging all variables and associated parameters between them. In a crossover move only a fraction of the variables with their associated parameters is exchanged between two randomly chosen chains. A mutation move is used to update a chain according to a reversible jump step as suggested by Green (1995), which includes addition, deletion, replacement, split and merge moves. The basic steps of the Flexible

Threshold-GARCH reversible jump algorithm that obtains samples (θ_m, m), $m \in M$, on spaces of varying dimension follows. The specific details can be found in the Appendix.

The reversible jump Algorithm

- `Initialize the chain and sweep over the following:`

  1. `Randomly add, delete or replace an index variable.`
  2. `For all index variables, randomly propose to split or merge the current threshold points.`
  3. `Update all remaining parameters in the current model through adaptive random walk Metropolis Hastings kernels.`

The above algorithm is enriched by applying a population algorithm. Denote by π be the reversible jump invariant distribution with states (m, θ_m). We construct two auxiliary distributions $\pi_\ell \propto \pi^{\zeta_\ell}$, $1 = \zeta_1 > \zeta_2 > 0$, with ζ_ℓ denoting the inverse temperature parameter in chain ℓ.

The population Algorithm

- Run 3 parallel Markov chains each one with target densities π^{ζ_ℓ}.

- Every 10 iterations choose randomly between

  - an exchange move which changes the states (m, θ_m) between two randomly chosen chains which are adjacent in terms of temperature.

  - a crossover move which changes a randomly chosen subset of variables included in m and its associated parameters θ_m between two randomly chosen chains.

- In the rest of the iterations perform a mutation move which updates the chains according to the reversible jump Algorithm.

The key intuition behind the population algorithm is that we want enough chains to explore efficiently the model space but not too many so that the algorithm is expensive in terms of CPU time. In our data the first (untempered) chain had addition/deletion acceptance rates for different exchange rates between 4.7% and 5.4%, whereas in the third tempered chain the corresponding proportions were 7% to 9% respectively. This led to satisfactory mixing of the chain; see the MCMC diagnostics in Section 5.3. For a sufficient intercommunication between chains we followed Jasra etal. (2007b) and exchanged information using either an exchange or a crossover move every ten iterations.

The temperature ladder is specified as $\zeta_\ell = z^{\ell-1}$, where the scalar z, $0 < z < 1$, is calibrated during the burn-in period of the algorithm as follows. We started from $z = 0.8$ and every 50 exchange move proposals we set $z' = z + \delta(0.5 - \alpha)$, where $\delta > 0$ is a pre-specified sensitivity parameter and α is the acceptance rate of the exchange move during the latest 50 exchange move sweeps. The constant $\delta = 0.1$ is chosen so that the exchange move is accepted about half of the time (Liu, 2001).

4.3. Prediction

We base all our predictions to Bayesian model averaging (see Raftery et al. (1997); Liu & Maheu (2009)). The posterior distribution of a quantity Δ, which in our case corresponds to the model averaging variance forecasts, given the data $\mathcal{D}$ and the models $m \in M$, is given as

$$f(\Delta \mid \mathcal{D}) = \sum_{m \in M} f(\Delta \mid m, \mathcal{D}) f(m \mid \mathcal{D}),$$

which is an average of the posterior predictive distribution under each model m,

$$f(\Delta \mid m, \mathcal{D}) = \int f(\Delta \mid \theta_m, m, \mathcal{D}) f(\theta_m \mid m, \mathcal{D}) d\theta_m$$

weighted by the posterior model probabilities $f(m \mid \mathcal{D})$.

In our context each state of the Markov chain corresponds to a model m and associated parameters θ_m, so each calculated volatility is being averaged across both model and parameter posterior densities. To avoid confusion, we call the Flexible Threshold-GARCH and Spline-GARCH as 'model specifications'. Model averaging takes place with respect to models m within a model specification.

In the empirical analysis we produce one day ahead out-of-sample forecasts based on monthly updated in-the-sample parameter estimates. This means that we reran the MCMC algorithm twelve times to produce daily out-of-sample forecasts for a year. Since we deal with scheduled announcements, when forming the out-of-sample forecasts the day of occurrence of each announcement is known a priori, thus we can evaluate the pre-announcement effect. For the days following the release, when the announcement outcome is known, we can directly perform predictions based on the estimated predictive density conditional on the announced surprise. In the exact date of the release the outcome is not known a priori, and therefore we cannot predict $I_{ij}(t_i^*)$, thus we propose two ways to perform forecasting. One, adopted in our main analysis, is to use an 'empirical Bayes' type estimator that estimates the probability of the not yet observed $I_{ij}(t_i^*)$ with its past (in-the-sample) sample average. The second estimator relies on the hypothesis that the news announcement will lie close to the 'Bloomberg consensus', so that $I_{i1}(t_i^*) = \mathbf{1}\left(c_{i1} \leq Z_{it_i^*} < c_{i,2}\right)$. In our empirical study it turned out that the former estimator is slightly better than the latter; but we suggest that choice between them should be based on the degree of belief one has to the first or the second assumption underlying their choices.

To evaluate the predictive performance of each model specification we rely first on robust loss functions evaluated based on a true volatility proxy. We adopt the MSE and QLIKE loss functions defined as

$$\begin{aligned} \text{MSE} &= \frac{1}{T}\sum_{t=1}^{T}\left(\hat{\sigma}_t^2 - h_t\right)^2 \\ \text{QLIKE} &= \frac{1}{T}\sum_{t=1}^{T}\left(\log h_t + \frac{\hat{\sigma}_t^2}{h_t}\right) \end{aligned}$$

where h_t denotes variance forecast for day t and the proxy $\hat{\sigma}^2$ is the realized variance obtained using 1-minute high frequency data; see Andersen & Bollerslev (1998a); Barndorff-Nielsen & Shephard (2002, 2004). Patton (2011) notes that MSE and QLIKE loss functions are robust in the sense that the ranking of any two volatility forecasts is the same whether it is done using the true conditional variance or some conditionally unbiased volatility proxy.

A second criterion used to evaluate the predictive ability of the models is based on the optimal pools methodology of Geweke & Amisano (2011), also employed by Jensen & Maheu (2010). Model comparison under the optimal prediction pool methodology mixes the predictive densities of each model specification and sets the mixing weights equal to the value that maximizes the log pooled predictive score. The larger the mixing weight the more important the model specification is in predicting future outcomes. Denote with $f(y_t|\mathcal{F}_{t-1}, A)$ the probability density for the exchange rate series y_t under a model specification A and the available information set $\mathcal{F}_{t-1}$. The respective log-predictive score is given as

$$LS = \sum_{t=1}^{T} \log(f(y_t|\mathcal{F}_{t-1}, A)).$$

Then, for a pool model specifications $A_1, ..., A_n$, we seek to find the weights w_i^* that maximize the log predictive score function

$$w_i^* = \arg\max_{w_i} \sum_{t=1}^{T} \log\left(\sum_{i=1}^{n} w_i f(y_t|\mathcal{F}_{t-1}, A_i)\right).$$

5. Empirical study

5.1. The data

Our data set consists of 3131 daily observations from 1/1/2002 (i.e. since the introduction of the euro currency) up to 31/12/2013 of the Euro-dollar (EURUSD), British pound-dollar (GBPUSD) and Dollar-Swiss franc (USDCHF) exchange rates. The data were obtained from Bloomberg. The diagnostics with respect to the macroeconomic announcements effects on the volatility of exchange rates are obtained on basis of the full sample. The predictive ability of the model specifications is evaluated for 261 observations in 2013 by forming one step ahead out-of-sample forecasts and estimating the parameters and model probabilities on a rolling window basis, as described in Section 4.3. For this exercise 1-minute high frequency data for the year 2013 (approximately 370,000 observations) over the three exchange rates were employed to calculate the required realized volatilities.

We used surprises of 15 monthly U.S. scheduled macroeconomic announcements, that we believed that they may affect the volatility process and/or have been found to explain volatility fluctuation by previous empirical studies. In the full sample (2002-2013) there are 144 observations for each macroeconomic announcement. The surprises are defined as the absolute percentage difference of the realised minus the Bloomberg consensus reported values. The macroeconomic announcements considered are presented in Table 1.

Under the Flexible Threshold-GARCH specification, model selection takes place with respect to the fifteen macroeconomic announcements and to the number and location of threshold points. We report in the last column of Table 1 the maximum number of regions, which coincides with the number of different (distinct) observations of each macroeconomic announcement.

Table 1: List of the macroeconomic announcements (K) and surprises ($Z_{it_i^*}$) descriptive statistics; minimum values are equal to zero.

	Macroeconomic announcement	Mean ($\times$100)	St. dev. ($\times$100)	Median ($\times$100)	Maximum ($\times$100)	Maximum number of regions (J_i) (different observations out of 144)
1	GDP QoQ (Annualized)	0.33	0.34	0.20	1.70	15
2	Industrial Production	0.29	0.26	0.25	2.00	14
3	Durable Goods Orders	1.77	1.57	1.35	8.20	44
4	Wholesale Inventories	0.41	0.32	0.30	1.80	15
5	Advance Retail Sales	0.39	0.35	0.30	1.80	14
6	Housing Starts	5.47	4.44	4.58	29.56	140
7	ISM Manufacturing	2.99	2.56	2.31	14.23	130
8	ISM Non-Manufacturing	3.72	2.95	3.29	15.10	138
9	Leading Indicators	0.14	0.13	0.10	0.50	6
10	Consumer Confidence	6.01	6.11	4.00	31.99	135
11	Consumer Price Index (MoM)	0.10	0.09	0.10	0.40	5
12	Producer Price Index (MoM)	0.37	0.33	0.30	1.70	15
13	Trade Balance	5.84	5.07	4.74	27.78	139
14	Unemployment Rate	0.12	0.11	0.10	0.50	6
15	Personal Income	0.19	0.27	0.10	1.80	13

5.2. Results

We first report the results with respect to the effect of the macroeconomic announcements on the volatility of the three exchange rates under the Flexible Threshold-GARCH and Spline-GARCH specifications. The four models with the highest estimated posterior probability within each specification and exchange rate are presented in Table 2. Although most of the macroeconomic announcements are included in the four highest posterior density models of all specifications, we observe that the Consumer Confidence (no. 10) is present in most of the models across all specifications and exchange rates and that Unemployment Rate (no. 14) is present in the four best models under the Flexible Threshold-GARCH specification in both EURUSD and USDCHF currencies.

A quantity usually adopted to help identifying important regressors in high dimensional regression type problems is the marginal probability of inclusion (see Barbieri & Berger, 2004), calculated as the percentage of times a variable is observed in the transdimensional MCMC sample. Denoting with $\hat{f}(m \mid \mathcal{D})$ the estimated posterior probability of a model m, the probability of inclusion of macroeconomic announcement i is given as

$$p_i = \sum_{m \in M^i} \hat{f}(m \mid \mathcal{D}),$$

Table 2: The four models with the highest posterior probability for each exchange rate and their estimated posterior probabilities under the Flexible Threshold-GARCH and Spline-GARCH specifications.

	Flexible Threshold-GARCH			Spline-GARCH		
EURUSD						
Model	Variables	Probability	No. of parameters*	Variables	Probability	No. of parameters
First	6,10,13,14	0.0253	402	7	0.1860	8
Second	4,10,11,13,14	0.0238	301	7,10	0.1483	11
Third	10,11,13,14	0.0195	290	6,7	0.1258	11
Fourth	10,11,14	0.0167	153	6,7,10	0.0536	14
GBPUSD						
First	7,10	0.1035	270	3,10	0.2601	11
Second	3,7,10	0.0639	323	10	0.2502	8
Third	7,8,10	0.0281	423	3,4,10	0.1671	14
Fourth	8,10	0.0264	244	4,10	0.0921	11
USDCHF						
First	6,8,10,14	0.0945	392	3,4,10	0.2161	14
Second	6,7,8,10,14	0.0701	543	10	0.0911	8
Third	6,8,10,14,15	0.0362	425	7,10	0.0646	11
Fourth	6,8,10,11,14	0.0241	419	3,10	0.0597	11
*The number of parameters under the threshold point combination with the highest posterior probability is reported.						

where M^i denotes all the models visited by the algorithm in which macroeconomic announcement i is present.

Table 3 lists the macroeconomic announcements according to their marginal probability of inclusion in the two model specifications. The results are in accordance to those analysed on basis of the four most probable models of Table 2. It is interesting to notice that different announcements affect each exchange rate. For example Unemployment rate, found as an important announcement in previous studies (Ederington & Lee, 1993; Andersen & Bollerslev, 1998b; Nikkinen & Sahlström, 2001), has marginal probability of inclusion 100% in EURUSD and USDCHF (Flexible Threshold-GARCH specification), but below 9% for GBPUSD. However, we note that Consumer Confidence has high probability of inclusion in all exchange rates. Another interesting observation is that announcements related to actual growth such as GDP, Industrial Production and Personal Income display low probability of inclusion (below 30%) under both specifications for all currencies. On the other hand sample surveys for the condition of the economy, such as ISM Manufacturing, ISM Non-Manufacturing and Consumer Confidence seem to affect more the volatility of exchange rates. Presumably, this can be attributed to the

Table 3: Marginal probability of inclusion of macroeconomic announcements

	Macroeconomic announcement	Flexible Threshold-GARCH			Spline-GARCH		
		EURUSD	GBPUSD	USDCHF	EURUSD	GBPUSD	USDCHF
1	GDP QoQ (Annualized)	0.1398	0.1264	0.1034	0.0688	0.0331	0.0272
2	Industrial Production	0.2525	0.0803	0.0596	0.0474	0.0210	0.0311
3	Durable Goods Orders	0.0349	0.3581	0.1961	0.0245	**0.5851**	**0.5134**
4	Wholesale Inventories	0.5648	0.2021	0.0753	0.0204	0.3262	0.4945
5	Advance Retail Sales	0.1628	0.1761	0.0558	0.0345	0.0308	0.0461
6	Housing Starts	0.3817	0.1312	0.7811	0.3130	0.0066	0.1439
7	ISM Manufacturing	0.3917	**0.7699**	0.5246	**0.9393**	0.0326	0.4247
8	ISM Non-Manufacturing	0.1400	0.2710	**1.0000**	0.0395	0.0199	0.0157
9	Leading Indicators	0.2412	0.1151	0.2330	0.0813	0.0195	0.0757
10	Consumer Confidence	**0.8311**	**1.0000**	0.7555	**0.4167**	**0.9999**	**0.9293**
11	Consumer Price Index (MoM)	0.5086	0.1085	0.2919	0.0430	0.0321	0.0374
12	Producer Price Index (MoM)	0.0703	0.0658	0.1406	0.0168	0.0160	0.0156
13	Trade Balance	0.4673	0.1072	0.0699	0.0048	0.0043	0.0209
14	Unemployment Rate	**1.0000**	0.0885	**1.0000**	0.0677	0.0471	0.0314
15	Personal Income	0.1088	0.0765	0.2829	0.1403	0.0154	0.0503

The two announcements with the highest marginal probability of inclusion for each case are denoted with bold.

fact that changes in actual growth have been mainly absorbed by results of indices that provide advance signs for the condition of the economy, such as confidence and sentiment reports, or other announcements with this characteristic such as Durable Goods Orders or Housing Starts, the latter being also related to the financial crisis of 2008.

It is interesting to investigate whether days with certain macroecoeconomic announcements provide better predictive ability when compared with non-announcements days. We examined the one-day-ahead predictive likelihood of the model specifications as in Jensen & Maheu (2013), but we did not observe any significant difference in days with largest one-day-ahead predicitive likelihoods in any announcement. As an example, see Figure 2 in which the Consumer Confidence and Unemployment rate announcements are depicted together with predictive likelihoods.

Table 4 presents summary statistics for the posterior densities of pre-announcement effects under the highest posterior probability model of each specification. There is a sign agreement in relevant specifications that represents a drop of volatility before the announcement of Consumer Confidence. It is evident that the pre-announcement effect depends on the announcement and the related exchange rate and can be positive or negative. To this respect we observe an increase in the volatility of EURUSD and USDCHF

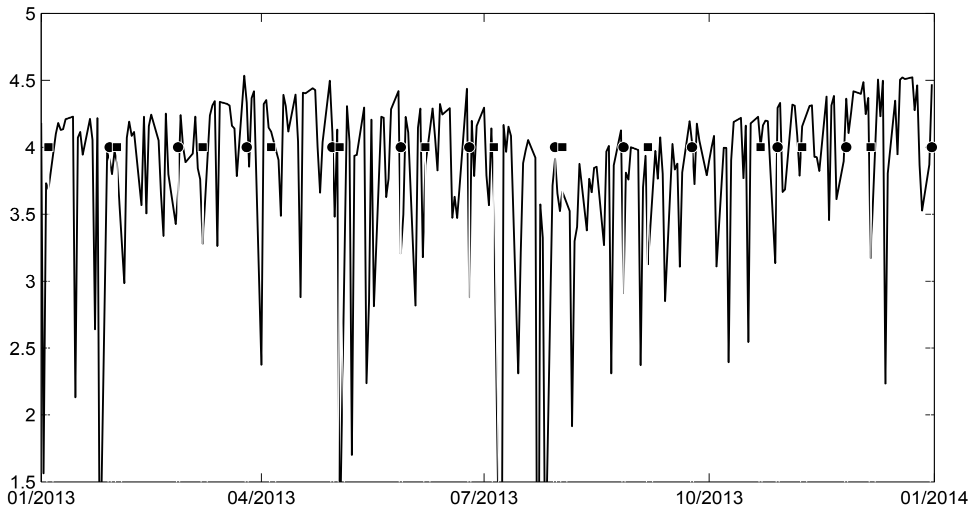

Figure 2: One-day-ahead predictive likelihood under the Flexible Threshold-GARCH model; EURUSD currency. Circles represent the announcement days of the Consumer Confidence index and squares of the Unemployment rate.

the day before the announcement of Unemployment rate.

Figures 3 and 4 depict estimates of the posterior means of jump sizes γ_{ij} and decay rates r_{ij} across all observed threshold points for the model with the highest posterior probability of Table 2, under the Flexible Threshold-GARCH specification for EURUSD exchange rate. It seems that the hypothesis that the larger the degree of surprise (defined on basis of the threshold points c_{ij}), the larger the size of jump (γ_{ij}) in the volatility is not confirmed. As noted in several event studies (Jones et al. (1998); Hautsch & Hess (2002); Laakkonen (2004); Andersen et al. (2007b)) since the effect of an announcement is dependent to past or anticipated announcements, it is hard to isolate it in a changing macroeconomic environment.

However certain interesting features are observed. The 197 estimated size of jump parameters range from -0.66 (66% decrease in the volatility at the day of announcement), up to 2.02 (202% increase). The 90% of these parameters range from -0.50 to 1.17. Respectively the 90% of rates of decay r_{ij} estimates range from 0.04 up to 29.4. The lower the value of the rate of decay the longer the impact of a given jump in volatility. Half of them lie below the unity, which corresponds to an impact of about 36.8% of the original jump at the day following the announcement, 13.5% two days after the

Table 4: Pre-announcement effect (ς_i) on the volatility based on models with the highest posterior probability; posterior standard deviations are in brackets.

Macroeconomic announcement		Flexible Threshold-GARCH			Spline-GARCH		
		EURUSD	GBPUSD	USDCHF	EURUSD	GBPUSD	USDCHF
3	Durable Goods Orders					-0.161 [0.130]	-0.395 [0.115]
4	Wholesale Inventories						-0.222 [0.116]
6	Housing Starts	0.373 [0.023]		0.0001 [0.0004]			
7	ISM Manufacturing		-0.088 [0.080]		0.325 [0.127]		
8	ISM Non-Manufacturing			0.149 [0.0003]			
10	Consumer Confidence	-0.340 [0.007]	-0.664 [0.086]	-0.337 [0.014]		-0.661 [0.125]	-0.377 [0.128]
13	Trade Balance	-0.008 [0.010]					
14	Unemployment Rate	0.447 [0.012]		0.111 [0.012]			

announcement and 5% three days after an announcement. The 25% of estimates lie below the value 0.23, which corresponds to an effect with rather long memory, having 10% of its initial magnitude even 10 days after an announcement. These results are in line with the findings of Evans & Lyons (2005, 2008) and Brenner et al. (2009), that found macroeconomic releases to affect volatility several days following the announcement.

We now turn to forecasting aspects of our specifications. We stress that forecasts are not based on one model but on model averaging. This fact allows borrowing strength between parameter estimates across models and therefore obtain robust volatility forecasts. For comparative reasons besides the Flexible-Threshold GARCH and Spline-GARCH, we report the results under a typical GARCH(1,1) specification.

Figure 5 depicts in-the-sample (2002-2012) variance estimates and out-of-sample (2013) variance predictions under each model specification for the EURUSD currency. These can be visually compared with squared residuals and realized volatility respectively. It is clear that specifications which use information from the macroeconomic announcements are more spiky than the usual smooth GARCH estimates. This is in line with the findings of various empirical studies (DeGennaro & Shrieves, 1997; Andersen & Bollerslev, 1998b; Bollerslev et al., 2000; Bauwens et al., 2005), in which the effects of news on the volatility are found to be stronger than that of GARCH effects.

In Figure 6 the out-of-sample volatility predictions are displayed for the three exchange rates under the Flexible Threshold-GARCH specification. The 'empirical Bayes' and 'Bloomberg consensus' estimators are nearly indistinguishable. In Table 5 we present the

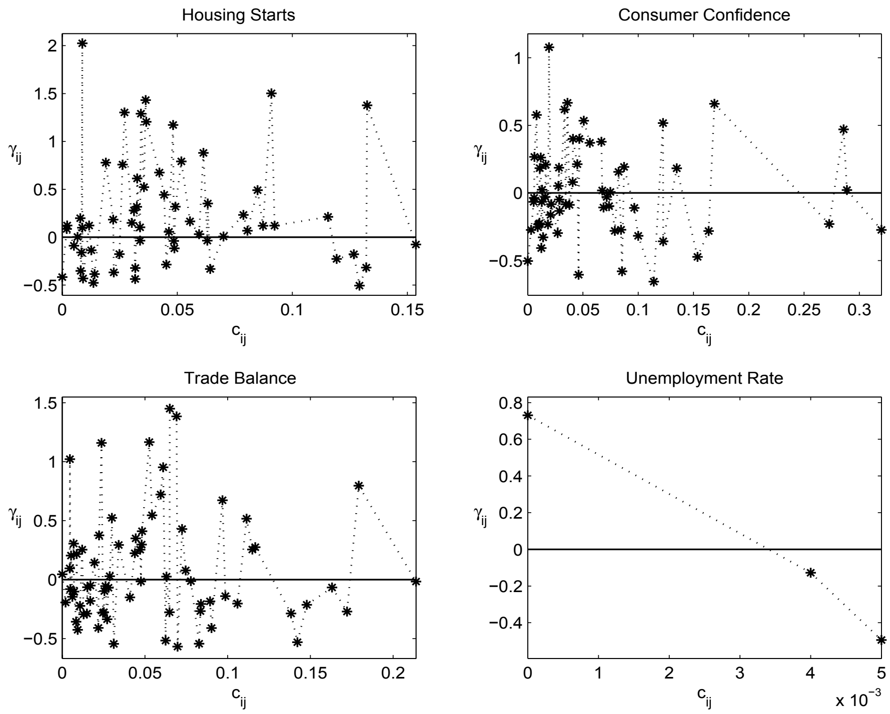

Figure 3: Posterior means of jump size parameters γ_{ij} against announcement surprises threshold points, under the first best model of the Flexible Threshold-GARCH specification for the EURUSD exchange rate.

forecast errors, log-predictive scores and optimal pool weights for all exchange rates under all specifications. We observe that under all criteria the Flexible Threshold-GARCH specification dominates over the Spline-GARCH and GARCH specifications for the EURUSD and USDCHF exchange rates. For GBPUSD the Flexible Threshold-GARCH and Spline-GARCH specifications perform approximately equally well when evaluated on basis of the QLIKE loss function and the log-predictive score. Flexible Threshold-GARCH displays lower Mean Square Errors (MSEs) than Spline-GARCH, but the latter has higher optimal pool weight. In view of that, one could use both specifications to predict GBPUSD volatility on a model averaging basis. Under all criteria the two specifications that use information from news announcements provide better forecasts compared to a typical GARCH(1,1) specification for all three exchange rates. These differences in the forecasting performance based on MSE and QLIKE among specifications do not change if we consider only announcement days or if we exclude the announcement days.

Table 5: Forecast errors, log-predictive scores and optimal pool weights. Bold denote the specification that provides better forecasts under the respective criterion; results based on 'empirical Bayes' estimators; 'Bloomberg consensus' estimators in brackets (see Section 4.3).

Model	Flexible Threshold-GARCH	Spline-GARCH	GARCH
	EURUSD		
	In-the-sample errors based on squared residuals		
MSE ($\times 10^{-10}$)	**49.5646**	51.2658	51.7309
QLIKE	**-9.2718**	-9.2330	-9.2214
	Out-of-sample errors based on squared residuals		
MSE ($\times 10^{-10}$)	**10.7935** [10.8402]	10.9648 [10.9693]	11.0077
QLIKE	**-9.7958** [-9.7908]	-9.7759 [-9.7755]	-9.7724
	Out-of-sample errors based on realized volatility		
MSE ($\times 10^{-10}$)	**1.9276** [1.9330]	2.1548 [2.1613]	2.1743
QLIKE	**-9.5981** [-9.5958]	-9.5843 [-9.5835]	-9.5839
	Out-of-sample log-predictive score		
LS ($\times 10^3$)	**1.0385** [1.0379]	1.0359 [1.0358]	1.0355
	Out-of-sample optimal pool weights		
w_i^*	**1.0000**	0.0000	0.0000
	GBPUSD		
	In-the-sample errors based on squared residuals		
MSE ($\times 10^{-10}$)	**44.8740**	45.8647	46.0003
QLIKE	**-9.4958**	-9.4660	-9.4519
	Out-of-sample errors based on squared residuals		
MSE ($\times 10^{-10}$)	**12.2226** [12.2605]	12.2693 [12.2563]	12.3659
QLIKE	**-9.7226** [-9.7199]	**-9.7230** [**-9.7241**]	-9.7142
	Out-of-sample errors based on realized volatility		
MSE ($\times 10^{-10}$)	**1.5697** [1.5774]	1.5739 [1.5709]	1.6054
QLIKE	**-9.6897** [-9.6885]	**-9.6897** [**-9.6900**]	-9.6874
	Out-of-sample log-predictive score		
LS ($\times 10^3$)	**1.0290** [1.0286]	**1.0290** [**1.0291**]	1.0278
	Out-of-sample optimal pool weights		
w_i^*	0.3479	**0.5461**	0.1060
	USDCHF		
	In-the-sample errors based on squared residuals		
MSE ($\times 10^{-10}$)	**314.3459**	341.6850	346.6429
QLIKE	**-9.1037**	-9.0463	-9.0308
	Out-of-sample errors based on squared residuals		
MSE ($\times 10^{-10}$)	**18.9276** [19.0211]	19.0142 [19.0284]	19.0126
QLIKE	**-9.4724** [-9.4682]	-9.4655 [-9.4651]	-9.4640
	Out-of-sample errors based on realized volatility		
MSE ($\times 10^{-10}$)	**4.8349** [4.8923]	4.9494 [4.9245]	5.0590
QLIKE	**-9.2579** [-9.2521]	-9.2536 [-9.2533]	-9.2502
	Out-of-sample log-predictive score		
LS ($\times 10^3$)	**0.9963** [0.9957]	0.9954 [0.9954]	0.9952
	Out-of-sample optimal pool weights		
w_i^*	**1.0000**	0.0000	0.0000

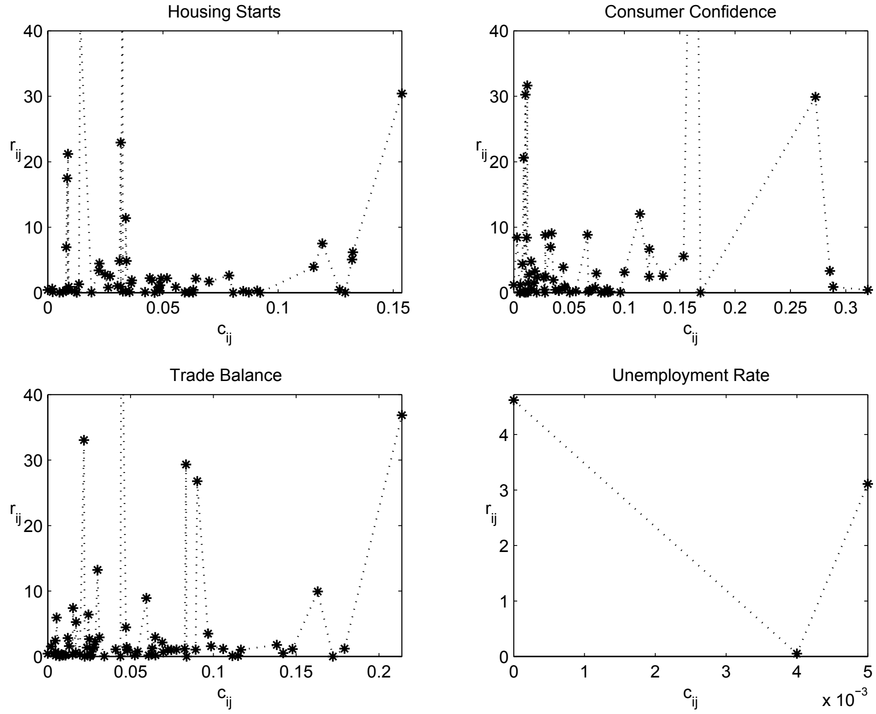

Figure 4: Posterior means of decay rates r_{ij} against announcement surprises threshold points, under the first best model of the Flexible Threshold-GARCH specification for the EURUSD exchange rate.

5.3. MCMC diagnostics and sensitivity analysis

Table 6 presents the associated acceptance rates and CPU time required under each specification for all exchange rates. As expected, the two auxiliary tempered chains display higher addition, deletion and replacement acceptance rates. The calibration of the temperature ladder seems to work well since the exchange move is accepted about half of the times. Finally, the adaptive MCMC works rather well with acceptance rates close to the targeted 0.234 in all cases. Figure 7 presents posterior model probabilities and associated traces for the Flexible Threshold-GARCH specification under all exchange rates, indicating that the mixing of the MCMC chain was good, with models varying between one and thirteen variables.

A proper Bayesian data analysis requires a sensitivity analysis on the effect of prior densities to the posteriors. As reported in Section 4.1, this is achieved by increasing the standard deviation of all parameter prior densities by factors of two, five and ten, and inspecting the changes in the resulting posterior densities. Table 7 demonstrates

Table 6: MCMC diagnostics

	EURUSD			GBPUSD			USDCHF		
Flexible Threshold-GARCH									
CPU time (hours)	138.6			59.3			99.2		
Acceptance rate	1^{st} chain	2^{nd} chain	3^{rd} chain	1^{st} chain	2^{nd} chain	3^{rd} chain	1^{st} chain	2^{nd} chain	3^{rd} chain
Addition	0.0539	0.0676	0.0879	0.0513	0.0665	0.0794	0.0478	0.0576	0.0720
Deletion	0.0522	0.0700	0.0877	0.0505	0.0656	0.0809	0.0466	0.0589	0.0710
Replacement	0.0059	0.0096	0.0160	0.0048	0.0083	0.0115	0.0042	0.0061	0.0099
Split	0.4114	0.3990	0.3943	0.4851	0.4933	0.5097	0.4398	0.4411	0.4367
Merge	0.4186	0.4070	0.4047	0.4819	0.4908	0.5023	0.4521	0.4546	0.4562
Update μ^m	0.2347	0.2350	0.2339	0.2326	0.2348	0.2341	0.2332	0.2326	0.2362
Update α_1^m, α_2^m	0.2323	0.2338	0.2376	0.2345	0.2325	0.2346	0.2361	0.2340	0.2327
Update σ^m	0.2351	0.2357	0.2338	0.2338	0.2343	0.2301	0.2348	0.2340	0.2329
Update s_i^m	0.2351	0.2309	0.2343	0.2350	0.2327	0.2278	0.2313	0.2336	0.2332
Update γ_{ij}^m	0.2339	0.2347	0.2324	0.2371	0.2283	0.2320	0.2341	0.2333	0.2342
Update r_{ij}^m	0.2367	0.2354	0.2344	0.2311	0.2349	0.2282	0.2302	0.2354	0.2232
Exchange		0.5054			0.5116			0.5031	
Crossover		0.1798			0.2068			0.1232	
Spline-GARCH									
CPU time (hours)	132.6			133.9			128.6		
Acceptance rate	1^{st} chain	2^{nd} chain	3^{rd} chain	1^{st} chain	2^{nd} chain	3^{rd} chain	1^{st} chain	2^{nd} chain	3^{rd} chain
Addition	0.0360	0.0533	0.0682	0.0326	0.0471	0.0653	0.0342	0.0467	0.0595
Deletion	0.0371	0.0554	0.0900	0.0327	0.0497	0.0724	0.0344	0.0526	0.0815
Replacement	0.0036	0.0076	0.0146	0.0025	0.0043	0.0094	0.0034	0.0063	0.0118
Update μ^m	0.2349	0.2353	0.2330	0.2336	0.2341	0.2346	0.2355	0.2355	0.2336
Update α_1^m, α_2^m	0.2314	0.2332	0.2332	0.2312	0.2319	0.2334	0.2333	0.2316	0.2363
Update σ^m	0.2330	0.2333	0.2331	0.2344	0.2346	0.2332	0.2345	0.2339	0.2358
Update w_0^m	0.2339	0.2359	0.2347	0.2320	0.2341	0.2349	0.2378	0.2314	0.2352
Update w_i^m	0.2356	0.2336	0.2307	0.2337	0.2334	0.2402	0.2364	0.2318	0.2382
Update γ_i^m	0.2331	0.2333	0.2370	0.2353	0.2353	0.2302	0.2320	0.2430	0.2311
Update s_i^m	0.2362	0.2321	0.2364	0.2321	0.2354	0.2341	0.2348	0.2330	0.2338
Exchange		0.4964			0.4936			0.5099	
Crossover		0.2751			0.2991			0.2300	

Note: Results based on a single run on the full sample. All algorithms ran for 100.000 iterations with 10.000 iterations burn-in on an Intel Corei7 CPU. For the out-of-sample forecasts we ran the algorithms 12 times on a rolling window basis (see Section 4.3).

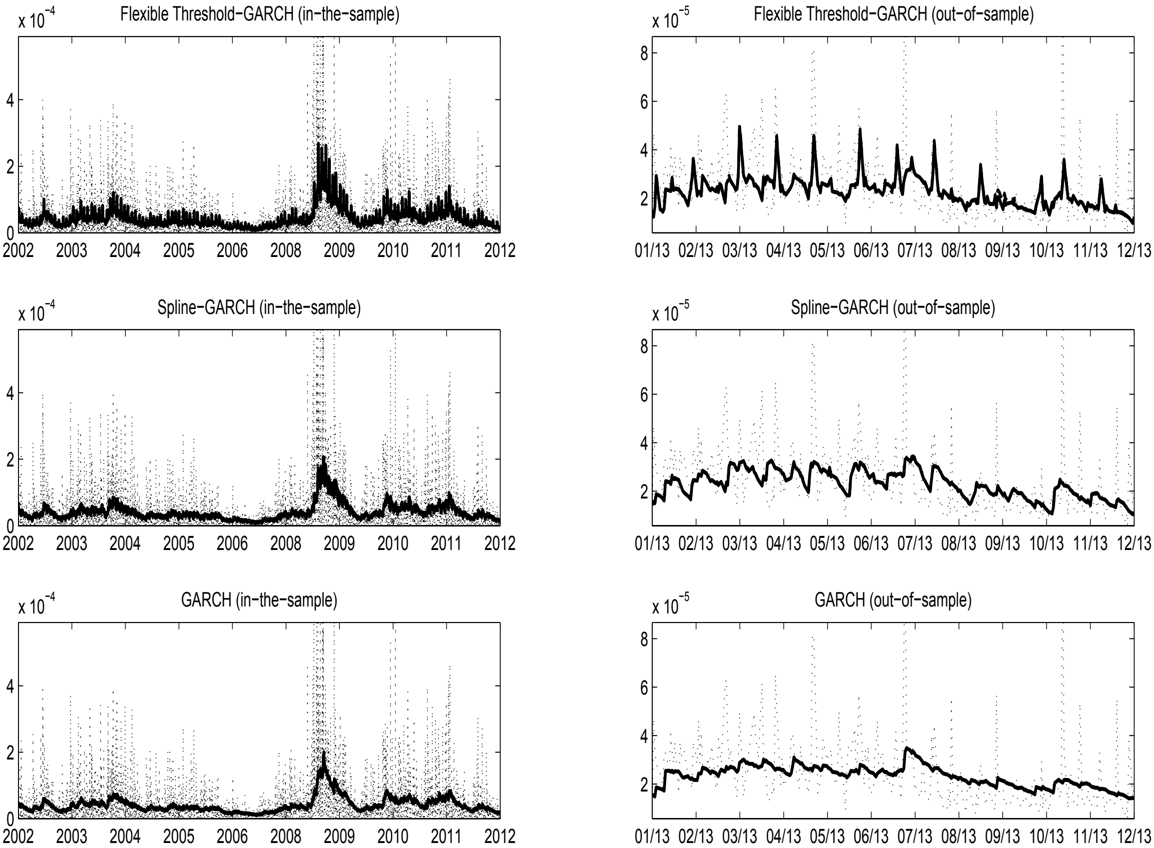

Figure 5: Estimated EURUSD volatilities in-the-sample and out-of-sample. Solid: volatility estimates (empirical Bayes estimator), dotted: squared residuals (left panel) and realized volatility (right panel)

that in the Flexible Threshold-GARCH specification, only marginal changes occur in the associated forecast errors. Moreover, the out-of-sample volatility forecasts remain almost identical under the two extreme prior specifications, see Figure 9. Similar results hold for the Spline-GARCH specification. In Table 7 we also report results based on varying MCMC samples that provide evidence that convergence has been achieved with sample size $100,000$. Plots of volatility forecasts (not reported here) illustrate that the forecasts are visually identical between $5,000$ and $100,000$ MCMC sample sizes.

6. Conclusions

We presented a new class of flexible threshold models for predicting volatilities of exchange rates that incorporate information from scheduled news announcements. For the dataset analysed, we have found strong evidence that suggests the use of such extra information can enrich the GARCH structures commonly used.

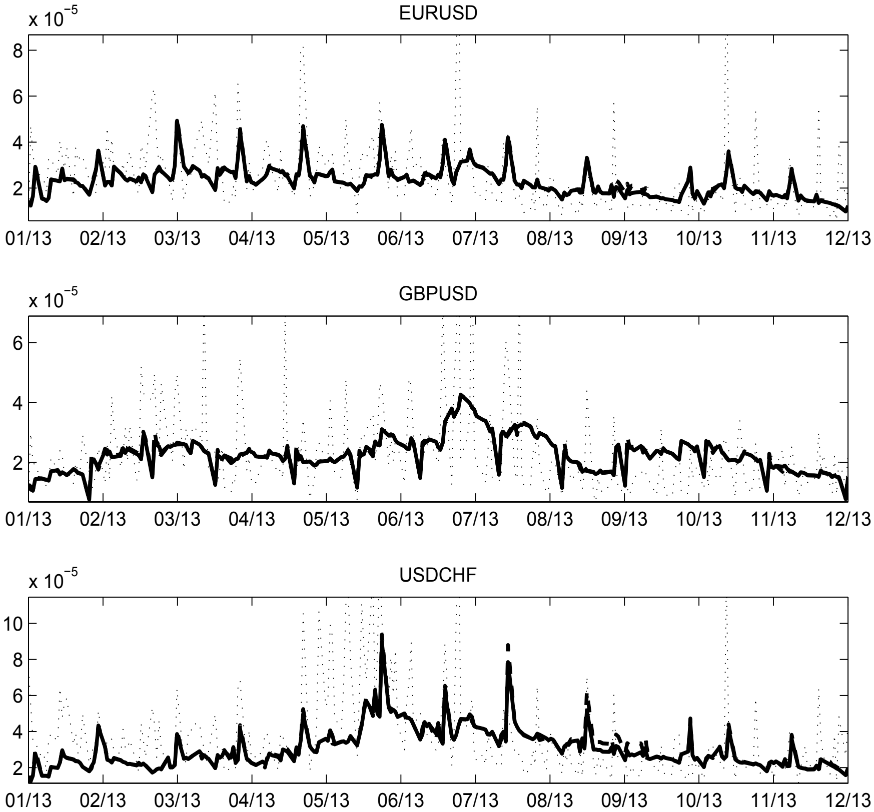

Figure 6: Out-of-sample volatility estimates under the Flexible Threshold GARCH specification for the three exchange rates. Solid: volatility estimates based on 'empirical Bayes' estimator, dashed: volatility estimates based on 'Bloomberg consensus' estimator, dotted: realized volatility

The use of our proposed threshold model is intuitively appealing, although it may suffer from the usual problem of over-fitting. For this purpose, we have advocated the use of model averaging prediction that exploits modern MCMC strategies that sample in a transdimensional space. Our detailed empirical analysis provided evidence that our model predicts rather well when compared with conditional volatility models and other nonparametric formulations such as splines.

Our statistical framework can be easily extended to the multivariate case, where given the variance estimates, once can employ a Dynamic Conditional Correlation (Engle, 2002) or a Cholesky-type parametrization (Dellaportas & Pourahmadi, 2012) to obtain covariance estimates; see Petralias (2010) for illustrations of such multivariate extensions.

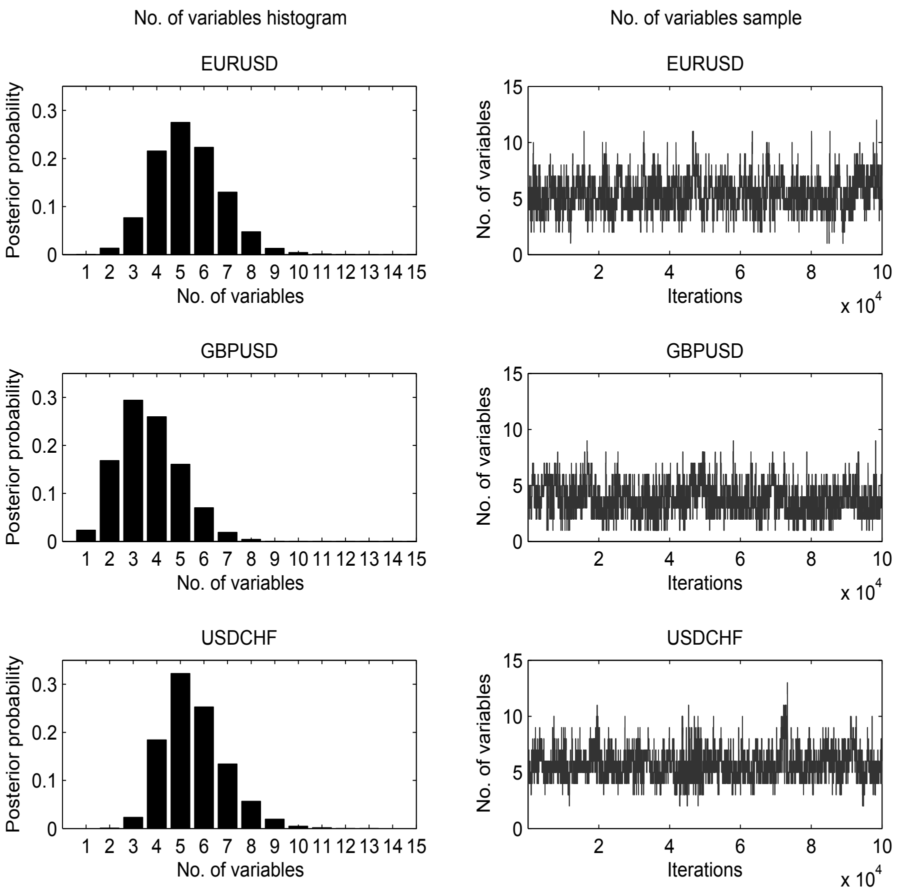

Figure 7: Estimated posterior probability for the number of variables in the model under the Flexible Threshold GARCH specification (left panel) and the number of variables across iterations (right panel).

Acknowledgments

We are grateful to Omiros Papaspiliopoulos for many useful comments that greatly improved a previous version of the paper. The second author was supported by the European Union, Seventh Framework Programme FP7/2007-2013 under grant agreement SYRTO-SSH-2012-320270.

References

Andersen, T. G., & Bollerslev, T. (1998a). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, *39*, 885–905.

Andersen, T. G., & Bollerslev, T. (1998b). Deutsche mark-dollar volatility: Intraday activity

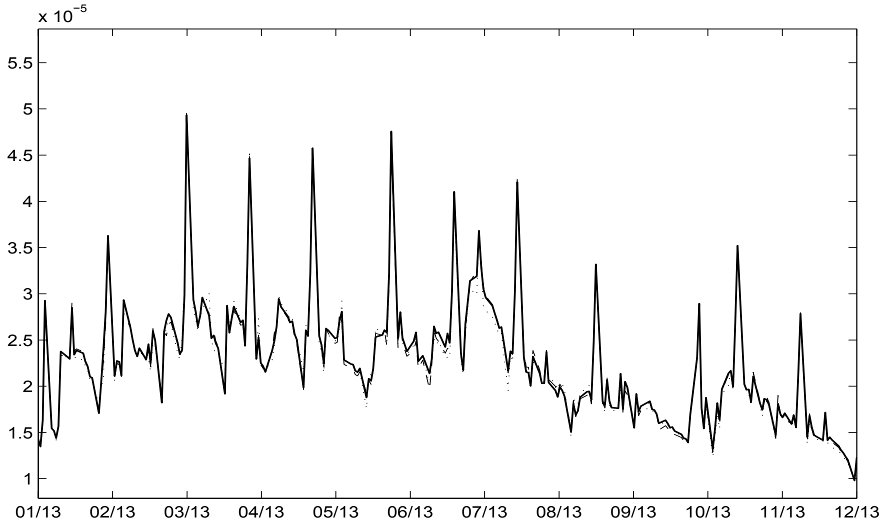

Figure 8: Out-of-sample EURUSD volatility forecasts of the Flexible Threshold-GARCH specification under different MCMC samples; solid: 100,000 sample, dashed: 10,000 sample; dotted: 5,000 sample. A burn-in period of 10,000 is used in all cases.

Table 7: Forecast errors of the Flexible Threshold-GARCH specification under different priors and MCMC samples; displayed are the out-of-sample errors based on EURUSD realized volatility

	MSE ($\times 10^{-10}$)	QLIKE
prior s.d. ×1, sample 100,000	1.9276	-9.5981
prior s.d. ×2, sample 100,000	1.9536	-9.5966
prior s.d. ×5, sample 100,000	1.9456	-9.5958
prior s.d. ×10, sample 100,000	1.9303	-9.5966
prior s.d. ×1, sample 5,000	1.9238	-9.5989
prior s.d. ×1, sample 10,000	1.9391	-9.5973

patterns, macroeconomic announcements, and longer run dependencies. *Journal of Finance*, *53*, 219–265.

Andersen, T. G., Bollerslev, T., & Diebold, F. X. (2007a). Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *The Review of Economics and Statistics*, *89*, 701–720.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Vega, C. (2007b). Real-time price discovery in global stock, bond and foreign exchange markets. *Journal of International Economics*, *73*, 251–277.

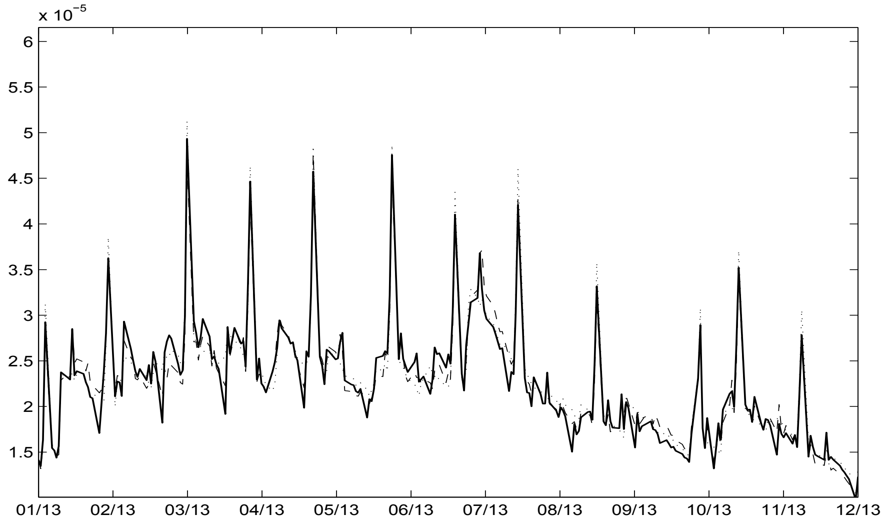

Figure 9: Out-of-sample EURUSD volatility forecasts of the Flexible Threshold-GARCH specification under different priors; solid: prior standard deviation, dashed: prior standard deviation $\times 2$; dotted: prior standard deviation $\times 10$.

Barbieri, M. M., & Berger, J. O. (2004). Optimal predictive model selection. *The Annals of Statistics*, *32*, 870–897.

Barndorff-Nielsen, O., & Shephard, N. (2001). Non-gaussian ornstein–uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society, Series B*, *63*, 167–241.

Barndorff-Nielsen, O. E., & Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B*, *64*, 253–280.

Barndorff-Nielsen, O. E., & Shephard, N. (2004). Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. *Econometrica*, *72*, 885–925.

Bartlett, M. S. (1957). Comment on d.v. lindley's statistical paradox. *Biometrika*, *44*, 533–534.

Bauwens, L., Omrane, W. B., & Giot, P. (2005). News announcements, market activity and volatility in the euro/dollar foreign exchange market. *Journal of International Money and Finance*, *24*, 1108–1125.

Bollerslev, T. (1986). Generalised autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*, 307–327.

Bollerslev, T., Cai, J., & Song, F. M. (2000). Intraday periodicity, long memory volatility, and macroeconomic announcement effects in the us treasury bond market. *Journal of Empirical Finance*, *7*, 37–55.

Bomfim, A. N. (2000). *Pre-Announcement Effects, News, and Volatility: Monetary Policy and the Stock Market*. FEDS Working Paper 00-50.

Brenner, M., Pasquariello, P., & Subrahmanyam, M. (2009). On the volatility and comovement of u.s. financial markets around macroeconomic news announcements. *Journal of Financial and Quantitative Analysis*, *44*, 1265–1289.

DeGennaro, R. P., & Shrieves, R. E. (1997). Public information releases, private information arrival and volatility in the foreign exchange market. *Journal of Empirical Finance*, *4*, 295–315.

Dellaportas, P., Denison, D. G. T., & Holmes, C. C. (2007). Flexible threshold models for modelling interest rate volatility. *Econometric Reviews*, *26*, 419–437.

Dellaportas, P., & Pourahmadi, M. (2012). Cholesky-garch models with applications to finance. *Statistics and Computing*, *22*, 849–855.

Ederington, L., & Lee, J. (1993). How markets process information: news releases and volatility. *Journal of Finance*, *48*, 1161–1191.

Ederington, L. H., & Lee, J. H. (1996). The creation and resolution of market uncertainty: The impact of information releases on implied volatility. *Journal of Financial and Quantitative Analysis*, *31*, 513–539.

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of united kingdom inflations. *Econometrica*, *50*, 987–1007.

Engle, R. F. (2002). Dynamic conditional correlation - a simple class of multivariate garch models. *Journal of Business and Economic Statistics*, *20*, 339–350.

Engle, R. F., & Rangel, J. G. (2008). The spline-garch model for low-frequency volatility and its global macroeconomic causes. *The Review of Financial Studies*, *21*, 1187–1222.

Evans, M. D. D., & Lyons, R. K. (2005). Do currency markets absorb news quickly? *Journal of International Money and Finance*, *24*, 197–217.

Evans, M. D. D., & Lyons, R. K. (2008). How is macro news transmitted to exchange rates? *Journal of Financial Economics*, *88*, 26–50.

Geweke, J., & Amisano, G. (2011). Optimal prediction pools. *Journal of Econometrics*, *164*, 130–141.

Geyer, C. J. (1991). Markov chain monte carlo maximum likelihood. In *Computing Science and Statistics: Proceedings of the 23rd Symposium of the Interface* (pp. 156–163). Interface Foundation of North America.

Gilks, W. R., Roberts, G. O., & George, E. I. (1994). Adaptive direction sampling. *The Statistician*, *43*, 179–189.

Green, P. J. (1995). Reversible jump mcmc computation and bayesian model determination. *Biometrika*, *82*, 711–732.

Hautsch, N., & Hess, D. (2002). The processing of non-anticipated information in financial markets: Analyzing the impact of surprises in the employment report. *European Finance Review*, *6*, 133–161.

Jasra, A., Stephens, D. A., & Holmes, C. C. (2007a). On population-based simulation for static inference. *Statistics and Computing*, *17*, 97–109.

Jasra, A., Stephens, D. A., & Holmes, C. C. (2007b). Population-based reversible jump markov chain monte carlo. *Biometrika*, *94*, 787–807.

Jensen, M., & Maheu, J. (2010). Bayesian semiparametric stochastic volatility modeling. *Journal of Econometrics*, *157*, 306–316.

Jensen, M., & Maheu, J. (2013). Bayesian semiparametric multivariate garch modeling. *Journal of Econometrics*, *173*, 3–17.

Jones, C. M., Lamont, O., & Lumsdaine, R. L. (1998). Macroeconomic news and bond market volatility. *Journal of Financial Economics*, *47*, 272–309.

Kim, S. J., McKenzie, M. D., & Faff, R. W. (2004). Macroeconomic news announcements and the role of expectations: evidence for us bond, stock and foreign exchange markets. *Journal of Multinational Financial Management*, *14*, 217–232.

Laakkonen, H. (2004). *The impact of macroeconomic news on exchange rate volatility*. Technical Report 24 Bank of Finland Discussion Papers.

Liang, F., & Wong, W. H. (2001). Real-parameter evolutionary monte carlo with applications to bayesian mixture models. *Journal of the American Statistical Association*, *96*, 653–666.

Liu, C., & Maheu, J. M. (2009). Forecasting realized volatility: A bayesian model-averaging approach. *Journal of Applied Econometrics*, *24*, 709–733.

Liu, J. S. (2001). *Monte Carlo Strategies in Scientific Computing*. New York: Springer.

Liu, J. S., Liang, F., & Wong, W. H. (2000). The multiple-try method and local optimisation in metropolis sampling. *Journal of the American Statistical Association*, *95*, 121–134.

Mitchell, M. L., & Mulherin, J. H. (1994). The impact of public information in the stock market. *Journal of Finance*, *49*, 923–950.

Nikkinen, J., Omran, M., Sahlström, P., & Äijö, J. (2006). Global stock market reactions to scheduled u.s. macroeconomic news announcements. *Global Finance Journal*, *17*, 92–104.

Nikkinen, J., & Sahlström, P. (2001). Impact of scheduled u.s. macroeconomic news on stock market uncertainty: A multinational perspective. *Multinational Finance Journal*, *5*, 129–149.

Patton, A. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, *160*, 246–256.

Petralias, A. (2010). *Bayesian Model Determination and Nonlinear Threshold Volatility Models*. Ph.D. thesis Department of Statistics, Athens University of Economics and Business.

Raftery, A. E., Madigan, D., & Hoeting, J. A. (1997). Bayesian model averaging for linear regression models. *Journal of the American Statistical Association*, *92*, 179–191.

Roberts, G., O., P., & Dellaportas, P. (2004). Bayesian inference for non-gaussian ornstein-uhlenbeck stochastic volatility processes. *Journal of the Royal Statistical Society, Series B*, *66*, 369–393.

Roberts, G., & Rosenthal, J. (2009). Examples of adaptive mcmc. *Journal of Computational and Graphical Statistics*, *9*, 349–367.

Appendix: The Population Reversible Jump algorithm

Reversible jump Algorithm

For parameters θ_m associated with model $m \in M$, the goal is to sample from the posterior distribution

$$f(\theta_m, m \mid \mathcal{D}) \propto f(\mathcal{D} \mid \theta_m, m) f(\theta_m \mid m) f(m),$$

where $\mathcal{D}$ represents the data based on T time points, $f(\mathcal{D} \mid \theta_m)$ is the likelihood of model m, and $f(\theta_m \mid m)$ is the prior density of the parameters θ_m conditional on model m. The likelihood based on model (2)-(5) is

$$f(\mathcal{D} \mid \theta_m, m) = (2\pi\sigma^2)^{-T/2} \left|V\right|^{-1/2} \exp\left(-\frac{\boldsymbol{\epsilon}' V^{-1} \boldsymbol{\epsilon}}{2\sigma^2}\right),$$

where $\boldsymbol{\epsilon} = (\epsilon_1, \dots \epsilon_T)'$ and V is a diagonal matrix with elements $G_t H_t$. The specification of the priors is given in Section 4.1, while the general structure of the algorithm can be found in Section 4.2.

Assume that at the current state (θ_m, m) of the Markov chain there exist K^m index variables. We select an addition, deletion or replacement move with probabilities q_m^+, q_m^- and q_m^0 respectively. We used equal probabilities taking care to change them accordingly when there are zero ($K^m = 0$) or the maximum number of variables ($K^m = K^{\max}$) in the current model m, with $K^{\max} = 15$ in the empirical application.

The *addition* move proceeds as follows. First we propose to move to model m' by adding an index variable not present in the current model m chosen with probability

$$j(m, m') = \frac{1}{(K^{\max} - K^m)} q_m^+.$$

The proposed parameters $\theta'_{m'} = (c_{ij}, g_{ij}, \rho_{ij}, \varsigma_i)$ are sampled from the proposal density $q(\theta'_{m'} \mid \theta_m, m, m')$ which is taken to be equal to the prior distribution $f(\theta'_{m'} \mid m')$ defined in Section 4.1. The *deletion* move proceeds in a similar way, where in the *replacement* move we select one variable present at the current model to be replaced with one not present with probability

$$j(m, m') = \frac{1}{K^m (K^{\max} - K^m)} q_m^0.$$

In all cases the proposed move to $(\theta'_{m'}, m')$ is accepted with probability

$$\alpha = min\left(1, \frac{f(\mathcal{D} \mid \theta'_{m'}, m') f(\theta'_{m'} \mid m') f(m') j(m', m)}{f(\mathcal{D} \mid \theta_m, m) f(\theta_m \mid m) f(m) q(\theta'_{m'} \mid \theta_m, m, m') j(m, m')}\right).$$

In the *split* move, for every i we propose to increase the number of threshold points c_{ij} by one with uniform probability

$$q(c'_{ij} \mid c_{ij}) = \frac{1}{J_i^{\max} - J_i^m},$$

where $J_i^{\max}$ the total number of possible threshold points, taken equal to the number of distinct observations of each variable $Z_{it_i^*}$. To achieve better mixing of the Markov chain, the parameters g_{ij} and ρ_{ij} are drawn from a density that dependents on the values of the current state of the chain. We draw u_1 and u_2 from the prior distributions of g_{ij} and ρ_{ij} respectively and set

$$\begin{aligned} g'_{ij} &= g_{i,j-1} + u_1 \\ g'_{i,j-1} &= g_{i,j-1} - u_1 \\ \rho'_{ij} &= \rho_{i,j-1} + u_2 \\ \rho'_{i,j-1} &= \rho_{i,j-1} - u_2. \end{aligned}$$

The move is then accepted with probability

$$\alpha = min\left(1, \frac{f(\mathcal{D} \mid \theta'_{m'}, m')f(\theta'_{m'} \mid m')f(m')q(c_{ij} \mid c'_{ij})q^M_{c'_{ij}}}{f(\mathcal{D} \mid \theta_m, m)f(\theta_m \mid m)f(m)q(c'_{ij} \mid c_{ij})q(u_1 \mid \theta_m, m)q(u_2 \mid \theta_m, m)q^S_{c_{ij}}}|J|^2\right),$$

where $q^S_{c_{ij}}$, $q^M_{c_{ij}}$ are the probabilities to perform a split or merge move, which are set equal (1/2) with appropriate changes when $J_i^m = 1$ or $J_i^m = J_i^{\max}$ and the Jacobian term for the g_{ij} parameters is

$$|J| = \left|\frac{\partial(g'_{i,j-1}, g'_{ij})}{\partial(g_{i,j-1}, u_1)}\right| = \begin{vmatrix} 1 & -1 \\ 1 & 1 \end{vmatrix} = 2.$$

Since the same Jacobian is obtained for the ρ_{ij} parameters, the term $|J|^2$ is included in the acceptance ratio. In the *merge* move we randomly select a threshold point to be deleted taking care so that the first threshold point ($j = 1$) is always included in c_{ij}. Then we set the new parameters as

$$\begin{aligned} g_{i,j-1} &= (g'_{ij} + g'_{i,j-1})/2 \\ \rho_{i,j-1} &= (\rho'_{ij} + \rho'_{i,j-1})/2. \end{aligned}$$

When performing random walk metropolis to update $\{\mu, \sigma, \alpha_1, \alpha_2, g_{ij}, \varsigma_i, \rho_{ij}\}$, we found useful to apply adaptive Metropolis methods, enforcing the diminishing adaptive condition which ensures asymptotic convergence and ergodicity and the bounded convergence condition so that the convergence time of the kernel is bounded in probability, as described in Roberts & Rosenthal (2009). To this respect the scale of the proposal for each parameter (denoted as τ_i for a parameter i) is adaptively updated as follows: Given a starting value for the proposal scale (acquired through a pilot run), set a global maximum (and minimum) bound, $M = max(\tau_i)$ (in our case we set $M = 10e{+}5$ for all parameters), so that $-\log(M) \leq \log(\tau_i) \leq \log(M)$. Then every b batches of 50 swaps add (subtract) to the logarithm of the proposal scale a quantity, $\log(\tau_i) + \delta(b)$, where $\delta(b) = min(b_0, b^{-1/2})$ and $b_0 = 0.1$, if the acceptance rate is higher (lower) than 0.234. The necessary condition $H_t > 0$ resulted to less than 5% rejection of the proposed values of g_{ij}, ς_i and ρ_{ij}.

Population Algorithm

In the *exchange* move we randomly select two adjacent, in terms of temperature, chains, and propose to swap their values. Denote with x_ℓ the state $(m_\ell, \theta^\ell_{m_\ell})$ of chain ℓ, with target distribution $\pi_\ell(x_\ell)$ and $\pi_\ell \propto \pi^{\zeta_\ell}$ (see Section 4.2). The Metropolis-Hastings acceptance ratio used to swap the states of chains 1 and 2 is of the form

$$\alpha = min\left(1, \frac{\pi_1(x_2)\pi_2(x_1)}{\pi_1(x_1)\pi_2(x_2)}\right).$$

We choose not to temper the prior distributions, thus they cancel out in the acceptance ratio.

The *crossover* move takes a fraction of the variables, along with their associated parameters present in the current chain and places them in another randomly selected chain. This move proceeds as follows. Select two random chains (not necessarily adjacent), say 1 and 2, with probability τ_1 and $\tau_{2|1}$ respectively. Then

if both chains have the same variables reject the move, else:

- draw a discrete uniform random variable $u \sim DU(1, \ldots, v)$, where v is the number of variables in chain 1 not included in chain 2.
- delete randomly u variables with their associated parameters from chain 1.
- add the u variables with their associated parameters to chain 2.
- accept the new state with probability

$$\alpha = min\left(1, \frac{\pi_1(x_1')\pi_2(x_2')q(x_1, x_2 \mid x_1', x_2')}{\pi_1(x_1)\pi_2(x_2)q(x_1', x_2' \mid x_1, x_2)}\right),$$

where $q(x_1', x_2' \mid x_1, x_2) = q(u \mid v)\tau_1\tau_{2|1}$, is the proposal density of the new states, $q(u \mid v)$ the uniform probability to select u from v available variables i, τ_1 is the probability to select chain 1 and $\tau_{2|1}$ is the probability to select chain 2 given we have selected chain 1.

The term $\tau_1 \tau_{2|1}$ cancels out in the acceptance ratio since these probabilities are set to be uniform.

The algorithm is coded in Matlab and is available from the first author upon request.