Conditional Heteroskedasticity in Qualitative Response Models of Time Series:A Gibbs Sampling Approach to the Bank Prime Rate.

Authors	Michael J. Dueker
Working Paper Number	1998-011A
Citable Link	https://doi.org/10.20955/wp.1998.011
Suggested Citation	Dueker, M.J.; Conditional Heteroskedasticity in Qualitative Response Models of Time Series:A Gibbs Sampling Approach to the Bank Prime Rate., Federal Reserve Bank of St. Louis Working Paper 1998-011. URL https://doi.org/10.20955/wp.1998.011

Published In	Journal of Business & Economic Statistics
Publisher Link	https://doi.org/10.1080/07350015.1999.10524834

CONDITIONAL HETEROSCEDASTICITY IN QUALITATIVE RESPONSE MODELS OF TIME SERIES: A GIBBS SAMPLING APPROACH TO THE BANK PRIME RATE

August 1998

Abstract

Previous time series applications of qualititative response models have ignored features of the data, such as conditional heteroscedasticity, that are routinely addressed in time-series econometrics of financial data. This article addresses this issue by adding Markov-switching heteroscedasticity to a dynamic ordered probit model of discrete changes in the bank prime lending rate and estimating via the Gibbs sampler. The dynamic ordered probit model of Eichengreen, Watson and Grossman (1985) allows for serial autocorrelation in probit analysis of a time series, and the present article demonstrates the relative simplicity of estimating a dynamic ordered probit using the Gibbs sampler instead of the Eichengreen et al. maximum-likelihood procedure. In addition, the extension to regime-switching parameters and conditional heteroscedasticity is easy to implement under Gibbs sampling. The article compares tests of goodness of fit between dynamic ordered probit models of the prime rate that have constant variance and conditional heteroscedasticity.

JEL Classification: C25, C22, G21

Keywords: Markov chain, data augmentation, discrete choice

Michael Dueker
Research Officer
Federal Reserve Bank of St. Louis
411 Locust Street
St. Louis, MO 63102
mdueker@stls.frb.org

Introduction

Early applications of qualitative response models, e.g., logits and probits, dealt exclusively with cross-sectional data [Goldberger (1964)]. In discrete-choice models of a cross section of individuals, a maintained assumption is that each individual's random utility shock is an independent draw from the population distribution. Qualitative-response models have also become popular for time series such as business recessions, financial crises and interest-rate changes [Estrella and Mishkin (1997); Frankel and Rose (1996); Bernard and Gerlach (1996); Eichengreen, Rose and Wyplosz (1995); Davutyan and Parke (1995)]. Yet insufficient attention has been given in this literature to the dependent nature of time series, in both first and second moments. Serial dependence in the mean has been addressed in the dynamic probit model of Eichengreen, Watson and Grossman (1985) and in the autoregressive conditional hazard model of Hamilton and Jorda (1997). With respect to the variance of a time series, however, there is often substantial evidence against the assumption that each disturbance term is drawn independently from a population with a constant variance. Conditional heteroscedasticity is an especially prevelant feature in financial data among interest rates, stock price tick changes and other price movements [Bollerslev, Chou, Kroner (1992); Hausman, Lo and McKinlay (1992)].

The chief obstacle to applying time-series methods to address conditional heteroscedasticity has been the latent nature of the residual shocks or perturbations in discrete-choice models. Most methods, such as autoregressive conditional heteroscedasticity [ARCH, Engle (1982)], model the conditional variance as a function of lagged squared residuals, which is not feasible if the residuals are unobservable. Broseta (1993) modifies the ARCH approach to fit probit-type models by substituting the squared value of the expected value of the residual for the squared residual. This expected value is not very informative, however, because it does not vary much in a discrete-choice model. Moreover, the square of the expected value of the residual may be a poor indicator of the expected value of the squared residual. In this article, I use data-augmentation methods to draw values of the latent variable, whereupon its conditional heteroscedasticity can be addressed with familiar methods, such as regime switching. Furthermore, the data-augmentation approach espoused in this article simplifies the estimation of the constant-variance dynamic ordered probit that serves as the base model in the analysis of weekly prime rate changes.

The bank prime lending rate of interest is a discrete variable that always changes by multiples of 25 basis points. Other examples of discretely-changing or "administered" interest rates include the Federal Reserve discount rate, the Federal Reserve's target federal funds rate [Hamilton and Jorda (1997)] and home mortgage rates quoted by individual lenders. The published weekly national mortgage rates represent survey averages, but any potential borrower following a handful of lenders would need to predict discrete changes. Changes in such interest rates naturally fit into the ordered-probit framework described in Maddala (1983), because the observations fall into a small number of categories that can be ordered from lowest to highest.

The next section briefly discusses the advantages of estimating a dynamic ordered probit model via Gibbs sampling, relative to the maximum-likelihood procedure of Eichengreen et al. (1985). The third section reviews the regime-switching approach to time-varying parameters popularized by Hamilton (1989). The fourth section presents the application of Gibbs-sampling methods to estimate a dynamic ordered probit model with regime-switching conditional heteroscedasticity. The fifth section presents estimation and forecast results for the bank prime rate in the form of posterior means from the Gibbs sampler.

II. The estimation of a dynamic ordered probit with and without data augmentation

In the dynamic ordered probit of Eichengreen, Watson and Grossman (1985), an observed variable, Y, changes each period by one of J different discrete amounts, including changes of zero. A latent 'desired' level, Y^*, is defined in terms of its own changes from period to period plus an initial level, Y_0^*, where the changes in the desired level are assumed to depend on a vector of lagged explanatory variables plus a disturbance as in an ordinary regression model:

$$\Delta Y_t^* = \Delta X_{t-1}'\beta + \epsilon_t \tag{1}$$

The desired change in Y at time t (as opposed to the change in the desired level) is denoted

Z_t and equals $Y_t^* - Y_{t-1}$. Equivalently,

$$Z_t = \Delta Y_t^* + Y_{t-1}^* - Y_{t-1} \tag{2}$$

In this way, the model allows for pressure for a change in Y to depend on past gaps between desired and actual levels, thereby accounting for serial dependence in the changes in Y. A vector of cut-off constants $(c_0, .., c_J)$ determines that the actual change, ΔY_t, is in category j if and only if

$$Z_t \in (c_{j-1}, c_j).$$

The maximum-likelihood estimation procedure of Eichengreen, Watson and Grossman (1985) requires numerical evaluation of an integral for each observation in order to obtain the density, h, of Y_t^* inside the relevant interval, where ϕ is the standard normal density and I_t is the information available up to time t:

$$h(Y_t^* \mid I_t) = 1/\sigma_\epsilon \int_{l_{t-1}}^{U_{t-1}} \phi(\Delta Y_t^*/\sigma_\epsilon) h(Y_{t-1}^* \mid I_t) dY_{t-1}^*, \tag{3}$$

where $\Delta Y_{t-1} \in$ cat.i and $l_{t-1} = Y_{t-2} + c_{i-1}; U_{t-1} = Y_{t-2} + c_i$ Because numerical evaluation of these integrals is time-consuming and approximate, it is not tractable under direct maximum-likelihood estimation to extend the model to include additional features, such as regime-switching parameters.

In cases like the dynamic ordered probit, where the joint density of Y_t^* and Y_{t-1}^* is difficult to evaluate, data augmentation via Gibbs sampling offers a tractable method to generate a sample of draws from a joint distribution through a sequence of draws from the respective conditional distributions. Data augmentation in the present context allows one to treat augmented values of $Y_s^*, s \neq t$, as observed data when evaluating the conditional density of Y_t^*. Thus, one conditions the density of Y_t^* on a *value*, instead of a *density*, of Y_{t-1}^*, making the problem much simpler than recursive evaluation of the integral in equation (3). Additional details on how the latent variables are sampled are in the fourth section and the appendix. Because the fourth section presents the full model with regime-switching conditional heteroscedasticity, we briefly review regime-switching models in the next section.

III. Regime-switching models

Hamilton's (1989, 1990) use of Markov-switching parameters has spawned many applications, because it offers a way to capture regime changes in economic data. In forecasting, Hamilton (1994) notes that an estimated regime-switching model permits predictions of a variable to take account of possible non-deterministic future changes in regime. Inferences regarding the dates of past changes in regime, the magnitude of the difference between regimes and forecasts of the regime at future dates are not readily apparent from the data. For this reason, Hamilton (1989,1990) assumes that the regimes or states are governed by an unobservable discrete state variable that follows a first-order Markov process. In many applications, the state variable, S_t, is assumed to be binary: $S_t \in \{0, 1\}$. For a first-order Markov process with constant transition probabilities, the serial dependence is easily summarized by the sum of the transition probabilities:

$$\begin{aligned}\text{Prob}(S_t = 0 \mid S_{t-1} = 0) &= p \\ \text{Prob}(S_t = 1 \mid S_{t-1} = 1) &= q\end{aligned}$$

If $p+q > 1$, the process has positive serial correlation; if $p+q < 1$, the process has negative serial correlation and if $p + q = 1$, there is no serial dependence. The unconditional value of $\text{Prob}(S_t = 0)$ is $(1 - q)/(2 - p - q)$.

IV. The Model and Gibbs Sampling Procedure

The dynamic ordered probit has seven categories, corresponding with the categories found in Table 1 for 1302 weekly prime rate observations (with the prime rate denoted as PR) and 223 non-zero changes between December 15, 1972 and December 5, 1997.

As discussed in the second section, the dynamic ordered probit model assumes that the change in the desired level of the prime rate depends on the changes in a vector of explantory variables, X, plus a disturbance:

$$\Delta PR_t^* = \Delta X_{t-1}'\beta + \epsilon_t. \tag{4}$$

The asterisk denotes the desired, as opposed to actual, level. The shock ϵ_t is normally distributed with variance σ_t^2, where the variance is not constant across time; it will change between two levels:

$$\mathrm{var}(\epsilon_t) = \sigma^2_{S_t} = \sigma^2_0(1 - S_t) + \sigma^2_1 S_t \tag{5}$$

In estimating β from equation (4), the observations from the high-volatility regime have less influence on the determination of the estimated values of the β coefficients through weighted least squares. The state variable $S_t \in \{0, 1\}$ is a binary random variable that follows a first-order Markov process. The transition probabilities for this process are:

$$\begin{aligned}\mathrm{Prob}(S_t = 0 \mid S_{t-1} = 0) &= p \\ \mathrm{Prob}(S_t = 1 \mid S_{t-1} = 1) &= q\end{aligned}$$

As in the standard dynamic ordered probit model of Eichengreen and Watson (1985), the actual change in the discrete choice variable is a function of a latent variable, Z, that is the sum of this period's change in the desired level and last period's gap between the actual level and the desired level:

$$Z_t = \Delta PR^*_t + PR^*_{t-1} - PR_{t-1} \tag{6}$$

The actual changes are assumed to be related to Z_t in the following way:

$$\begin{aligned}\Delta PR < -.50 &\leftrightarrow Z_t < -.750 \\ \Delta PR = -.50 &\leftrightarrow -.750 \leq Z_t < -.375 \\ \Delta PR = -.25 &\leftrightarrow -.375 \leq Z_t \leq -.125 \\ \Delta PR = 0.00 &\leftrightarrow -.125 < Z_t < +.125\end{aligned}$$

$$\begin{aligned} \Delta PR = +.25 &\leftrightarrow +.125 \leq Z_t \leq +.375 \\ \Delta PR = +.50 &\leftrightarrow +.375 < Z_t \leq +.750 \\ \Delta PR > +.50 &\leftrightarrow Z_t > +.750 \end{aligned} \tag{7}$$

The choice of the cut-off constants, $c = (-.75, -.375, -.125, .125, .375, .75)$, assumes that the actual discrete changes in the prime rate correspond to the 25-basis point increment closest to the 'desired' continuous change.

Gibbs sampling

The model is estimated via the Gibbs sampler in order to take advantage of the data augmentation that generates samples of the state variables and the latent 'desired' levels of the prime rate. These variables are included in a chain of parameters to be simulated from a full set of conditional distributions. The parameter groupings for the Gibbs chain are:

$$\begin{aligned} \varrho_1 &= \{PR_t^*\}, t = 1, ..., T \qquad \text{latent 'desired' level} \\ \varrho_2 &= (\sigma_0^2, \sigma_1^2) \\ \varrho_3 &= \{S_t\}, t = 1, .., T \qquad \text{States} \\ \varrho_4 &= \beta \qquad \text{regression coefficients} \\ \varrho_5 &= (p, q) \qquad \text{transition probs.} \end{aligned}$$

Gibbs sampling is an attractive approach because it is relatively easy to sample from these conditional distributions, as outlined in Albert and Chib (1993) for the binary and ordered probit. Gibbs sampling consists of iterating through cycles of draws of parameter values from conditional distributions as follows:

$$\begin{array}{ll}
f(\varrho_1^{(i+1)} & | \quad \varrho_2^{(i)}, \varrho_3^{(i)}, \varrho_4^{(i)}, \varrho_5^{(i)}, Y_T) \\
f(\varrho_2^{(i+1)} & | \quad \varrho_1^{(i+1)}, \varrho_3^{(i)}, \varrho_4^{(i)}, \varrho_5^{(i)}, Y_T) \\
f(\varrho_3^{(i+1)} & | \quad \varrho_1^{(i+1)}, \varrho_2^{(i+1)}, \varrho_4^{(i)}, \varrho_5^{(i)}, Y_Y) \\
f(\varrho_4^{(i+1)} & | \quad \varrho_1^{(i+1)}, \varrho_2^{(i+1)}, \varrho_3^{(i+1)}, \varrho_5^{(i)}, Y_T) \\
f(\varrho_5^{(i+1)} & | \quad \varrho_1^{(i+1)}, \varrho_2^{(i+1)}, \varrho_3^{(i+1)}, \varrho_4^{(i+1)}, Y_T).
\end{array} \tag{8}$$

where Y_T stands for the entire history of the data and superscript i indicates run number i through the Gibbs sampler. At each step, a value of ϱ is drawn from its conditional distribution. As discussed in the appendix, all of the necessary conditional distributions can be standard statistical distributions, given appropriate choices for prior distributions. The key idea behind Gibbs sampling is that after a sufficient number of iterations, the draws from the respective conditional distributions jointly represent a draw from the joint posterior distribution, which often cannot be evaluated directly [Gelfand and Smith (1990)].

Prior and posterior conditional distributions for $\varrho_j, j = 1,..,5$ are in the appendix. The Gibbs sampler was run for a total of 8000 iterations in each estimation. The first 3000 iterations were discarded to make sure the sampler had converged to the posterior distribution.

Explanatory variables

The conditioning variables, X_{t-1} from equation (4), are lagged changes in the federal funds rate and lagged values of the change in the spread between the three-month commercial paper rate and the three-month Treasury bill rate. Changes in the federal funds rate indicate general swings in short-term interest rates, and the prime rate has generally keyed off the federal funds target rate for the past several years. The paper-bill spread, on the other hand, serves as a business-cycle indicator, as suggested by Friedman and Kuttner (1993), who find that the paper-bill spread rises above normal when the economy is poised to enter a recessionary phase. These increases in the paper-bill spread reflect a quality spread due to the possibility that some issuers of commercial paper will not survive a recession without defaulting. For this reason, the paper-bill spread is a predictor of recessions, but not of the strength of ongoing expansions, so we only consider changes in the paper-bill spread when the spread is above its median value by multiplying changes in the spread by the appropriate dummy variable.

We expect all of the β coefficients from equation (4) to have positive coefficients. An increase in the federal funds rate signifies an upswing in short-term interest rates that the prime rate would be expected to follow. An increase in the paper-bill spread signals an increase in quality spreads that would imply an increase in bank customer lending rates for any given interbank lending rate, the federal funds rate. Because we are primarily interested in the overall, multi-period response of the prime rate to changes in the explanatory variables, coefficients representing sums of lag coefficients are presented. Thus, if the original model were written with Γs denoting coefficients on individual lags of the j^{th} regressor variable:

$$\Gamma_{j1}X_{j,t-1} + \Gamma_{j2}X_{j,t-2} + ... + \Gamma_{jk}X_{j,t-k},$$

then the reported β_j coefficients are defined as $\beta_{j1} = (\Gamma_{j1} + ... + \Gamma_{jk}), \beta_{j2} = (\Gamma_{j2} + ... + \Gamma_{jk}), ..., \beta_{jk} = \Gamma_{jk}$. The lag lengths, k, were chosen informally, based on when the β coefficients lost significance, which is at lag four for the federal funds rate changes and lag three for the changes in the paper-bill spread. The explanatory variables are all lagged at least one period, so they are pre-determined relative to this week's change in the dependent variable.

V. Estimation results

The posterior means for the regression coefficients, found in Table 2, have the expected positive signs and are significant, relative to their empirical confidence intervals from the Gibbs sampling. The probability values at the bottom of Table 1 are the posterior means of the p-values of Wald test statistics for the joint significance of the regression coefficients taken at each iteration of the Gibbs sampler. The significantly positive coefficients on the changes in the paper-bill spread suggest that the prime rate displays a countercyclical mark-up vis-a-vis the federal funds rate, that is, the quality spread increases in recessions when customer default rates rise.

In the model with Markov switching, the posterior mean of the standard deviation of the shocks is more than five times greater in the high-variance state than in the low-variance state. The unconditional probability of being in the low-volatility state is 83.3 percent $[(1-q)/(2-p-q)]$. Figure 1 shows the Gibbs posterior means for the volatility state variable across the sample period. As one would expect, the high-volatility state is concentrated between October 1979 and October 1982 when the Federal Reserve targeted non-borrowed reserves and induced greater volatility in short-term interest rates. Other occurrences of the high-volatility state outside of the 1979-82 period more closely resemble spikes than regimes. Accordingly, the transition probability q from Table 1 suggests that the half-life of the high-volatility state is about two weeks.

A tell-tale sign of conditional heteroscedasticity in a time series is fat tails or leptokurtosis. Conditional heteroscedasticity is related to leptokurtosis because the latter serves as a measure of the "variance of the variance" when a mixture of non-leptokurtic conditional distributions is causing fat tails in the unconditional distribution. Table 1 shows that the interest-rate changes appear too leptokurtic for the constant variance model, because the sample kurtosis of its residuals has a posterior mean of 6.59, which is well above 3, the kurtosis of level of normal random variables. This measure is a Gibbs posterior mean because the sample kurtosis was calculated at each Gibbs draw of the vector of PR^*. The Markov-switching model posits that a mixture of two normal random variables with different variances can generate the leptokurtosis observed in the interest-rate changes. Table 1 shows that the Gibbs posterior mean of the sample kurtosis of the standardized residuals from the switching model, ϵ_t/σ_{S_t}, is only 3.58, which is only slightly above 3. Thus, it appears that the simple Markov mixture model is adequate for explaining the conditional heteroscedasticity observed in the prime-rate changes. A formal test of goodness of fit is presented next.

In comparing the model with Markov-switching variance to the constant-variance model, I use a test of goodness of fit that checks how closely the model-impled cumulative density function describes the actual distribution of the residuals. Following Vlaar and Palm (1993), I divide the model-implied cumulative density functions into 100 percentile groups by cutting the vertical axis of the CDF into 100 pieces from 0.01 to 1.0. If the model fits the data well, the expected number of residuals in each percentile group is one percent of the sample (or about 13 observations). The actual count of residuals in percentile group i of the model-implied CDF, $\Phi(.)$, is

$$n_i = \sum_{t=1}^{T} I_{it} \qquad \text{where} \qquad \begin{aligned} I_{it} &= 1, \qquad \text{if} \quad \frac{(i-1)}{100} < \Phi(\epsilon_t/\sigma_{S_t}) \leq \frac{i}{100} \\ &= 0, \qquad \text{otherwise} \end{aligned}$$

The residual counts are tabulated at each iteration of the Gibbs sampler and the average count, $\overline{n_i}$, is taken across all the iterations. Figure 2 plots the actual $\overline{n_i}$ counts across all 100 percentile groups for the dynamic ordered probit models with and without Markov switching in the variance. Figure 2 shows that the residuals from the constant-variance model are too leptokurtic, relative to the model-implied distribution: the numbers of residuals near the center and the tails are too high relative to the model-implied expected value of 13, and the number in the shoulders of the distribution is too low. The residuals from the Markov-switching model, in contrast, are much more uniformly distributed across the percentile groups, without major deviations from the expected value of 13. A chi-square test of goodness of fit formalizes the visual evidence from Figure 2 in favor of the conditionally heteroscedastic model. The test statistic equals $100/T \sum_{i=1}^{100}(n_i - T/100)^2$ and is distributed χ^2_{99} under the null. This statistic, like the count n_i, is calculated at each iteration of the Gibbs sampler, and its probability value is averaged across the Gibbs iterations to arrive at a posterior mean for the p-value of the test. For the constant-variance model, the posterior mean p-value for the test of goodness of fit is less than 1e-05, whereas it equals 0.373 for the conditionally heteroscedastic model with Markov-switching variance. Thus, the differences between the two models in matching their CDFs to the data in Figure 2 lead to significantly different results in the test of goodness of fit.

Conclusions

This article presents methods for greatly simplifying estimation of the dynamic ordered probit model of Eichengreen et al. (1985) via the Gibbs sampler and its data augmentation. I also extend the model to include treatment of conditional heteroscedasticity through regime switching. These methods are applicable to a wide variety of qualitative-response models of time series. In general, with the methods presented here, one can add almost any time-series feature to the latent variable governing a qualitative-response process.

In the changes in the bank prime lending rate studied here, conditional heteroscedasticity is shown to be an important feature of this time series, and the heteroscedasticity is addressed through Markov switching in the variance parameters. The high-volatility state holds about 17% of the time on average, wherein the standard deviation of the shocks is more than five times as large as in the low-variance state. The Gibbs posterior means for the state variable give the expected result that the most important high-variance episode coincided with the 1979-82 period when the Federal Reserve experimented with non-borrowed reserves targeting and induced greater interest-rate volatility. The treatment of conditional heteroscedasticity makes the difference between a dynamic ordered probit model that can pass a test of goodness of fit and a constant-variance specification that badly fails the test.

Appendix: Gibbs sampling distributions

Priors and posteriors for transition probabilities

The likelihood function for a discrete binary random variable that is governed by a first-order Markov process is

$$L(p,q) = p^{n_{00}}(1-p)^{n_{01}}q^{n_{11}}(1-q)^{n_{10}} \tag{9}$$

where n_{ij} is the number of transitions between $S_{t-1} = i$ and $S_t = j$.

The prior is to assign parameters u_{ij}, where the ratio between u_{00} and u_{01}, for example, represents a prior guess for the ratio between the corresponding numbers of actual transitions, n_{00}/n_{01}. The magnitudes of the u_{ij} relative to the sample size indicate the strength of the prior. As a weak prior, I set $u_{00} = 4, u_{01} = 1, u_{10} = 1$, and $u_{11} = 4$, such that the sum of the u_{ij} is low relative to the sample size.

The beta distribution is conjugate to itself, so the posterior is also beta and is the product of the prior and the likelihood of the observed transitions, so that we may draw transition probabilities from

$$\begin{aligned} p \mid \tilde{S}_T &\sim \text{beta}(u_{00}+n_{00}, u_{01}+n_{01}) & (10)\\ q \mid \tilde{S}_T &\sim \text{beta}(u_{11}+n_{11}, u_{10}+n_{10}), & (11)\end{aligned}$$

where $\tilde{S}_T = \{S_t\}, t = 1, ..., T$. The initial values for p and q at the start of the Gibbs sampling were p=0.95 and q= 0.9.

Priors and posteriors for Markov state variables

We wish to sample the states in reverse order from the following probability, where Y_T stands for the entire history of the observed and latent data and y_t is the observed and latent data at a point in time:

$$P(S_t = 0 \mid S_{t+1}, ..., S_T, Y_T) \tag{12}$$

By Bayes theorem, and as outlined in Chib (1996),

$$\begin{aligned} P(S_t = 0 \mid S_{t+1},...,S_T,Y_T) &\propto f(y_{t+1},...,y_T,S_{t+1},...,S_T \mid y_1,...,y_t,S_t) \times P(S_t \mid y_1,...,y_t)\\ &\propto f(y_{t+1},...,y_T,S_{t+2},...,S_T \mid y_1,...,y_t,S_t,S_{t+1}) \times\\ &\quad P(S_{t+1} \mid S_t) \times P(S_t \mid y_1,...,y_t)\\ &\propto P(S_{t+1} \mid S_t) \times P(S_t \mid y_1,...,y_t) & (13)\end{aligned}$$

The first and second proportions in (13) are simply applications of Bayes' theorem. Because the density $f(y_{t+1}, ..., y_T, S_{t+2}, ..., S_T \mid y_1, ..., y_t, S_t, S_{t+1})$ is independent of S_t, it can be subsumed into the constant of proportionality, which can easily be recovered in order to draw states. As shown in (13), the only necessary inputs are the transition probabilities and the filtered probabilities conditional on the contemporaneous data.

Priors and posteriors for variances

Following Kim and Nelson (1998), the prior distribution for the variances is chosen to be inverted gamma, because the inverted gamma is conjugate to itself: $\sigma_i^2 \sim IG(\nu_i, \tau_i)$ with $\nu_0 = 2$, $\nu_1 = .4$, $\tau_0 = 1$ and $\tau_1 = 1$. The ν_i are shape parameters and the prior with $\nu_1 < \nu_0$ implies a more diffuse distribution with thicker tails for σ_1^2 because one of the variance states will occur less often and therefore the variance parameter for that state will be estimated less precisely. A higher value of ν_i implies a stronger prior for σ_i^2. The posterior is also inverted gamma, $\sigma_i^2 \sim IG(\frac{\nu_i+T}{2}, \frac{\tau_i}{2} + \frac{1}{2}\sum_{t|S_t=i} \epsilon_t^2)$. Initial values for σ_0^2 and σ_1^2 to start the Gibbs sampling were set at 0.025 and 0.20, respectively.

Priors and posteriors for β coefficients

The prior for β is diffuse and the initial value for β in the first cycle of the Gibbs sampler is the ordinary least square estimate from the regression on the actual (as opposed to the desired) prime rate changes. With Σ_T denoting the diagonal matrix with entries from the vector $(\sigma_{S_t}^2, t = 1, ..., T)$, the posterior distribution for β is the multivariate normal distribution for generalized least squares coefficients:

$$\beta \sim N(X'\Sigma_T^{-1}X)^{-1}X'\Sigma_T^{-1}\Delta PR^*, X'\Sigma_T^{-1}X)^{-1}).$$

Generating latent variables, PR_t^*

The initial values of $PR_t^*, t = 1, .., T$ are drawn from $f(PR_t^* \mid PR_{t-1}^*, \Delta PR_t \in \text{cat.}j)$, where PR_0^* is held fixed at PR_0. In this case,

$$PR_t^* \sim N(PR_{t-1}^* + \Delta X_{t-1}', \sigma_{S_t}^2)$$

with truncation such that $PR_t^* \in (PR_{t-1} + c_{j-1}, PR_{t-1} + c_j)$. From equation (7), the constants $c_0, .., c_7$ are $(-\infty, -.75, -.375, -.125, .125, .375, .75, \infty)$.

We take subsequent draws from

$$f(PR_t^{*(i+1)} \mid PR_{t-1}^{*\,(i+1)}, PR_{t+1}^{*\,(i)}, \Delta PR_t \in \text{cat.}j), \tag{14}$$

where, as in equation (8), superscript i denotes the i^{th} cycle of the Gibbs sampler. We use the density from equation (14), because sampling the entire vector jointly from $f(PR_1^*, ..., PR_T^* \mid Y_T)$ would require evaluation of a density equivalent to the cumbersome likelihood function from equation (3). To draw from (14), we note that unconditionally $(\epsilon_t, \epsilon_{t+1})$ are distributed as independent, bivariate normals with mean zero:

$$f(\epsilon_t, \epsilon_{t+1}) = \frac{1}{2\pi\sigma_{S_t}\sigma_{S_{t+1}}} \exp\left\{-.5\epsilon_t^2/\sigma_{S_t}^2 - .5\epsilon_{t+1}^2/\sigma_{S_{t+1}}^2\right\} \tag{15}$$

Given equation (4), we can relate ϵ_t and ϵ_{t+1} to PR_t^* as follows:

$$\begin{aligned} \epsilon_t &= PR_t^* - A \\ & \quad A = PR_{t-1}^* + \Delta X_{t-1}'\beta \\ \epsilon_{t+1} &= -PR_t^* + B \\ & \quad B = PR_{t+1}^* - \Delta X_t'\beta \end{aligned} \tag{16}$$

In other words, PR_t^* is the only unknown affecting the values of both $\epsilon_t, \epsilon_{t+1}$, and the conditional density of PR_t^* must take into account the extent to which a particular value of PR_t^* would create an outlier value in either or both of $(\epsilon_t, \epsilon_{t+1})$. Substituting (16) into (15), we end up (after some algebra) with a univariate density for PR_t^* conditional on PR_{t-1}^*, PR_{t+1}^* such that

$$PR_t^* \sim N\left(\frac{\sigma_{S_{t+1}}^2 A + \sigma_{S_t}^2 B}{\sigma_{S_{t+1}}^2 + \sigma_{S_t}^2}, \frac{\sigma_{S_{t+1}}^2 \sigma_{S_t}^2}{\sigma_{S_{t+1}}^2 + \sigma_{S_t}^2}\right). \tag{17}$$

Conditional on $\Delta PR_t \in \text{cat.}j$, the distribution of PR^* is truncated so that $PR_t^* \in (PR_{t-1} + c_{j-1}, PR_{t-1} + c_j)$. From (17), it is interesting to note how the variances affect the mean of PR_t^*. If the variances were the same across both time periods, the mean of PR_t^* is simply $(A + B)/2$, the average of the two values implied by knowing PR_{t-1}^* and PR_{t+1}^*, respectively. If, on the other hand, the variances are unequal, such that $\sigma_{S_t}^2 > \sigma_{S_{t+1}}^2$, then A receives less weight than B in determining the mean of PR_t^*. Similarly, the variance of PR_t^* is $\sigma_{S_t}^2/2$, the variance of the average, if $S_t = S_{t+1}$.

References

Albert, J.H. and Chib, S. (1993), "Bayesian Analysis of Binary and Polychotomous Response Data," *Journal of the American Statistical Association* 88, pp. 669-79.

Bernard, H. and Gerlach. S. (1996), "Does the Term Structure Predict Recessions? The International Evidence," working paper, Bank for International Settlements.

Bollerslev, T., Chou, R.Y. and Kroner, K.F. (1992), "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence," *Journal of Econometrics* 52, pp. 5-59.

Broseta, B. (1993), "Estimation of a Game-Theoretic Model of Learning: An Autoregressive Conditional Heteroscedasticity Approach," UC-San Diego Economics Discussion Paper 93-35.

Chib, S. (1996), "Calculating Posterior Distributions and Modal Estimates in Markov Mixture Models," *Journal of Econometrics* 75, pp. 79-97.

Davutyan, N. and Parke, W.R. (1995), "The Operations of the Bank of England 1890-1908: A Dynamic Ordered Probit Approach," *Journal of Money, Credit, and Banking* 27, pp. 1099-1112.

Eichengreen, B., Rose, A.K. and Wyplosz, C. (1995), "Exchange Rate Mayhem. The Antecedents and Aftermath of Speculative Attack," *Economic Policy* 21, pp. 249-312.

Eichengreen, B., Watson, M.W. and Grossman, R.S. (1985), "Bank Rate Policy Under the Interwar Gold Standard," *Economic Journal* 95, pp. 725-45.

Engle, R. F. (1982), "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica* 50, pp. 987-1007.

Estrella, A. and Mishkin, F.S. (1997), "The Predictive Power of the Term Structure of Interest Rates in Europe and the United States: Implications for the European Central Bank," *European Economic Review* 41, pp. 1375-1401.

Frankel, J.A. and Rose, A.K. (1996), "Currency Crashes in Emerging Markets. An Empirical Treatment," *Journal of International Economics* 41, pp. 351-66.

Friedman, B. and Kuttner, K. "Why Does the Paper-Bill Spread Predict Real Economic Activity?" in James H. Stock and Mark W. Watson (eds.) **Business Cycles, Indicators, and Forecasting.** University of Chicago Press, 1993.

Gelfand, A.E. and Smith, A.F.M. (1990), "Sampling-Based Approaches to Calculating Marginal Densities," *Journal of the American Statistical Association* 85, pp. 398-409.

Goldberger, A.S. (1964), **Econometric Theory**. New York: Wiley.

Hamilton, J.D. and Jorda, O. (1997), "A Model for the Federal Funds Target," UC-San Diego Working Paper.

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

Hamilton, J.D. (1990), "Analysis of Time Series Subject to Changes in Regime," *Journal of Econometrics* 45, No. 1/2, pp. 39-70.

Hamilton, J.D. (1989), "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica* 57, pp. 357-84.

Hausman, J.A., Lo, A.W. and McKinlay, A.C. (1992), "An Ordered Probit Analysis of Transaction Stock Prices," *Journal of Financial Economics* 31, pp. 319-380.

Kim, C.J. and Nelson, C.R. (1998), "Business Cycle Turning Points, A New Coincident Index, and Tests of Duration Dependence Based on a Dynamic Factor Model with Regime Switching," *Review of Economics and Statistics* 80, pp. 188-201.

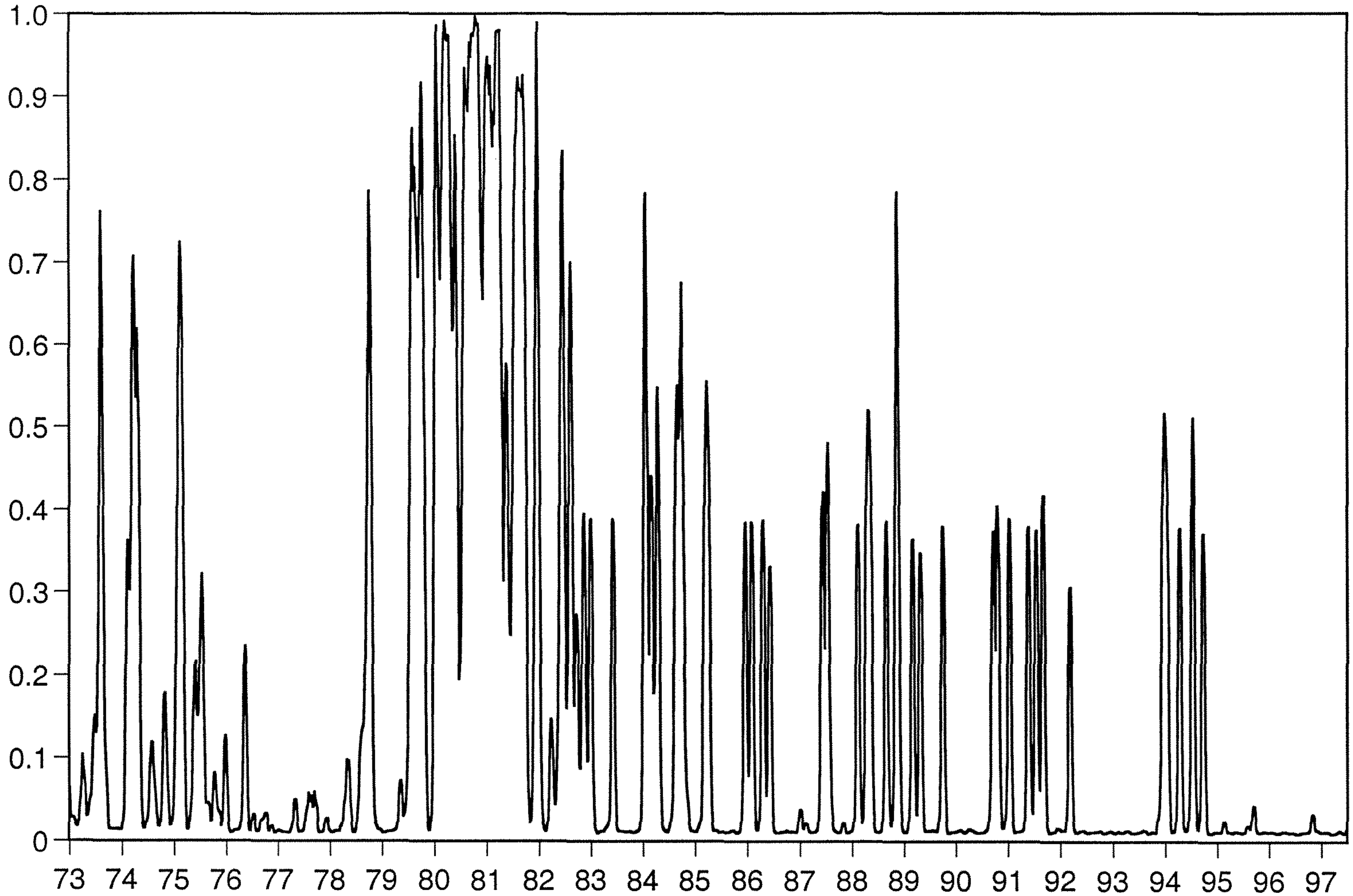
Maddala, G.S. (1983), **Limited-Dependent and Qualitative Variables in Econometrics.** New York: Cambridge Press.

Vlaar, P.J. and Palm, F.C. (1993), "The Message in Weekly Exchange Rates in the European Monetary System: Mean Reversion, Conditional Heteroscedasticity, and Jumps," *Journal of Business and Economic Statistics* 11, pp. 351-60.

Table 1: Observation categories based on size of prime rate change. Sample period: 12/5/72 – 12/13/97		
category	*criterion*	*frequency*
1	$\Delta PR > +.5$	15
2	$\Delta PR = +.5$	35
3	$\Delta PR = +.25$	68
4	$\Delta PR = 0$	1078
5	$\Delta PR = -.25$	49
6	$\Delta PR = -.5$	44
7	$\Delta PR < -.5$	12

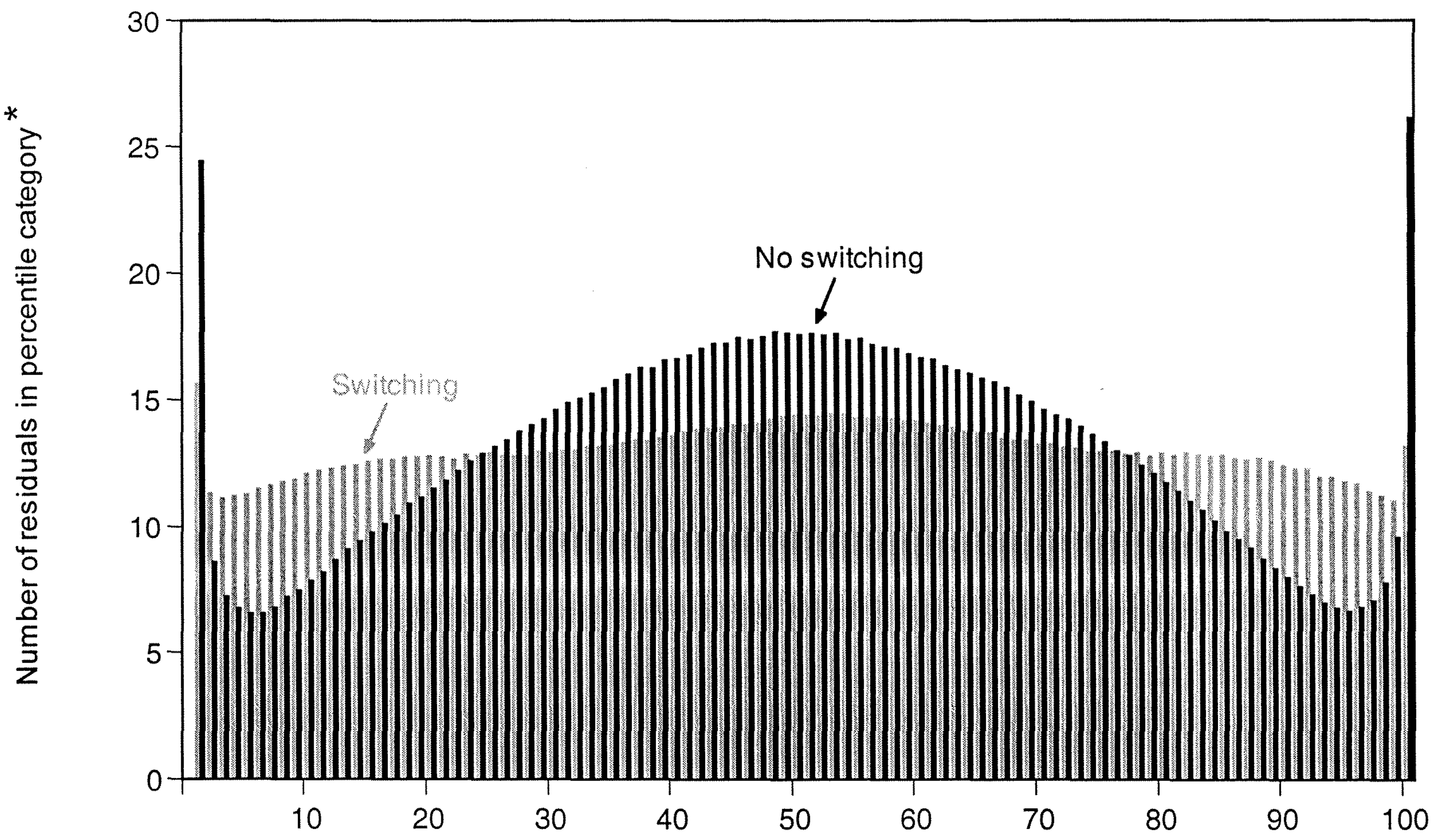
Table 2: Posterior distributions of parameters from 1302 weekly observations and 223 changes		
parameter	**Markov switching in variance**	**constant variance**
$\sigma^2_{S=0}$ post. mean	.0051	.0223
95% confid. interval	(.0036,.0073)	(.020,.025)
$\sigma^2_{S=1}$ post. mean	.1649	.0223
95% confid. interval	(.123,.223)	(.020,.025)
post. mean kurtosis of residuals	3.54	6.59
trans. prob. p post. mean	.9378	n.a.
95% confid. interval	(.913,.959)	n.a.
trans. prob. q post. mean	.6876	n.a.
95% confid. interval	(.563,.796)	n.a.
Intercept post. mean	.0069	.0091
95% confid. interval	(-.023,.036)	(001,.017)
Coefficients for changes in fed. funds rate are β_{1i}		
β_{11} post. mean	.3657	.6192
95% confid. interval	(.301,.435)	(.577,.661)
β_{12} post. mean	.2652	.4687
95% confid. interval	(.210,.322)	(.431,.506)
β_{13} post. mean	.1583	.2849
95% confid. interval	(.113,.206)	(.252,.318)
β_{14} post. mean	.0729	.1310
95% confid. interval	(.0466,.100)	(.109,.153)
Coefficients for changes in paper-bill spread are β_{2i}		
β_{21} post. mean	.1039	.1106
95% confid. interval	(.023,.182)	(.012,.207)
β_{22} post. mean	.0670	.0867
95% confid. interval	(.003,.137)	(.010,.164)
β_{23} post. mean	.0650	.1019
95% confid. interval	(.002,.129)	(.037,.166)
p-value: $\beta_{1i} = 0 \quad \forall i$	.000	.000
p-value: $\beta_{2i} = 0 \quad \forall i$	.013	.026
Note that the β parameters are sums of lag coefficients.		

Figure 1: Posterior mean of probability of high-variance state*



* Four-week moving average of the probability.

Figure 2: Goodness of fit by percentile of model-implied CDFs



* Average is 13 and would be uniformly equal to 13 in ideal model.