TI 2006-094/2
Tinbergen Institute Discussion Paper

Modeling Portfolio Defaults using Hidden Markov Models with Covariates

Konrad Banachewicz
Aad van der Vaart
*André Lucas**

Vrije Universiteit Amsterdam.

** Tinbergen Institute.*

Tinbergen Institute
The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Most TI discussion papers can be downloaded at http://www.tinbergen.nl.

Modeling Portfolio Defaults Using Hidden Markov Models with Covariates

Konrad Banachewicz,* Aad van der Vaart,

Vrije Universiteit Amsterdam Mathematics Department,

André Lucas

Vrije Universiteit Amsterdam, Department of Finance and Tinbergen Institute

October 24, 2006

Abstract

We extend the Hidden Markov Model for defaults of Crowder, Davis, and Giampieri (2005) to include covariates. The covariates enhance the prediction of transition probabilities from high to low default regimes. To estimate the model, we extend the EM estimating equations to account for the time varying nature of the conditional likelihoods due to sample attrition and extension. Using empirical U.S. default data, we find that GDP growth, the term structure of interest rates and stock market returns impact the state transition probabilities. The impact, however, is not uniform across industries. We only find a weak correspondence between industry credit cycle dynamics and general business cycles.

Key words: defaults; Markov switching; default regimes.

*We thank Standard & Poor's for making the CreditPro 7 database available for this research. We also thank the Dutch National Science Foundation (NWO/STO) for financial support. Corresponding author: Konrad Banachewicz, phone +31-20-5987832, mail konradb@few.vu.nl.

1 Introduction

Dependence of defaults within a given portfolio is an important issue in credit risk management. Proper modeling requires capturing the timing, as well as the total number of defaults. In recent years both the academic and the industrial research community have put forward dynamic models of default rates for at least three reasons. First, rating transition probabilities in general, and default probabilities in particular, vary over time and tend to co-move with general economic conditions. As a result, capital buffers required to cope with credit losses need to vary as well. Second, increased liquidity in financial markets for credit-related products has led to a shift in management paradigm: from counter-party credit risk assessment to dynamic, active management of credit portfolios. Third, the New Capital Accord of the Basel Committee on Banking Supervision (2005) allows banks to perform a larger part of risk management tasks using internal models. Here dynamic models for default probabilities can provide more efficient use of capital over stages of the business cycle.

Figure 1 shows the number of defaults among U.S. industrials recorded quarterly between 1981 and 2005. A visual inspection of the graph suggests, that the defaults are not independent over time – high and low default quarters appear to be clustered.

<INSERT FIGURE 1 HERE>

In this paper we capture these dynamics through a hidden Markov model (HMM), with the regime-switching probabilities depending on observed macro variables. In the credit risk context latent state models, such as the HMM in this paper, have an advantage over those based solely on observables. Often there is little or no theory as to which factors would be optimal as proxies for systematic credit risk. This problem is avoided in the HMM setting. The (hidden) states of the Markov chain in this paper correspond to the state of the credit market. We distinguish between normal and risky ("excited") credit market conditions. In risky credit market states, default probabilities increase. The transitions between the states are governed by probabilities that depend on selected macroeconomic variables. This is in line with the empirical evidence of Bangia et al. (2002), who show that defaults – and downgrades in general – are more likely during recessions than expansions.

In line with most credit risk portfolio models in the literature, the model we propose can be thought of as conditionally stationary, in the sense that non-stationarity in the general economic condition is captured by the conditioning macro variables. Conditional on the realization of the latent Markov process, credit exposures within the portfolio default independently. This results in a binomial distribution, with parameters depending on the risk state and the number of credits surviving up to time t.

We address a number of issues arising in the portfolio credit risk context. First, we investigate how much information on credit markets is contained in macro variables like GDP, interest rates and financial markets returns. Intuitively, in an expanding economy we should observe a decrease in default risk. We verify this by estimating the model in two versions. One model has a a simple latent risk state as in Crowder et al. (2005). The other model has an observable part (macro factors) and a hidden part (interpreted as condition of the credit market). In addition we analyze whether the industry sector influences the impact of economic conditions on default risk. The current approach with a latent component is less prone to the choice of incorrect macro proxies, as signaled in Lucas and Klaasen (2006). Because the credit state is a latent component, we can still have a systematic credit risk factor at the portfolio level even though all macro variables are incorrect. The model will then collapse to the basic HMM of Crowder et al. (2005).

Based on the HMM for defaults, we can construct an early warning mechanism for high default probability regimes. The importance of such regimes for setting capital buffers was clearly illustrated in Bangia et al. (2002) in a switching model. We confront our predictions with the NBER classification of business cycle states.

In this way we can analyze the dependence between the business cycle (behavior of the general economic variables) and the credit cycle (fluctuations of the recovered hidden credit risk state process), both in particular sectors and the economy as a whole. We find only mild correspondence between the two, indicating that credit and business cycles can have their own separate dynamics. Our current model classifies credit market conditions into a finite number of different levels for default intensities. This makes the model easy to estimate using standard methods like the EM algorithm. Our approach complements related papers that either use observed rather than hidden regimes, such as Bangia et al. (2002) and Nickell, Perraudin, and Varotto (2000), or a continuous number of states for economy-wide default intensities, see Koopman et al. (2005) and Duffie et al. (2006).

The paper is organized as follows. Section 2 describes the basic setup along with the proposed extensions and the statistical methods employed. The details of the calculations are deferred to the appendix. Empirical results are described in section 3. Section 4 concludes and suggests directions for future research.

2 Model formulation

2.1 Theoretical setup

At time $t = 1, \dots, T,$, we consider a portfolio of N_t units. Each unit can be thought of as a defaultable counterparty. We are interested in modeling the number of defaults over time. Underlying the default dynamics is a latent, discrete process W_t capturing current credit market conditions. In the basic version of the model, W_t is modeled as a time-homogeneous Markov chain taking values in the set $\{1, \dots, s\}$. State values 1 and s correspond to the lowest and highest default regimes, respectively. Conditional on the hidden process W_t the units behave independently. Hence, the total number of defaults D_t at time t has a binomial distribution. For the number of defaults this means that $D_t|(W_t = i) \sim \text{Bin}(N_t, \alpha_i)$, where α_i denotes the default probability in state i.

The hidden Markov chain W_t is characterized by its transition matrix $\boldsymbol{Q}_t$ together with the distribution of the initial state $\Pr(W_1 = i) = \pi_i$. As motivated in the introduction, we model the matrix $\boldsymbol{Q}_t$ as a function of observed covariates X_t, where X_t may contain macroeconomic variables such as growth or interest rates. This extends the framework of Crowder et al. (2005) that uses constant $\boldsymbol{Q}_t \equiv \boldsymbol{Q}$. It also departs from the setup of Bangia et al. (2002) or Nickell et al. (2000), where the latent W_t is replaced directly by an observed variable. We choose the standard (multinomial) logistic link function between X_t and the entries of $\boldsymbol{Q}_t$,

$$q_{ij,t} = \Pr(W_{t+1} = j | W_t = i, X_t) = \frac{\exp(\Phi'_{ij} X_t + \eta_{ij})}{\sum_{j=1}^{s} \exp(\Phi'_{ij} X_t + \eta_{ij})}, \tag{1}$$

Other link functions, such a probit are of course also possible. For identification, we have to restrict one of the Φ_{ij}s and η_{ij}s per row i, for example $\Phi_{ii} \equiv 0$ and $\eta_{ii} \equiv 0$. If $\Phi_{ij} \equiv 0$ for all i and j, then a multistate version of the model of Crowder et al. (2005) is recovered.

The number of units N_t at time t is affected by the number of defaults over the previous period, given by D_{t-1}. In addition, N_t may also increase because units enter the sample (births), or decrease because units leave the sample for reasons other than default. Obvious examples of the latter in our current context are firms that merged or were acquired. In this paper we follow the common approach in the literature and treat births and withdrawals as exogenous, see, e.g., Bangia et al. (2002) and Nickell et al. (2000). Alternatively, births and deaths could be modeled in a similar way to D_t. See Duffie et al. (2006) for an analogue in the setting of intensity models.

2.2 Estimation

We gather all the model parameters into a single vector θ. In order to estimate θ, we use the Expectation-Maximization (EM) algorithm. Details are provided in the appendix. The main idea is the following. We can easily write down the joint log-likelihood $\ln p_\theta(D, W|X)$ of the observed default sequence D_t and the associated hidden state evolution W_t conditioned on the sample path of the macro process X_t. As we can only observe D_t, N_t, and X_t, we have to integrate out W_t. We perform the integration using an initial estimate $\tilde{\theta}$ of θ. A new estimate of the parameter vector is given by

$$\theta^* = \operatorname*{argmax}_\theta \mathrm{E}_{\tilde{\theta}} \left[\sum_{t=1}^{T} \ln p_\theta(D_t, W_t | N_t, W_{t-1}, X_{t-1}) \right]. \tag{2}$$

The integration can be done analytically. In the next step we update $\tilde{\theta}$ to the value θ^* and again compute a new estimate using (2). This iterative procedure is repeated until convergence. Usually this type of estimation problem in the Markov switching models class is solved using the Baum-Welch algorithm, see Rabiner (1989). The algorithm is not directly applicable in our current setting. The distribution of the output D_t changes over time due to fluctuations in the number of survivors N_t, whereas the BW method makes extensive use of the time-invariance property. Therefore, we propose a modification of the algorithm taking into account this variation, as well as the presence of covariates. The general approach is similar. We proceed by defining forward and backward variables as in Rabiner (1989). The resulting update formulas for parameter estimates are, however, different from the standard case.

We obtain standard errors of the parameter estimates by using the *block bootstrap*. We split the original series into blocks, with block lengths generated from a geometric distribution. The blocks are then concatenated, until a new series of length T is formed. We create M new replicate series in this way, sample standard deviation of the estimates over the M replications serve as standard errors. For details of the implementation procedure, see Politis and Romano (1994).

2.3 Recovered latent state distribution

Given the probabilistic structure of our model, we can analyze the evolution of the latent state process W_t. We use the recursive algorithm provided by Hamilton (1994). The method uses three different conditional distributions of W_t (based on varying information sets). Define the information set as

$$\mathcal{F}_t = \{X_t, X_{t-1}, \ldots, D_t, D_{t-1}, \ldots, N_t, N_{t-1}, \ldots\}$$

and let

$$\xi_{t|\tau} = \begin{bmatrix} \Pr(W_t = 1|\mathcal{F}_\tau) \\ \ldots \\ \Pr(W_t = s|\mathcal{F}_\tau) \end{bmatrix} = \begin{bmatrix} \xi_{t|\tau}(1) \\ \ldots \\ \xi_{t|\tau}(s) \end{bmatrix}. \tag{3}$$

By substituting $\tau = t-1, \tau = t$, and $\tau = T$ we obtain *predicted*, *filtered*, and *smoothed* estimates of the distribution of the latent state W_t, respectively. Details of the recursive procedure to compute $\xi_{t|\tau}$ in each of these cases are provided in the appendix.

2.4 Forecasting

Using the notation in (3) we can construct a one-step prediction for the default process

$$\begin{aligned}\Pr(D_{T+1} = k|\mathcal{F}_T, N_{T+1}) &= \sum_{i=1}^{s} \Pr(D_{T+1} = k|W_{T+1} = i, N_{T+1}, \mathcal{F}_T) \cdot \Pr(W_{T+1} = i|\mathcal{F}_T) \\ &= \sum_{i=1}^{s} \mathrm{dbin}(k; N_{T+1}, \alpha_i) \cdot \xi_{T+1|T}(i), \end{aligned} \tag{4}$$

where $k \in \{0, 1, \ldots, N_{T+1}\}$, $\mathrm{dbin}(k; N, p)$ denotes the binomial probability of k successes when drawing N trials with success probability p. As N_{T+1} is the number of units at the *start* of the $(T+1)$th period, it is known at time T. Note that to obtain more-than-one-step-ahead forecasts, an auxiliary model is needed for the processes N_t and X_t. For N_t, an obvious choice is to abstract from the birth and withdrawal processes and update N_{T+1} by the predicted default distribution of D_{T+1} only, i.e.

$$\begin{aligned}
&\Pr(D_{T+2} = k | \mathcal{F}_T, X_{T+1}, N_{T+1}) = \\
&= \sum_{i=1}^{s} \sum_{j=1}^{s} \sum_{l=0}^{N_{T+1}} \Pr(D_{T+2} = k | X_{T+1}, N_{T+2} = N_{T+1} - l, W_{T+2} = j, \mathcal{F}_T) \\
&\cdot \Pr(D_{T+1} = l | W_{T+1} = i, \mathcal{F}_T, N_{T+1}) \cdot q_{ijt} \cdot \xi_{T+1|T}(i) \qquad (5)
\end{aligned}$$

To integrate out the X_t process (in this case X_{T+1}), a time series model for X_t is needed. Obvious choices include simple univariate time series models as in Duffie et al. (2006) or multivariate models such as vector autoregressive (VAR) models, see Kavvathas (2001) . Note that the forecast performance in that case becomes contingent on the quality of both the time series model for D_t and that for X_t.

3 Empirical results

3.1 Data description

In this section we apply the previous methodology to a two-state HMM. We distinguish a high and a low default regime. The data for our case study comes from the CreditPro 7.0. database of Standard & Poor's. The time series of interest consist of registered defaults in the U.S. economy between January 1981 and July 2005, sampled quarterly. The sample period encompasses both expansions and contractions. This is important, as part of our interest concerns the difference of conditional default probabilities between economic regimes.

For the sake of comparison, we group the observations into four industry blocks. We distinguish industrials (automotive / metal / capital goods, energy and natural resources, forest and building products / homebuilders, health care / chemicals, utility), services (consumer / service sector, leisure time / media, transportation), financials (financial institutions, insurance, real estate), and high-tech (high technology / computers / office equipment, telecommunication). The classification in parentheses follows the original industry classification provided in the database. The number of observations in each industry can be found in Table 1. A graph of the time series is presented in Figure 2.

<INSERT TABLE 1 HERE>

<INSERT FIGURE 2 HERE>

In line with previous empirical work in this area, we define the number of exposures N_t for each sector as the number of active companies on January 1 of year t minus the number of withdrawals over the subsequent year. A withdrawal is defined as the event of a company leaving the database for other reasons than default. If a company first withdraws and later defaults, this is recorded in the database. In such cases, we skip the withdrawal event and only account for the default event. In this way, we mitigate any biases due to strategic default behavior.

From earlier studies on the impact of macroeconomic variables on aggregate default rates see Couderc and Renault (2004), we take four candidate macro time series. We consider the real GDP growth rate over the past year, the slope of the term structure (measured by the difference between the 5 year and 3 month treasury yields), and the S&P 500 stock index returns over the past year. All series were sampled quarterly to match the quarterly default sample. We also experimented with a realized volatility variable based on daily S&P 500 index returns. This variable, however, was insignificant in all of our estimation results. We therefore exclude it from the remainder of the discussion.

3.2 Parameter estimates

First, we estimate the baseline, time homogeneous HMM model of Crowder et al. (2005) for our data set. The results are given in the top panel of Table 2. The estimation results for the total default series clearly indicate that there are two default regimes. The default probability in the low regime (0.19%) is only one third of its counterpart in the high default regime (0.61%). The standard errors indicate that the difference is statistically significant. The probability of remaining in the low default regime is high, 92.88%. This implies an expected duration of high default regimes of about 13 quarters. The high default regime has a lower probability of remaining in this regime (85.30%), and a correspondingly lower expected duration of about 6 quarters. The results are robust over industries. The most persistent estimate is for the low default regime for the services industry. The estimate implies an expected regime duration of somewhat more than 8 years. In all cases, there is an apparent asymmetry between low default regimes and shorter-lived high default regimes. This corresponds with results for similar models estimated for business cycles, see Hamilton (1994). The expected durations in our data set for default regimes, however, are much lower than those found for business cycles. The most interesting jump in the default probability α_i is for the high-tech industry. The high regime α_i is almost eightfold the size of its low default regime counterpart. This is mainly due to the crash of the technology bubble in the early 2000s.

<INSERT TABLE 2 HERE>

We now consider the model with covariates. The estimation results are presented in the lower panel of Table 2. Again, we notice the existence of two significantly different default regimes. The regime default probabilities α_i are very stable with respect to the results for the time homogeneous model. For the total default series, GDP growth and stock returns are significant for the low default regime probability. Both variables have the expected sign: high growth and high stock returns cause the low default regime to persist longer. The converse effect is seen for the high default regimes. High growth rates cause the high default regime to become less persistent. The stock return is not significant for the high default regime transition probability.

When we consider the results for the individual industries, we see a similar pattern. Except for the high-tech sector, GDP growth has a significant effect on low default regime persistence. The magnitude of the coefficient is largest for the financial industry. For the high default regime, we again see that GDP growth decreases regime persistence. The effect is only significant, however, for the industrials and high-tech sectors.

The stock returns and interest rate spread variables reveal a more mixed pattern. When significant, they have the correct sign: positive for low and negative for high default regime persistence. The difference in significance of the different variables for the various industries might be attributable to differences in leads and lags of the industry default cycle with the general business cycle, see also Figure 2. For example, GDP growth typically coincides with the business cycle, whereas the interest rate spread typically leads. Such timing differences may prove an important element in portfolio credit risk modeling.

A final result emerging from the bottom panel in Table 2 is that it is more difficult to predict the high than the low default regime. The number of significant covariates is much smaller for $q_{22,t}$ (high default persistence) than for $q_{11,t}$ (low default persistence).

3.3 Default state estimates

We continue our discussion of the results by presenting the estimates of the predicted and smoothed default regime probabilities $\xi_{t|\tau}(2)$. The results are presented in Figures 3 and 4.

<INSERT FIGURE 3 HERE>

We immediately see in Figure 3 that the introduction of covariates makes the predicted probabilities of the next regime more extreme. For the model without covariates, the probability of the next quarter being a high default regime varies between 85.30% and 7.12% (=100 - 92.88). This immediately follows from the estimation results in Table 2. If the smoothed probability of the high default regime is 100%, we obtain the maximum predicted probability. The converse holds if this smoothed probability is 0%. By contrast, if we add covariates, also the value of X_t becomes relevant for the predicted regime probability. As an example, consider the total series and the GDP variable. Even if the smoothed probability of the current state being in the high default regime is very high (low), an extremely high (low) GDP growth rate may cause the predicted probability of a subsequent high default regime to be much lower (higher). As high growth is correlated with low default regimes according to Table 2, we see the general effect that high regime probabilities from the time homogeneous model are driven even further towards 100% because of the X_ts additional impact, e.g., GDP growth rates being typically low at that time.

When we move on to the smoothed default probabilities in Figure 4, the differences between the model with and without covariates are quite small. The main differences are the high (low) regime in the mid 1980s for the basic (extended) model, and the low (high) regime in 1999 for the basic (extended model). The sample appears too short to obtain definite statements on whether one model outperforms the other. This is also supported by the industry wise comparison of predicted probabilities in Figure 5. Now the differences between the basic and extended model are more pronounced. For some years and industries, the extended model gives a clearer and more timely signal.

<INSERT FIGURE 4 HERE>

<INSERT FIGURE 5 HERE>

One of the interesting things about Figure 5 is the difference in timing and duration of high default regimes between industries. Differences are so large in several cases that one can clearly question the validity of pooling across industries to get a better grasp of systematic portfolio credit risk factors as is done in much empirical work.

Finally, we check the regime classifications from our model with the NBER business cycle classifications to see whether high default regimes coincide with business cycle recessions. This is done in Figures 7 and 6. We consider the smoothed probabilities. Figure 7 shows that for the total default series, the high smoothed probabilities generally contain the NBER classified recession. The high default regimes are, however, much more prolonged.

<INSERT FIGURE 7 HERE>

<INSERT FIGURE 6 HERE>

If we disaggregate these results by industry in Figure 6, we again see a more diverse picture emerging. The recession in the early 1980s only has some correspondence to the high default regime for financials during that period. The recession, however, starts and ends earlier than the high default regime. For the recession in the early 1990s, we see more likeness across industries. In all industries we see high default regimes surrounding these years. Again, the financials are somewhat departing in that the high default regime is much longer than the recession and starts much earlier, and ends much later. At the opposite side of the spectrum, the high-tech companies show an increased probability of a high default regime in the years directly adjacent to the recession, while the recession itself is classified as low default. For the recession in the early 2000s, all industries have a high default regime. That for the financials is very short lived, while those for the high-tech and industrials are longest. In short, there appears to be some correspondence with business cycle classifications and high default regimes, but again the pattern is mixed.

In terms of durations, starting and ending dates, and consistency across industries, there appears no clear match between business cycle and credit cycle dynamics. This is in line with findings by Koopman et al. (2006) and Duffie et al. (2006) using different modeling approaches, and may have important implications for credit risk management.

4 Conclusions

In this paper we developed a Hidden Markov Model (HMM) to model and predict corporate default frequencies. This model contributes to the literature in that it is relatively easy to estimate using the EM extension to Rabiner (1989) as proposed in this paper. Moreover, it allows for differences between credit and business cycles by making the state of the credit cycle (high or low) a latent process.

The model enables us to address a number of relevant issues. First, we find asymmetry in the credit cycle dynamics. Both in the basic (without covariates) and extended versions of the model, the persistence of the good and bad credit states differs significantly. Moreover, macro variables affect the persistence of a low default regime much more than that of a high default regime. Due to an improvement in explanatory power, we also see that sharper predictions are obtained for one-quarter-ahead predicted default rates based on observable macro variables.

Second, we present empirical evidence of differences in exposure to systematic risk across economic sectors. In particular, we found that the systematic credit risk factor varies considerably across industries. This cautions the use of pooling across industries in single factor portfolio credit risk models.

Finally, we can draw conclusions about the interdependence between credit and business cycles. Our findings suggest, that the two are correlated, but the correlation is far from perfect. This implies, that credit portfolio models that only condition on business cycle proxies may miss out on a significant part of systematic portfolio credit risk.

References

Bangia, A., F. X. Diebold, A. Kronimus, C. Schagen, and T. Schuermann (2002). Ratings migration and the business cycle, with application to credit portfolio stress testing. *Journal of Banking and Finance 26(2-3)*, 445–474.

Couderc, F. and O. Renault (2004). Times-to-default: Life cycle, global and industry cycle impacts. Technical Report 141, FAME.

Crowder, M., M. Davis, and G. Giampieri (2005). A hidden markov model of default interaction. *Quantitative Finance 5*, 27–34.

Duffie, D., A. Eckner, G. Horel, and L. Saita (2006). Frailty correlated default. Working paper, Graduate School of Business, Stanford University.

Duffie, D., L. Saita, and K. Wang (2006). Multi-period corporate default prediction with stochastic covariates. *Journal of Financial Economics*, forthcoming.

Hamilton, J. D. (1994). State - space models. In *Handbook of Econometrics, Vol IV.* Elsevier Science B.V.

Kavvathas, D. (2001). Estimating credit transition probabilities for corporate bonds. In *New Orleans Meetings*. AFA.

Koopman, S. J., R. Kräussl, A. Lucas, and A. Monteiro (2006). Credit cycles and macro fundamentals. Discussion Paper 2006-023/2, Tinbergen Institute.

Koopman, S. J., A. Lucas, and R. Daniels (2005). A non-gaussian panel time series model for estimating and decomposing default risk. Discussion Paper 2005-060/4, Tinbergen Institute.

Lucas, A. and P. Klaasen (2006). Discrete versus continuous state switching models for portfolio credit risk. *Journal of Banking and Finance 30(1)*, 23–35.

Nickell, P., W. Perraudin, and S. Varotto (2000). Stability of rating transitions. *Journal of Banking and Finance 24*, 203–227.

Politis, D. N. and J. P. Romano (1994). The stationary bootstrap. *Journal of the American Statistical Association 89*, 1303–1313.

Rabiner, L. R. (1989, February). A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE 77*(2), 257–286.

Appendix

EM algorithm

As usual, we start with writing the down the likelihood of an observed sequence of defaults $D = \{D_1, \dots, D_T\}$ along with an associated hidden state sequence $W = \{W_1, \dots, W_T\}$ and macro process $X = \{X_1, \dots, X_T\}$. Define $\tilde{W}_{it}$ as an indicator variable taking the value 1 if $W_t = i$ and 0 otherwise. Sticking to the notation of Section 2.1, we have the log-likelihood

$$\ln p_\theta(D, W|X) = \sum_{i=1}^{s} \tilde{W}_{it} \ln \pi_i + \sum_{t=2}^{T} \sum_{i=1}^{s} \sum_{j=1}^{s} \tilde{W}_{i,t-1} \tilde{W}_{j,t} \ln q_{ij,t-1} \qquad \text{(A1)}$$
$$+ \sum_{t=1}^{T} \left[\ln \binom{N_t}{D_t} + \sum_{i=1}^{s} \tilde{W}_{it} \left[D_j \ln \alpha_i + (N_j - D_j) \ln(1 - \alpha + i) \right] \right] .$$

Assume that we have obtained some initial estimate θ_0 of the model's parameters. Given θ_0, we can compute so called *forward* and *backward* variables, which simplify the computations and clarify the notation (see Rabiner (1989)). In our case, define

- the probability of a particular number of defaults, given a state of the Markov Chain and the model parameters,

$$b_j(d_t) = P(D_t = d_t | W_t = j, X) = \binom{N_t}{d_t} \alpha_j^{d_t} (1 - \alpha_j)^{N_t - d_t} \ ;$$

- the probability of the partial observation sequence $d_1, \dots, d_t$ and state i at time t, given the model parameters θ,

$$\bar{\alpha}_j(1) = \pi_j b_j(d_1) \ , \quad j = 1, \ldots, s \ ,$$
$$\bar{\alpha}_j(t) = \left[\sum_{i=1}^{s} \bar{\alpha}_i(t-1) q_{ij,t}\right] b_j(d_t) \ , \quad j = 1, \ldots, s \quad t = 2, \ldots, T \ ;$$

- the probablity of a partial observation sequence from $t+1$ to the end, given state i at time t and the model parameters,

$$\bar{\beta}_j(T) = 1 \quad j = 1, \ldots, s \ ,$$
$$\bar{\beta}_j(t) = \sum_{i=1}^{s} q_{ij,t} b_i(d_{t+1}) \bar{\beta}_{t+1}(i) \ , \quad t = T-1, \ldots, 1 \ ;$$

- the probability of being in state i at time t, given the observed sequence and model parameters,

$$\gamma_j(t) = \frac{\bar{\alpha}_j(t)\bar{\beta}_j(t)}{\sum_{i=1}^{s} \bar{\alpha}_i(t)\bar{\beta}_i(t)} \ , \quad j = 1, \ldots, s \quad t = 1, \ldots, T \ ;$$

- the probability of the Markov chain being in state i at time t and state j at $t+1$, given the model parameters and the observation sequence,

$$\xi_t(i,j) = \frac{\bar{\alpha}_t(i) q_{ij,t} b_j(d_{t+1}) \bar{\beta}_{t+1}(j)}{\sum_{i=1}^{s} \sum_{j=1}^{s} \bar{\alpha}_i(t) q_{ij,t} b_j(d_{t+1}) \bar{\beta}_j(t+1)} \ , \ i,j = 1, \ldots, s \quad t = 1, \ldots, T-1 \ .$$

Let $\mathcal{S}$ denote the space of all possible sample paths of the latent process. The E-step of the algorithm requires computing the expectation

$$\mathrm{E}_{\theta_0}[\ln p_\theta(D, W)|D, X] = \sum_{w \in \mathcal{S}} \ln p_\theta(D, w|X) p_{\theta_0}(D, w|D, X) \ .$$

For a fixed sequence of states (sample path) w, we have

$$p_\theta(D, w|X) = \pi_{w_0} \prod_{t=2}^{T} q_{w_{t-1} w_t, t} b_{w_t}(d_t) \ ,$$

such that

$$\mathrm{E}_{\theta_0}[\ln p_\theta(D,W)|D,X] = \sum_{w\in\mathcal{S}} (\ln \pi_{w_0}) p_{\theta_0}(D,w|D,X) \tag{A2}$$
$$+ \sum_{w\in\mathcal{S}} \left(\sum_{t=2}^{T} \ln p(w_{t-1}|w_t)\right) p_{\theta_0}(D,w|D,X) + \sum_{w\in\mathcal{S}} \left(\sum_{t=2}^{T} \ln b_{w_t}(d_t)\right) p_{\theta_0}(D,w|D,X) \ .$$

Our objective is to maximize (A2) with respect to θ. Since the parameters appear in groups, we can seek the maximum of each of the three sums in the above display separately. We start with decomposing the component related to π:

$$f(\pi) = \sum_{w\in\mathcal{S}} \ln \pi_{w_0} p_{\theta_0}(D,w) = \sum_{i=1}^{s} \ln \pi_i \mathrm{P}_{\theta_0}(w_0 = i, D, X) \ .$$

A maximal value of f under the constraint $\sum_{i=1}^{s} \pi_i = 1$ is obtained for $\pi = \gamma_i(1)$.

The second component of the likelihood contains the time-dependent probabilities $\boldsymbol{Q}_{t,ij}$. Those are described by the coefficients Φ_{ij} and η_{ij}. As the maximum likelihood estimators do not have a closed form expression, we resort to numerical maximization for this part of the likelihood.

Finally, the updates for α_i's can be obtained by finding the optimal argument values for the function

$$h(\alpha_1,\ldots,\alpha_s) = \sum_{w\in\mathcal{S}} \left(\sum_{t=2}^{T} \ln b_{w_t}(d_t)\right) p_{\theta_0}(D,w) = \sum_{i=1}^{s}\sum_{t=2}^{T} (\ln b_i(d_t)) \mathrm{P}_{\theta_0}(w_t = i, D, X) \ .$$

The optimal arguments result in update formulas

$$\alpha_i^* = \frac{\sum_{t=2}^{T} d_t \gamma_i(t)}{\sum_{t=2}^{T} N_t \gamma_i(t)} \ .$$

Hamilton algorithm

Define iteratively for $t = 1, \dots, T$ vectors of *prediction probabilities*

$$\xi_{t|t-1} = \begin{bmatrix} P_\theta(W_t = 1|\mathcal{F}_{t-1}) \\ \dots \\ P_\theta(W_t = s|\mathcal{F}_{t-1}) \end{bmatrix} \quad \text{, where} \quad \xi_{1|0} = \begin{bmatrix} \pi_1 \\ \dots \\ \pi_s \end{bmatrix} \quad .$$

For $t = 1, \dots, T$ it holds that

$$\xi_{t+1|t} = \frac{\boldsymbol{Q}'_t(\xi_{t|t-1} \circ \eta_t)}{\mathbf{1}`(\xi_{t|t-1} \circ \eta_t)} \quad \text{, with} \quad \eta_t = \begin{bmatrix} P_\theta(D_t = d_t|W_t = 1, \mathcal{F}_{t-1}) \\ \dots \\ P_\theta(D_t = d_t|W_t = s, \mathcal{F}_{t-1}) \end{bmatrix},$$

with $\mathbf{1}$ denoting a vector of ones, and $\circ$ denoting an element-by-element multiplication. As a by-product we obtain *filtered probabilities*, representing the distribution of the latent process at time t based on the information available at that time,

$$\xi_{t|t} = \frac{\xi_{t|t-1} \circ \eta_t}{\mathbf{1}'(\xi_{t|t-1} \circ \eta_t)} \quad .$$

Furthermore, we can compute *smoothed probabilities*, useful for reproducing the evolution of the hidden state process,

$$\xi_{t|T} = \begin{bmatrix} P_\theta(W_t = 1|\mathcal{F}_T) \\ \dots \\ P_\theta(W_t = s|\mathcal{F}_T) \end{bmatrix} \quad ,$$

with $\xi_{T|T}$ obviously equal to the most recent filtered probability. The probabilities $\xi_{t|T}$ are obtained through a backward recursion

$$\xi_{t|T} = \xi_{t|t} \circ \left\{ \boldsymbol{Q}_t \left(\xi_{t+1|T} \div \xi_{t+1|t} \right) \right\}, \quad t = T-1, \dots, 1 \quad , \tag{A3}$$

where $\div$ denotes element-by-element division.

Table 1: Registered defaults within U.S. economy – sector split according to the CreditPro 7.0 database classification of economy branches

Economy sector	**Defaults**
Consumer / service sector	290
Aerospace / automotive / capital goods / metal	209
Leisure Time / Media	139
Telecommunications	85
Energy and natural resources	77
Health care / chemicals	76
Transportation	62
Forest and building products / homebuilders	59
Financial Institutions	54
High technology / computers / office equipment	50
Utility	30
Insurance	25
Real Estate	8

Table 2: Estimation results

The table contains the estimation results for the HMM model with (bottom panel) an without (top panel) covariates. The regimes i correspond to the low (1) and high (2) default regime. Columns with parameter values are labeled accordingly. We consider the complete sample of defaults (Total) as well as defaults distinguished across industries (Industrials, Services, Financials, High-tech). The model for transition probabilities is given by

$$q_{ijt} = \Pr(W_{t+1} = j | W_t = i, X_t) = \exp(\Phi'_{ij} X_t + \eta_{ij}) \Big/ \sum_{j=1}^{s} \exp(\Phi'_{ij} X_t + \eta_{ij}),$$

for $i, j = 1, 2$, where we set $\Phi_{12} = \Phi_{21} = 0$ and $\eta_{12} = \eta_{21} = 0$ for identification. X_t contains GDP growth rates (with coefficients Φ_{11}^{GDP} and Φ_{22}^{GDP}), the term structure of interest rate (defined as the 5 year minus the 3 month rate, with coefficients Φ_{ii}^{SPRD}), and the return on the S&P500 index (Φ_{ii}^{SP500}). The default probability in a specific regime is given by α_i. Both q_{ii} and α_i are measured in percentages. Standard errors obtained from the block-bootstrap with replications are in parentheses.

	Total		Industrials		Services		Financials		High-tech	
	low default	high default	low default	high default	low default	high default	low default	high default	low default	high default
	Time homogeneous model									
q_{ii}	92.88	85.30	92.49	87.25	97.00	90.36	93.78	83.93	96.63	86.06
	(2.35)	(5.04)	(3.43)	(15.19)	(3.53)	(5.91)	(11.92)	(8.63)	(2.77)	(19.98)
α_i	0.19	0.61	0.15	0.60	0.36	1.12	0.04	0.43	0.18	1.41
	(0.02)	(0.05)	(0.01)	(0.03)	(0.03)	(0.09)	(0.02)	(0.05)	(0.02)	(0.10)
	Model with covariates									
Φ_{ii}^{GDP}	4.65	-1.96	1.54	-1.50	1.93	-1.19	7.02	-4.56	2.16	-5.18
	(0.71)	(0.68)	(0.11)	(0.17)	(0.86)	(2.46)	(1.05)	(1.98)	(1.30)	(1.94)
Φ_{ii}^{SPRD}	1.38	-1.20	3.98	0.78	1.47	3.84	-2.07	2.27	1.31	-5.77
	(2.27)	(0.78)	(6.12)	(1.00)	(0.37)	(6.10)	(3.37)	(6.08)	(0.62)	(2.64)
Φ_{ii}^{SP500}	1.19	0.46	-0.58	0.24	-0.21	-0.78	1.59	0.31	-3.26	0.80
	(0.43)	(0.46)	(0.70)	(1.13)	(0.79)	(0.79)	(1.11)	(1.66)	(2.10)	(1.46)
α_i	0.17	0.57	0.18	0.69	0.39	1.18	0.05	0.45	0.18	1.31
	(0.02)	(0.04)	(0.04)	(0.08)	(0.03)	(0.06)	(0.02)	(0.03)	(0.03)	(0.15)

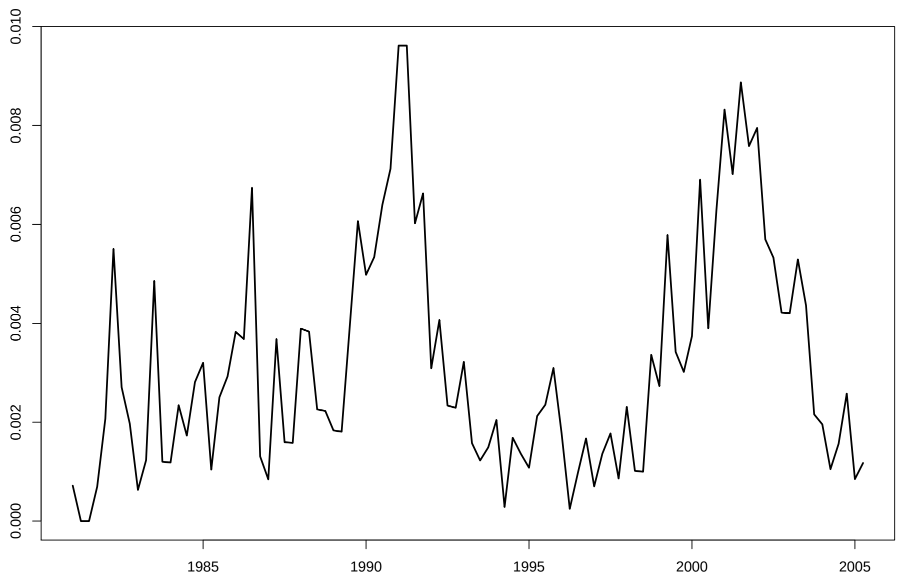

Figure 1: Default rate in the U.S. economy over the sample period

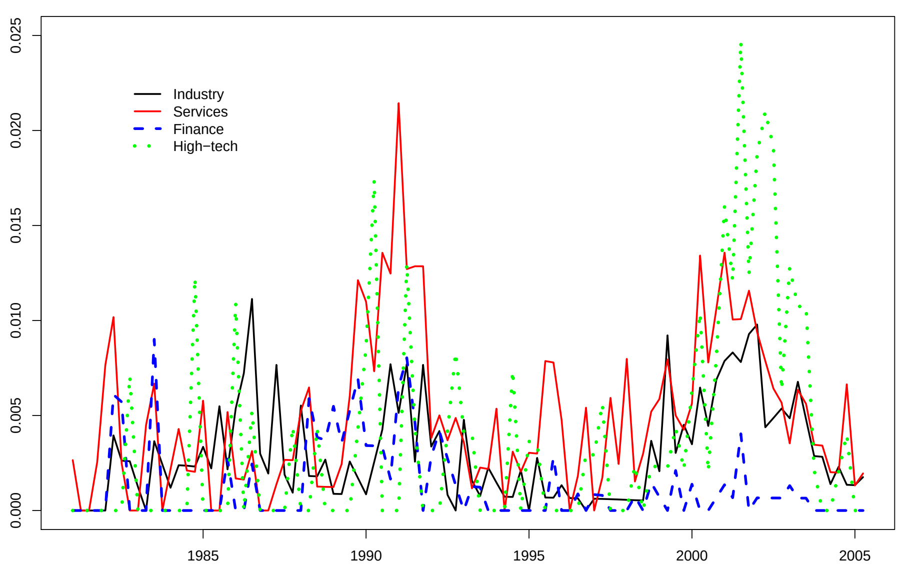

Figure 2: Default rates in the U.S. industry over the sample period – sector split

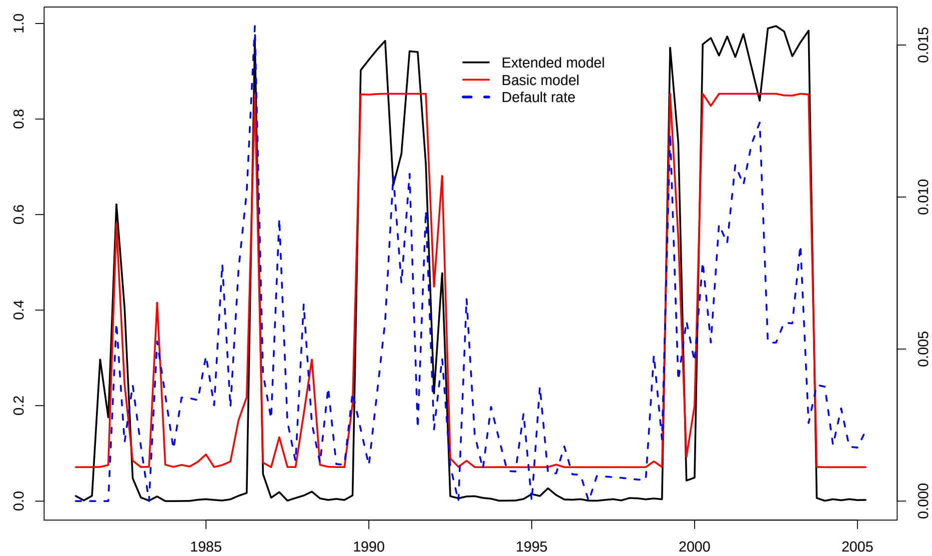

Figure 3: Predicted probability of crisis: basic and extended model vs realized default rate on the entire sample

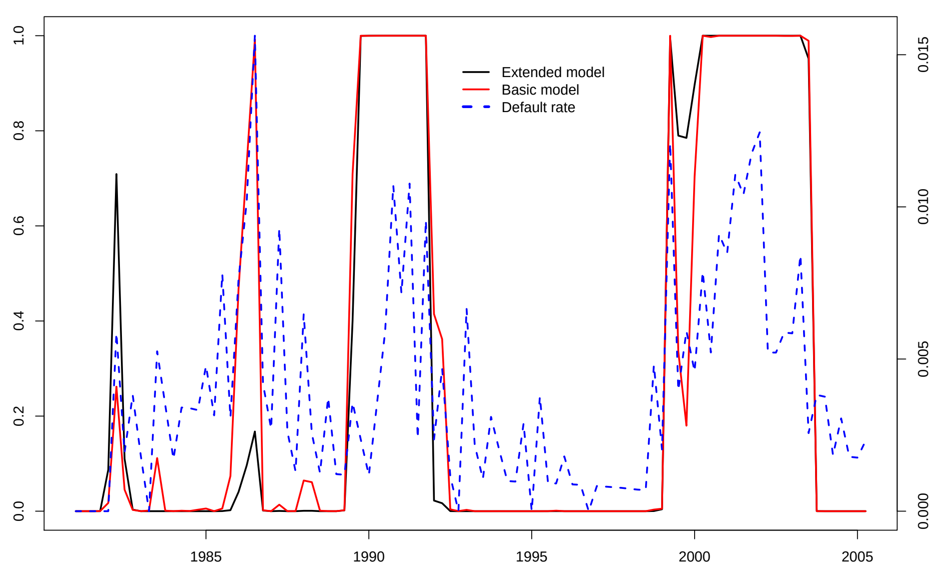

Figure 4: Smoothed probability of crisis: basic and extended model vs realized default rate on the entire sample

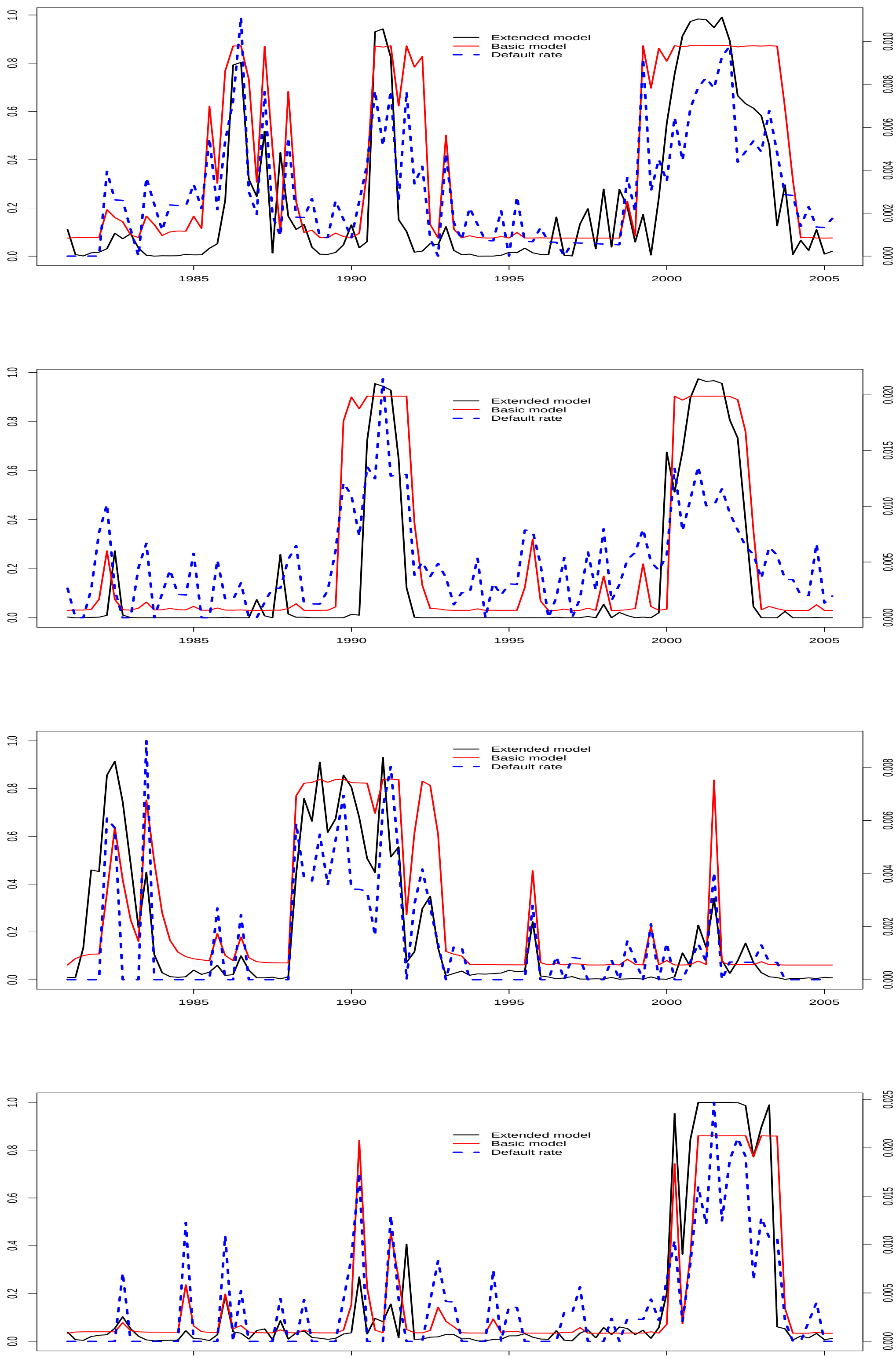

Figure 5: Predicted probability of crisis: basic and extended model vs realized default rate; top to bottom: industry, services, finance and high-tech

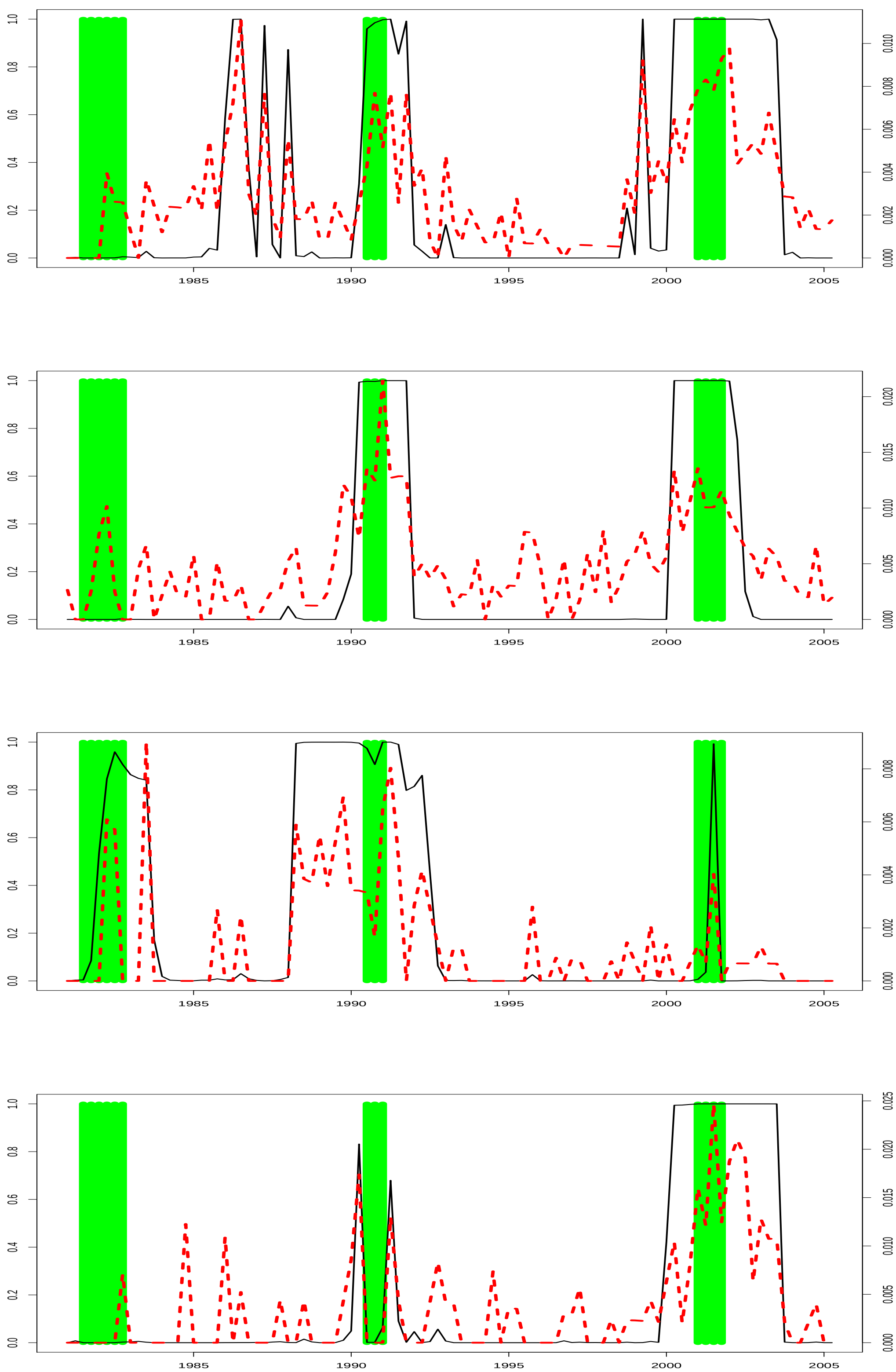

Figure 6: Realized default rate (dashed line) and smoothed crisis probability (solid line) vs NBER recession periods; top to bottom: industry, services, finance and high-tech

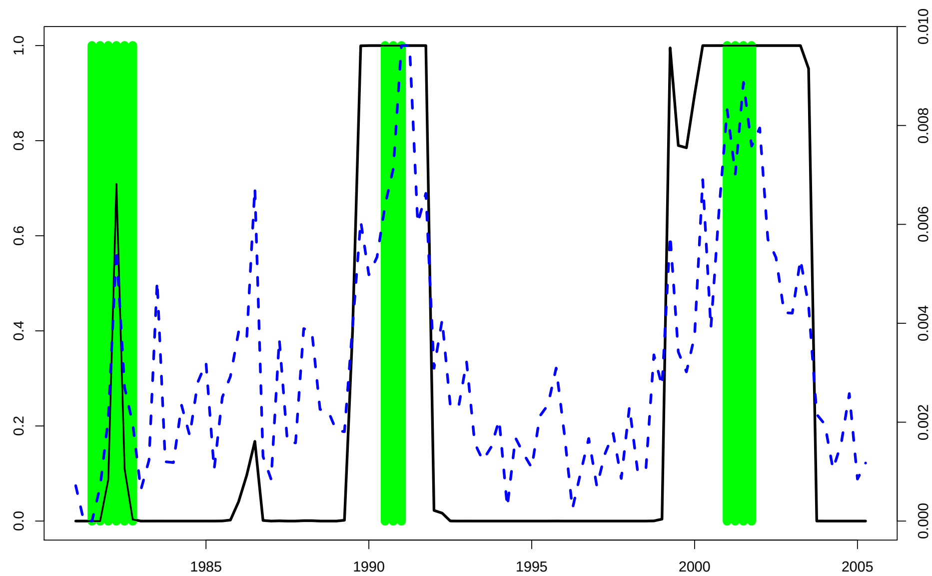

Figure 7: Realized default rate (dashed line) and smoothed crisis probability (solid line) for the entire sample vs NBER recession periods