Multiperiod Corporate Default Prediction with Partially-Conditioned Forward Intensity

Jin-Chuan Duan & Andras Fulop

@NYU Volatility Institute

Discussion by
Sanjiv R Das

A tale of 3 papers

DSW

Duffie, D., L. Saita, and K. Wang (2007). Multi-Period Corporate Default Prediction with Stochastic Covariates. *Journal of Financial Economics 83*, 635–665.

FITS

Multiperiod Corporate Default Prediction – A Forward Intensity Approach

JoEconometrics 2012

Jin-Chuan Duan, Jie Sun and Tao Wang*

PC-FITS

Multiperiod Corporate Default Prediction with the Partially-Conditioned Forward Intensity

Jin-Chuan Duan* and Andras Fulop†

DSW

A maximum likelihood approach to estimate the term structure of survival probabilities.

Related to a default stopping time τ driven by a multidimensional Markov process *X*, and an intensity that is a function of *X* and some parameters $\theta = (\beta, \Upsilon)$, some of which are observable.

Once the parameters are estimated, then we can obtain the required term structures of survival probabilities or conversely, the conditional probabilities of default.

DSW System

$$\theta = (\beta, \gamma)$$

$$\lambda_{it} = \lambda(X_t; \beta_i)$$

Firm by firm default intensities as a function of state variables, i.e., covariates X, and parameters β.

$$\phi(X_{t+1}|X_t; \gamma)$$

Transition density functions for the state variables with parameters γ.

$$\mathcal{L}(X_1, X_2, ..., X_k, \tau_1, \tau_2, ..., \tau_n; \beta, \gamma)$$

Joint likelihood of covariates and default stopping times.

DSW Decomposition

$$\mathcal{L}(X_1, X_2, ..., X_k, \tau_1, \tau_2, ..., \tau_n; \beta, \gamma)$$

$$= \mathcal{L}(\tau_1, ..., \tau_n; X_1, ..., X_k, \beta) \times \mathcal{L}(X_1, ..., X_K; \gamma)$$

Scales by *n*, *k* Scales by k

Separates the stopping time likelihood from the covariates dynamics.

Requires the doubly stochastic assumption and Bayes theorem. Note that the dynamics of the state variables are not affected by default.

DSW MLE

Given the doubly stochastic assumption the stopping times for n firms, conditional on X, becomes independent, allowing for an easy solution to the default likelihood:

$$\mathcal{L}(\tau_1, ..., \tau_n | X_t, \beta) = \prod_{i=1}^{n} e^{-\int_0^{\tau_i} \lambda_i(X_t;\beta)\, dt} \lambda_i(X_{\tau_i}; \beta)$$

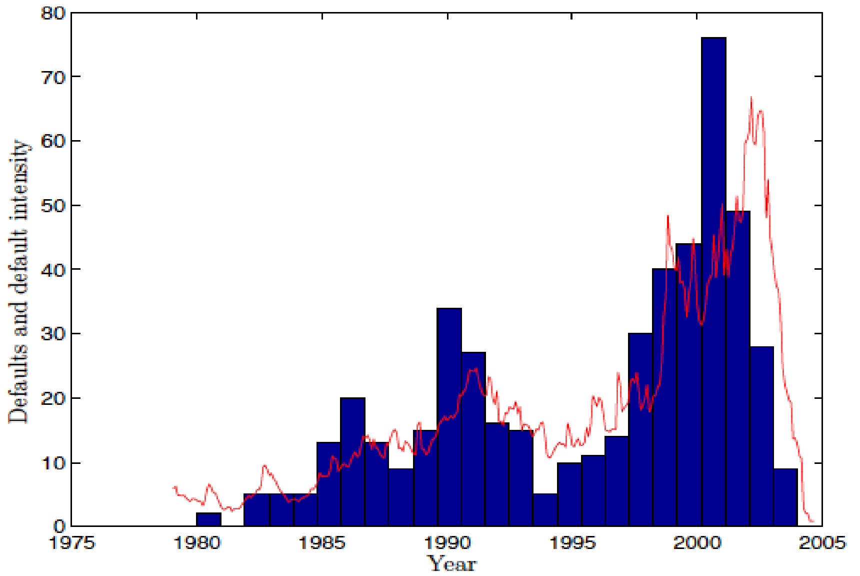
And because X is Markovian, we can simply write

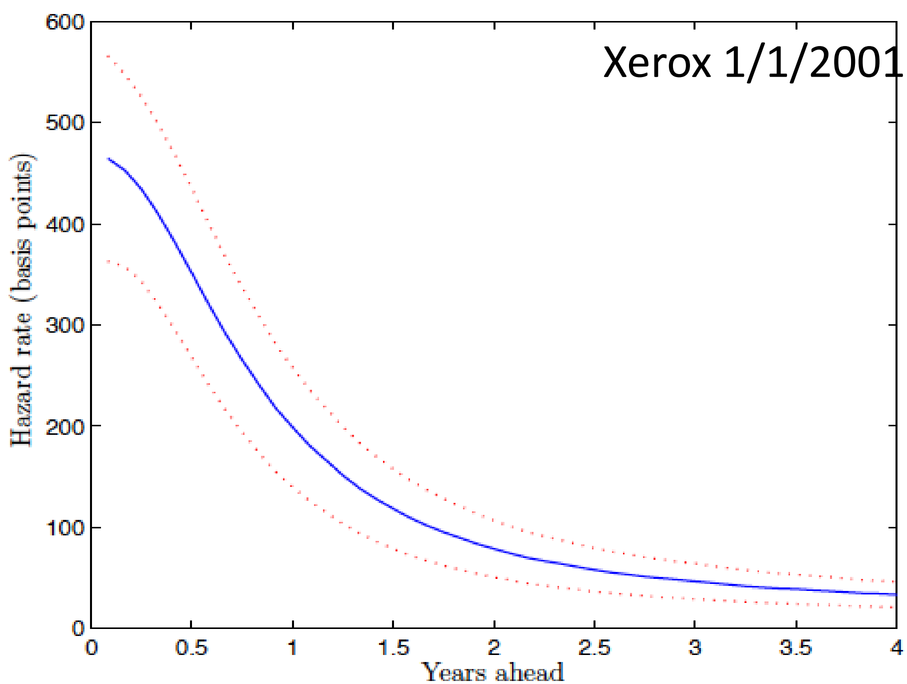
$$\mathcal{L}(X; \gamma) = \prod_{j=1}^{k} \phi(X_j | X_{j+1}; \gamma)$$

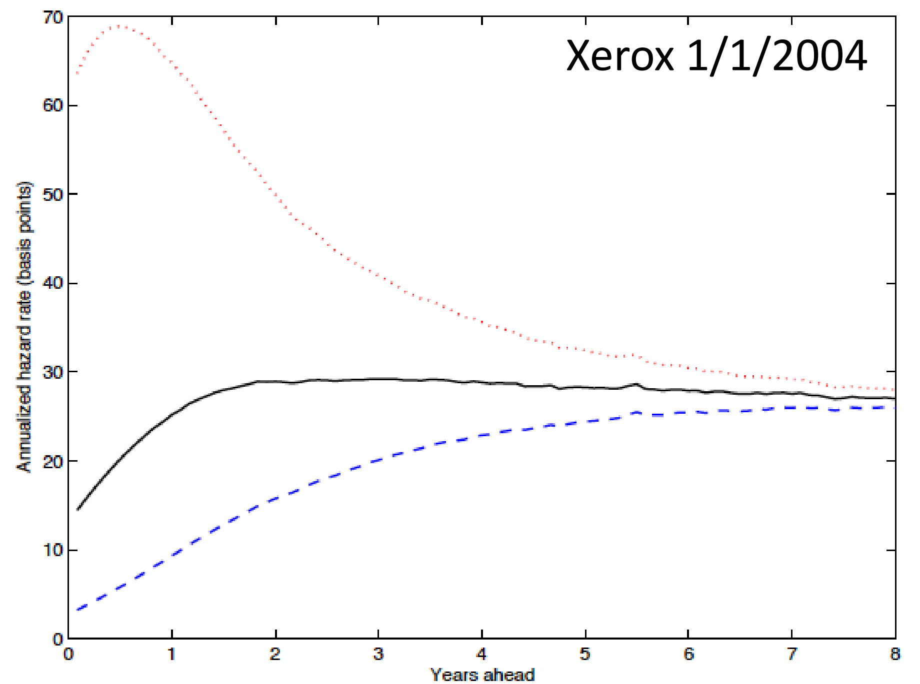
This is nested in a competing hazard framework given a firm exits for reasons other than default/bankruptcy.

DSW Results

Exit type	constant	DTD	return	3-mo. r	SPX
bankruptcy	−3.099	−1.089	−0.930	−0.153	1.074
	(0.198)	(0.062)	(0.141)	(0.037)	(0.489)
default	−2.156	−1.129	−0.694	−0.105	1.203
	(0.113)	(0.036)	(0.075)	(0.021)	(0.289)
failure	−2.148	−1.129	−0.692	−0.106	1.185
	(0.113)	(0.036)	(0.074)	(0.021)	(0.289)
merger	−3.220	0.021	0.310	−0.137	1.442
	(0.098)	(0.013)	(0.050)	(0.014)	(0.241)
other	−2.773	−0.072	0.677	−0.167	0.674
	(0.095)	(0.014)	(0.040)	(0.015)	(0.231)



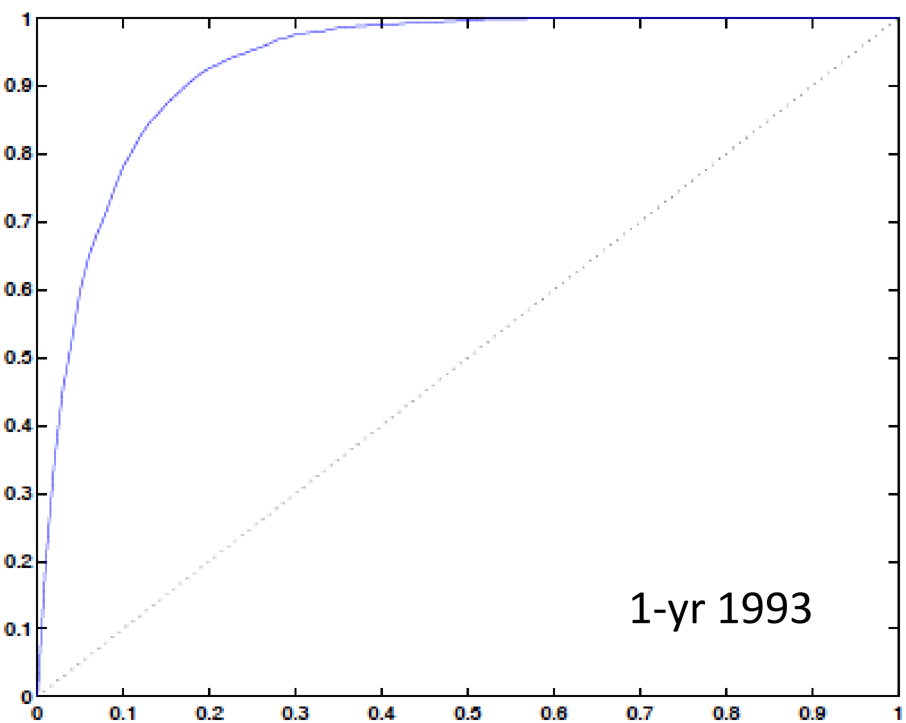


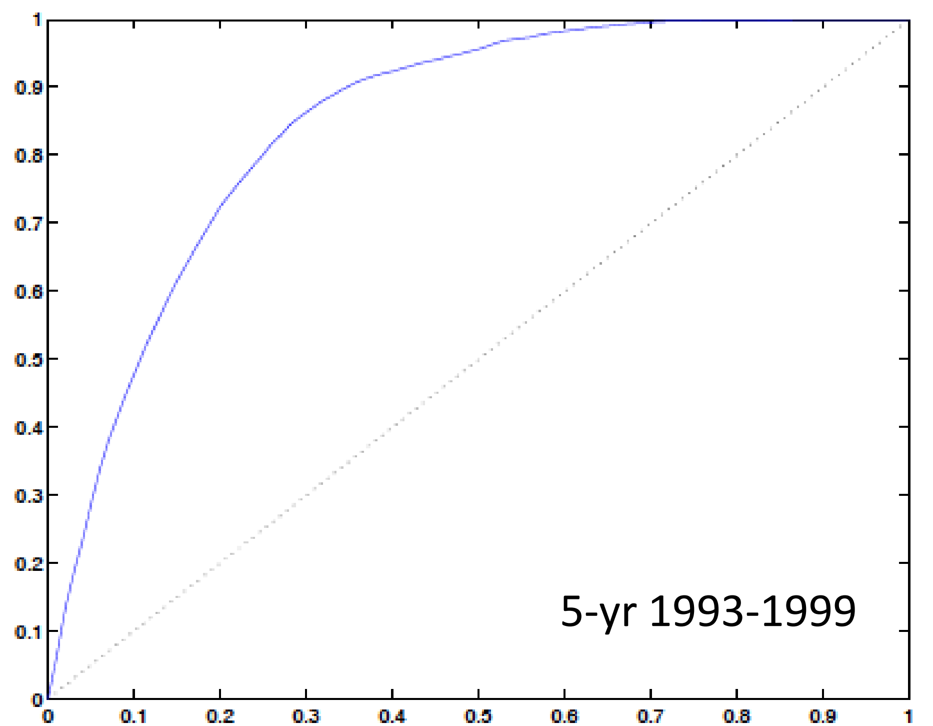


DSW Metrics

	AR1	AR5
bankruptcy	0.88	0.69
default	0.87	0.70
failure	0.86	0.70
merger	−0.03	−0.01
other exit	0.19	0.11

1-yr: 1993-2004
5-yr: 1993-2000





FITS Approach

$$\text{Current time } t, \text{ and forward horizon } \tau$$

$$F_{it}(\tau): \text{ Conditional distribution of exit time at } t + \tau$$

$$\lambda_{it}: \text{ instantaneous intensity}$$

$$\psi_{it}: \text{ average intensity over } \tau$$

$$1 - F_{it}(\tau) = e^{-\psi_{it}(\tau)\tau} = E_t\left[\exp\left(-\int_t^{t+\tau} \lambda_{is}\,ds\right)\right]$$

Integral over conditional exit times.

Forward Intensities $f_{it}(\tau)$

Forward default probability at time t for the period $[t+\tau, t+\tau+1]$:

$$P_t(t+\tau < \tau_{Di} = \tau_{Ci} \leq t+\tau+1) = e^{-\psi_{it}(\tau)\tau\Delta t}\left(1 - e^{-f_{it}(\tau)\Delta t}\right)$$

$$f_{it}(\tau) \quad = \quad \exp\left(\alpha_0(\tau) + \alpha_1(\tau)x_{it,1} + \alpha_2(\tau)x_{it,2} + \cdots \alpha_k(\tau)x_{it,k}\right)$$

DSW: when $\tau = 0$ Not really??

Cumulative default probability at time t for the period $[t, t+\tau]$:

$$P_t(t < \tau_{Di} = \tau_{Ci} \leq t+\tau) = \sum_{s=0}^{\tau-1} e^{-\psi_{it}(s)s\Delta t}\left(1 - e^{-f_{it}(s)\Delta t}\right)$$

Decomposing the Likelihood

$$\mathscr{L}(\alpha(s)) = \prod_{i=1}^{N} \prod_{t=0}^{T-s-1} \mathscr{L}_{i,t}(\alpha(s)), \quad s = 0, 1, \cdots, \tau - 1$$

where

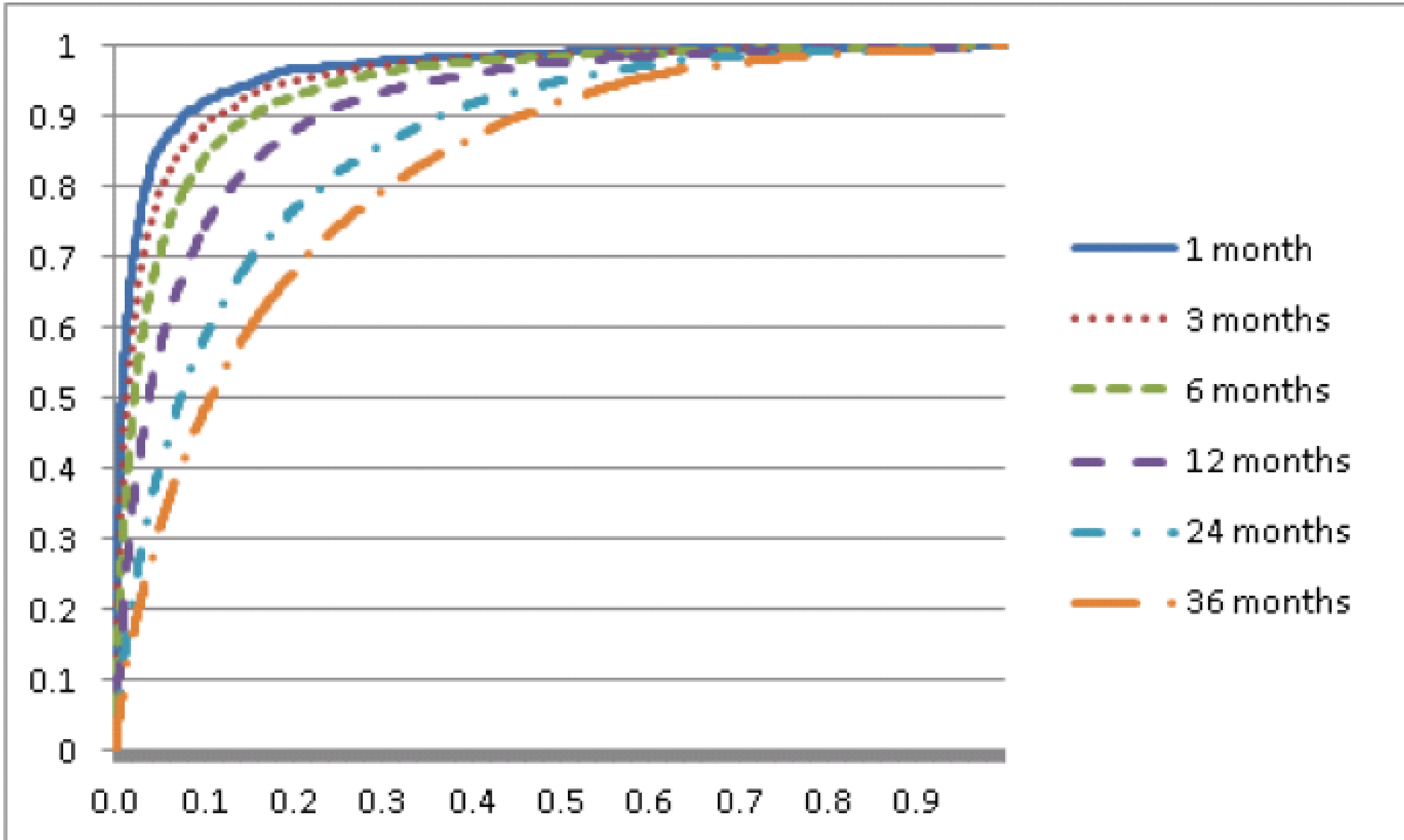
$$\begin{aligned}\mathscr{L}_{i,t}(\alpha(s)) =& 1_{\{t_{0i} \leq t, \tau_{Ci} > t+s+1\}} \exp\left[-f_{it}(s)\Delta t\right] \\ &+ 1_{\{t_{0i} \leq t, \tau_{Di} = \tau_{Ci} = t+s+1\}} \left\{1 - \exp\left[-f_{it}(s)\Delta t\right]\right\} \\ &+ 1_{\{t_{0i} \leq t, \tau_{Di} \neq \tau_{Ci}, \tau_{Ci} = t+s+1\}} \exp\left[-f_{it}(s)\Delta t\right] \\ &+ 1_{\{t_{0i} > t\}} + 1_{\{\tau_{Ci} < t+s+1\}}\end{aligned}$$

FITS (+)

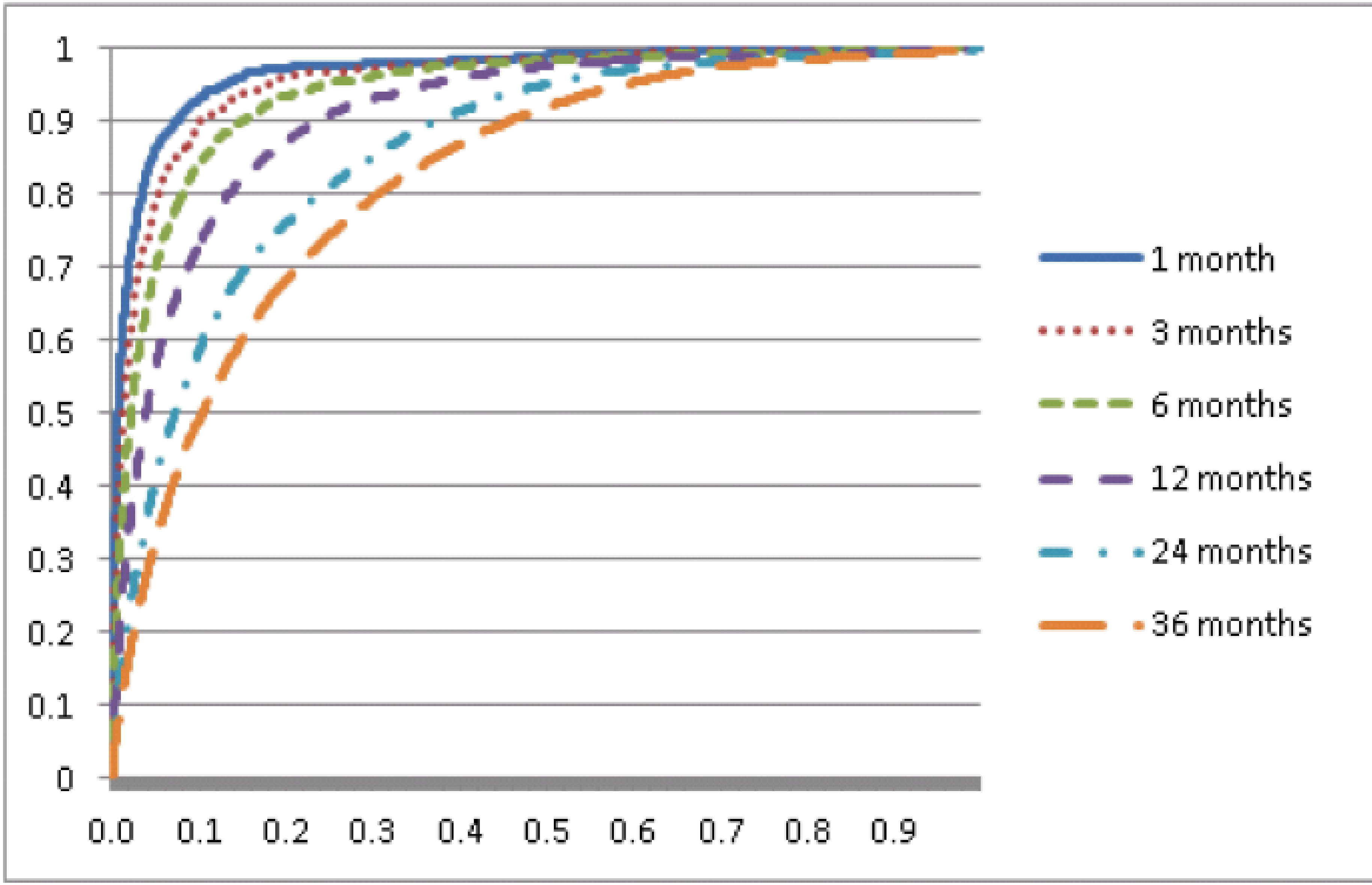
1. Does not need a choice of the specification of the stochastic process for state variables.
2. Can account for competing risks.
3. Generates term structure of PDs.
4. Based on *t*-filtration with no look-ahead bias.
5. Uses overlapping data to compute a “pseudo-MLE” estimator.
6. Decomposability of likelihood for future periods, lending itself to parallel computing.
7. Aggregation across names and time feasible (with implicit assumptions) from decomposability.
8. Fewer parameters (uses the Nelson-Siegel construction to make time-dependent functions project on a small set of parameters).

FITS: In-sample

This figure shows the in-sample cumulative accuracy profiles (power curves) based on all firms and the entire sample period (1991 to 2011) for different prediction horizons.



FITS: Out-of-sample



Accuracy Ratios: w/o partial conditioning

Panel A: In-sample result (1991-2011)						
	1 month	3 months	6 months	12 months	24 months	36 months
DSW (2007)	91.95%	90.06%	88.14%	85.37%	80.54%	77.22%
Restricted DSW	91.95%	89.96%	87.24%	81.72%	71.28%	63.85%
Forward Intensity	91.95%	89.63%	86.78%	81.43%	71.43%	64.01%
Panel B: In-sample result (2001-2011)						
	1 month	3 months	6 months	12 months	24 months	36 months
DSW (2007)	92.26%	91.08%	89.19%	86.58%	81.22%	77.58%
Restricted DSW	92.26%	91.12%	88.91%	84.58%	75.04%	68.98%
Forward Intensity	92.26%	90.85%	88.56%	84.68%	76.15%	70.39%
Panel C: Out-of-sample (over time) result (2001-2011)						
	1 month	3 months	6 months	12 months	24 months	36 months
DSW (2007)	91.97%	91.38%	87.43%	77.50%	60.33%	51.87%
Restricted DSW	91.97%	90.80%	88.44%	83.52%	71.66%	65.04%
Forward Intensity	91.97%	90.50%	88.04%	83.77%	74.67%	70.31%

Restricted DSW has same mean reversion parameter across all firms. Rolling one month and then out to end of sample. In-sample DSW does better as it uses more information, out-of-sample it does worse as it's probably over-fitting.

FITS (-)

1. Not amenable to generating conditional forward distributions of outcomes for single names, i.e., dynamics.
2. Akin to “functional” bootstrapping yield/spread curves, but variation not embedded.
3. For credit portfolios, the absence of covariates makes inducing dependence harder.

PCFITS

FITS

$$f_{it}(\tau) \;=\; \exp\left(\alpha_0(\tau) + \alpha_1(\tau)x_{it,1} + \alpha_2(\tau)x_{it,2} + \cdots \alpha_k(\tau)x_{it,k}\right)$$

PCFITS

$$\begin{aligned} f_{it}(\tau; \mathbf{Z}_{t+\tau}) \;=\; & \exp[\alpha_0(\tau) + \alpha_1(\tau)x_{it,1} + \cdots + \alpha_k(\tau)x_{it,k} + \theta_1(\tau)z_{t,1} + \cdots + \theta_m(\tau)z_{t,m} \\ & + \theta_1^*(\tau)(z_{t+\tau,1} - z_{t,1}) + \cdots + \theta_m^*(\tau)(z_{t+\tau,m} - z_{t,m})] \end{aligned}$$

The forward intensity is now conditioned on a partial set of common variables that captures dynamics through to the forward horizon. A hybrid version of FITS+DSW.

Requires a dynamic model for common factors *z* so that the distribution of *f* can be obtained. Keep the dimension of *z* low.

z can include a latent common frailty factor.

Accuracy Ratios under PCFITS

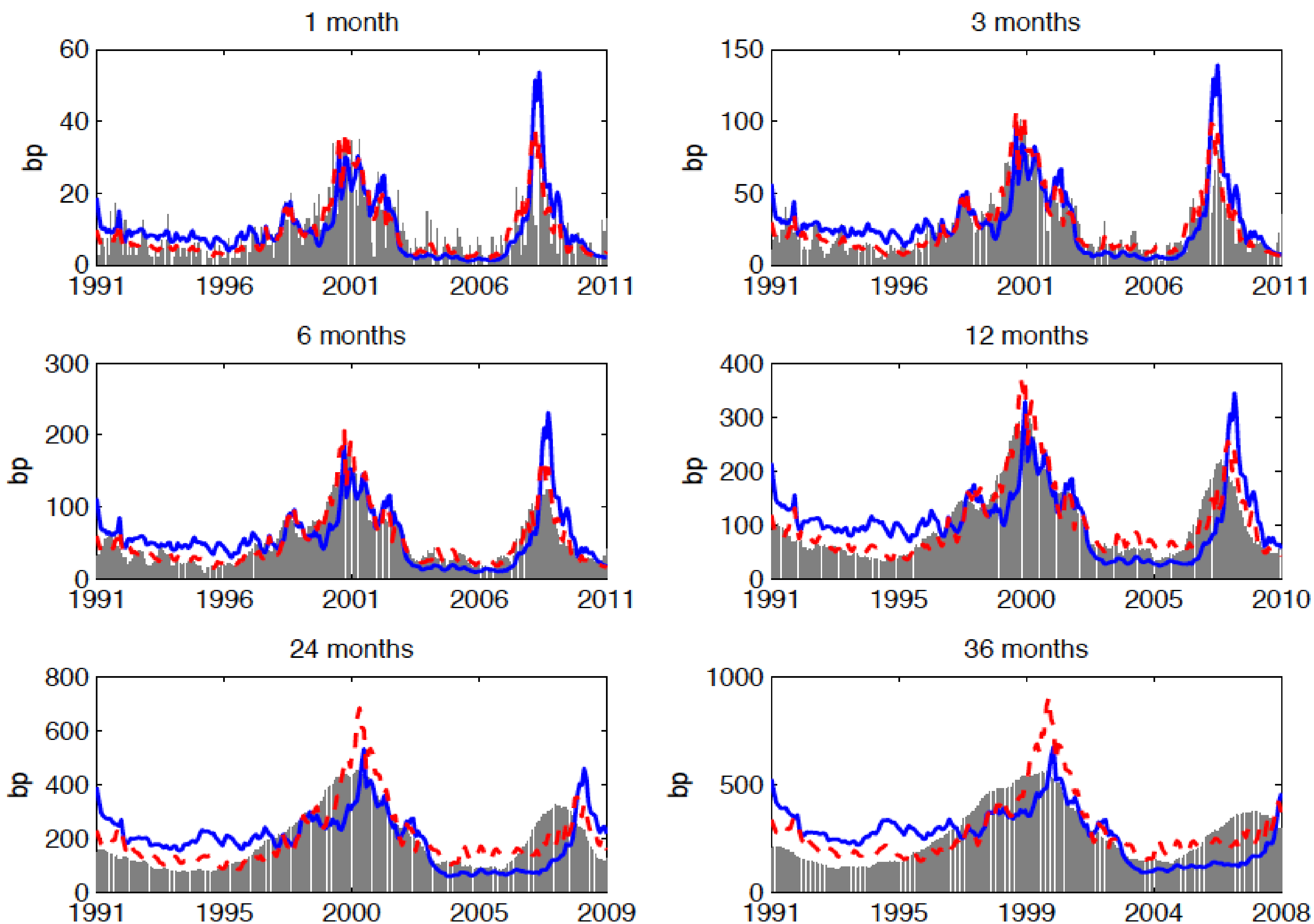
	1 month	3 months	6 months	12 months	24 months	36 months
Panel A: In-sample results for the whole sample						
DSW	93.02%	91.13%	88.49%	83.4%	73.92%	66.49%
DSW-F	93.66%	91.54%	88.84%	84.13%	75.31%	67.6%
PC-F	93.5%	91.49%	88.91%	84.29%	75.51%	68.05%
PC-M	93.48%	91.47%	88.89%	84.27%	75.45%	67.82%
Panel B: In-sample results for the non-financial subsample						
DSW	93.08%	91.1%	88.26%	82.95%	73.87%	66.76%
DSW-F	93.7%	91.53%	88.72%	83.91%	75.42%	67.78%
PC-F	93.57%	91.51%	88.8%	84.03%	75.6%	68.29%
PC-M	93.58%	91.52%	88.81%	84.04%	75.59%	68.1%
Panel C: In-sample results for the financial subsample						
DSW	92.49%	91.18%	90.29%	86.87%	73.51%	60.13%
DSW-F	93.53%	91.85%	90.34%	87.17%	76.96%	67.05%
PC-F	93.09%	91.49%	90.26%	87.36%	77.09%	66.87%
PC-M	93.08%	91.47%	90.19%	87.26%	77.05%	67.14%
Panel D: Out-of-sample (over time) results for the whole sample						
DSW	92.85%	91.31%	88.95%	85.02%	77.15%	72.26%
DSW-F	93.54%	91.93%	89.64%	85.92%	77.46%	70.27%
PC-F	93.37%	91.87%	89.7%	86.2%	78.63%	72.04%
PC-M	93.46%	91.95%	89.75%	86.19%	78.52%	71.61%

DSW – old model, DSW-F: only additional factor is the frailty but only at time t. PC-F : frailty at t and beyond; PC-M: macro factor at t and beyond.

(a) Out of sample better?

(b) (b) Not sure why the PC versions don't over fit like we see in DSW (2007)?

Figure 1: Aggregate default rate predictions of the DSW and PC-F models



This figure presents the predicted default rates for two models: DSW (solid blue curve) and PC-F (dashed red curve). Realized default rates are the gray bars.

Significantly different?

Final comments …

- Stability of parameters across time? Report the estimates month to month given there is a rolling procedure in place.
- Confusion matrix.
- Ratings predictions.
- Blochlinger (JFQA 2012): tests for both "discrimination" (rank ordering of default) and "calibration"(expected default).
- Choose a fatter tailed innovation for the frailty process.
- Can the approach be modified to extract the forward recovery rate?
- CVA? Estimate the dependence on exposure.