Web Appendix to Global Credit Risk: World, Country and Industry Factors

Bernd Schwaab,[(a)] *Siem Jan Koopman,*[(b,c)] *André Lucas*[(b)]

[(a)] European Central Bank, Financial Research

[(b)] VU University Amsterdam and Tinbergen Institute

[(c)] CREATES, Aarhus University

July 31, 2015

Appendix A: International default data

Figure A.1 visualizes and tabulates the main regions and countries we distinguish in the main paper. We take into account data from 16,360 rated firms in the U.S., 903 firms in the U.K., 2087 firms in euro area countries, and 1517 firms in the Asia-Pacific region. In total, we consider 20,867 firms, worldwide. The corresponding number of defaults are 1660, 64, 106, and 72, respectively, totaling 1902 default events. We focus on 35 years of quarterly data from 1980Q1 to 2014Q4.

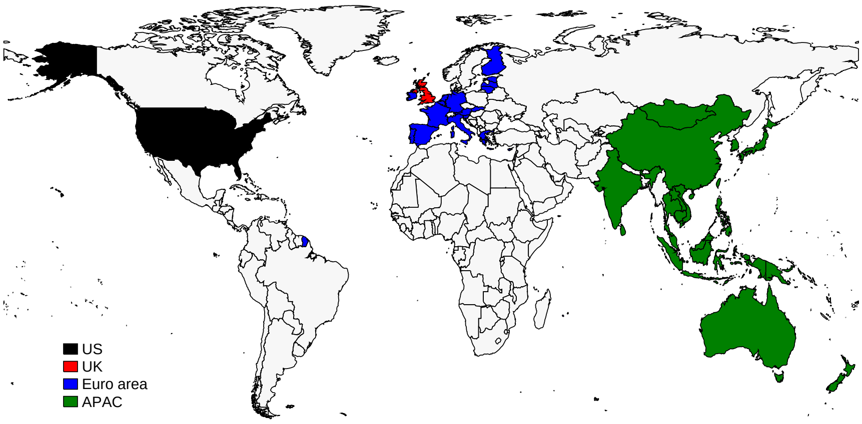
We apply several standard filters when counting default events and firms at risk: We consider only the first default event when there are multiple defaults for the same firm. We exclude firms that are in the database for less than 100 days. We also exclude firms that enter the database with a default as a first event. If a firm defaults, we ignore a previous rating withdrawal.

Our country selection and grouping in Figure A.1 is in part motivated by the availability of EDF data to augment the count data from Moody's extensive default and recovery database (DRD). Other groupings are possible as well.[1]

Figure A.2 plots the international default data in two ways. The top panel reports the total number of defaults, the total number of firms at risk, and the respective default fractions. The bottom four panels present aggregate default and firm counts, as well as observed default fractions over time for each economic region. Most defaults are centered around a few global recession periods in each region. The highest default fractions are observed approximately around (U.S.) recession years such as 1990-1991, 2001-2002, and 2007-2009. Exceptions exist: there are a substantial number of defaults in the euro area during the most acute phase of the sovereign debt crisis from 2010-2013.

[1] For example, it would in principle be possible to group Canada with the U.S. into a "North America" region. We do not do so because we don't have EDF data on Canadian firms. We do not consider countries from, say, Latin America for the same reason.

Figure A.1: Country sample and grouping



Region 1: U.S.
U.S.A.
U.S. Territories

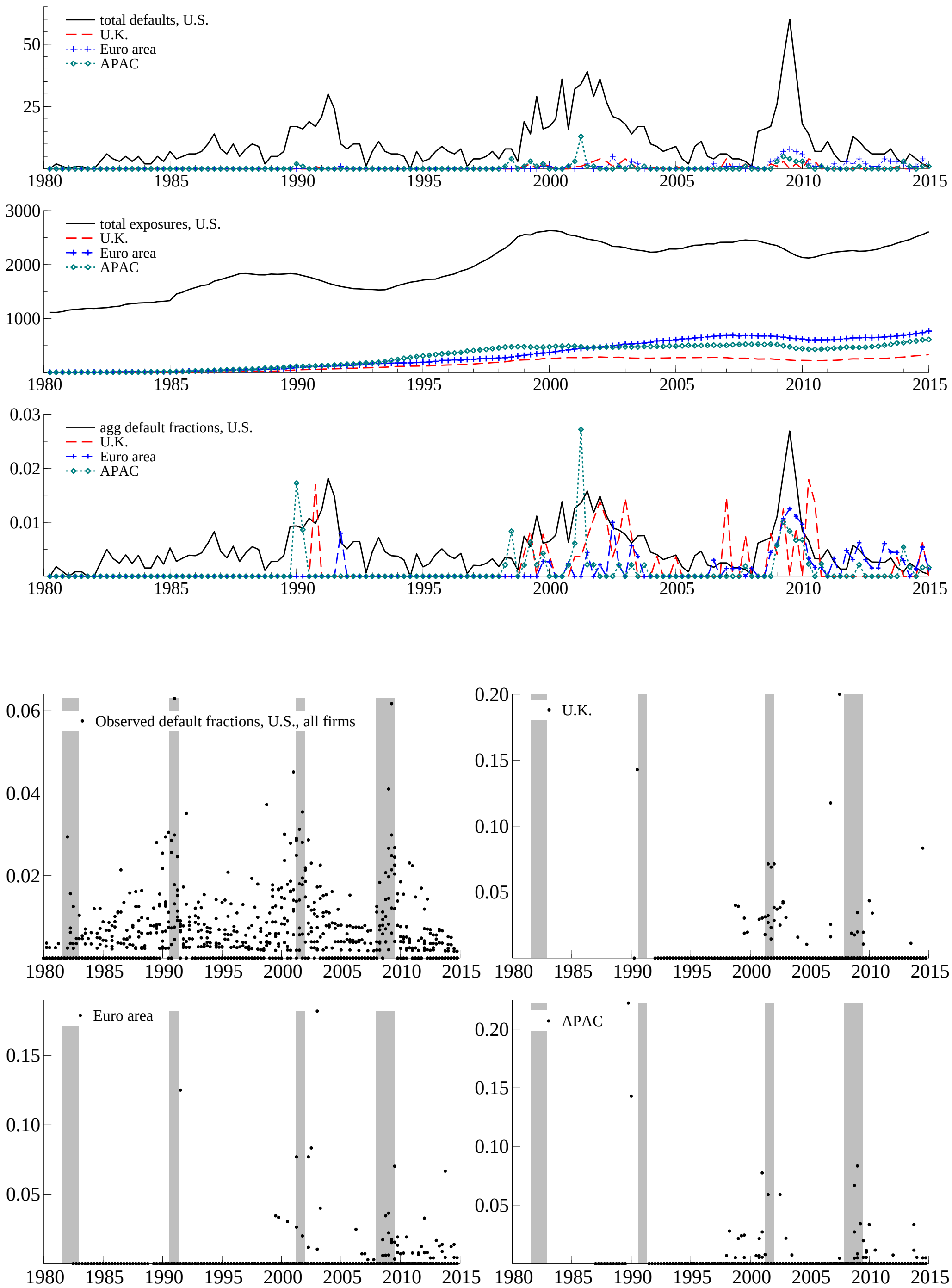
Region 2: U.K.
United Kingdom
British Virgin Islands
Isle of Man

Region 3: Euro area	
Austria	Latvia
Belgium	Lithuania
Cyprus	Luxembourg
Estonia	Malta
Finland	Netherlands
France	Portugal
Germany	Slovakia
Greece	Slovenia
Ireland	Spain
Italy	

Region 4: Asia-Pacific	
Australia	Mongolia
Cambodia	New Zealand
China	Papua New Guinea
Hong Kong	Philippines
India	Singapore
Indonesia	South Korea
Japan	Sri Lanka
Laos	Taiwan
Macau	Thailand
Malaysia	Vietnam

Figure A.2: Historical default and firm counts

The first panel plots time series data of the total default counts $\sum_j y_{r,j,t}$ aggregated to a univariate series (top), the total number of firms at risk $\sum_j k_{r,j,t}$ (middle), as well as the aggregate default fractions $\sum_j y_{r,j,t}$ / $\sum_j k_{r,j,t}$ over time (bottom). The second panel plots observed default fractions at the industry level for four different economic regions, distinguishing firms from the U.S., the United Kingdom, the euro area, and the Asia-Pacific region. Light-shaded areas are NBER recession times for the U.S. for reference purposes only.



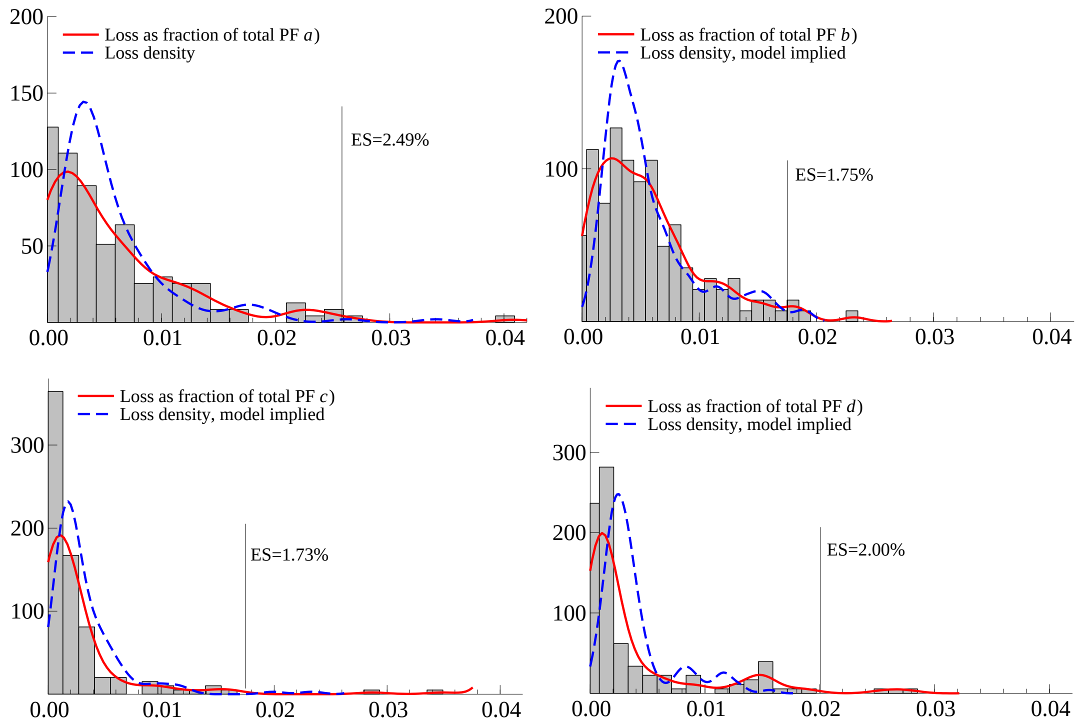
Appendix B: Observed default clustering in international credit portfolios

This section motivates our study of global default risk factors by providing a preliminary study of the benefits and limits of credit risk diversification across national borders. Figures A.2 and the EDF data suggest that world factors are an important determinant of national default rates. Hence we expect that this feature is also reflected by the historical default experience of diversified credit portfolios. We explore this intuition by studying the risk of successively more diversified portfolios while holding marginal risks (ratings) constant.

We focus on four credit portfolios. For simplicity, each portfolio is of equal size (1000$), and all loans have a maturity of one quarter. Loans are extended to firms that are active at the beginning of each quarter during our sample. At all times, loans are split equally between investment and speculative grade firms. The rating profiles of firms, as approximate measures of firm-specific default risk, are constant and the same across portfolios. As a result, differences in portfolio risk are only due to changes in the default *dependence*, or systematic default risk, across firms. Portfolio *a*) is an industry and region concentrated loan portfolio, containing 1000 loans of 1$ each to 1000 U.S. industrial firms. Portfolio *b*) is region-specific, but diversified over industries and contains 1000 loans of 1$ each to U.S. firms from five non-financial industries (transportation, industrials, technology, retail & distribution, consumer goods). Each industry has an exposure of 200 loans. Portfolio *c*) is industry-specific, but diversified over regions and contains 1000 loans of 1$ each to industrial firms located in each of the four regions (U.S., U.K., the euro area, and Asia-Pacific). Each region has an exposure of 250 firms. Finally, portfolio *d*) is diversified both across industries and regions. It contains 50 loans of 1$ each for each combination of the four regions and five non-financial industries in our sample.

Figure B.1 plots the quarterly historical loss rates for portfolios *a*) to *d*). An estimate of the 95% expected shortfall (ES) is indicated in each panel of Figure B.1 as a vertical line. Table B.1 reports the respective portfolio risk measures for the four loan portfolios. Overall, more credit risk diversification across national borders does not necessarily decrease portfolio

Figure B.1: Unconditional portfolio loss densities $a)$ to $d)$
The figure plots the histogram of quarterly default losses for portfolios $a)$ to $d)$ as discussed in the main text. The red line is a (Epanechnikov) kernel density estimator of the loss rate distribution. The blue dashed line indicates the unconditional loss distribution as implied by the full model estimated in the main paper. The 95% expected shortfall is indicated in each panel as a vertical line.



risk.

Based on the 95% ES, successively reducing portfolio concentration moves the quarterly portfolio risk from 2.47% to 1.75% to 1.73%, and then to 2.00%. The pattern is analogous when portfolio risk is alternatively measured as the 95% Value-at-Risk (VaR) or as the mean portfolio loss. The raw data therefore suggests that, in the presence of global factors, cross-border risk diversification is not necessarily beneficial if the initial portfolio is already somewhat diversified across industries.[2]

When global factors are the main source of international default clustering, two effects work in opposite directions. First, if country-specific (regional) macro-financial and other

[2] As a caveat, portfolio risk measures such as ES and VaR are subject to substantial estimation uncertainty, see (McNeil, Frey, and Embrechts 2005), and the risk measures in Table B.1 are not necessarily statistically different from each other. In addition, the historical default experience may be different for different areas, even conditional on rating groups.

Table B.1: **Credit portfolio risk measures**

The table reports portfolio risk measures for losses from four portfolios *a*) to *d*) as discussed in the main text. The risk measures refer to quarterly credit losses due to default, assuming a loss-given-default of 100%. The Value-at-Risk at 95% is the 95% quantile of the quarterly observed losses from 1980Q1 to 2014Q4 (140 quarters). The expected shortfall at 95% is approximated as the unweighted average over the 95% to 99% empirical quantiles of the observed quarterly losses.

Observed losses	mean loss	Var 95%	ES 95%
PF1 [most concentrated]	0.58%	1.74%	2.47%
PF2 [country concentration]	0.57%	1.45%	1.75%
PF3 [industry concentration]	0.27%	0.97%	1.73%
PF4 [fully diversified]	0.39%	1.51%	2.00%

factors are imperfectly correlated, then expanding the credit portfolio across borders *decreases* dependence across firms, and therefore decreases portfolio credit risk. Conversely, portfolio diversification across borders can *increase* dependence if it leads to risk exposures that load relatively more heavily on the global factors, such as, for example, global macro and global default-specific factors. Figure B.1 suggests that the first effect appears to dominate when moving from portfolio *a*) to portfolios *b*), *c*), and *d*). The second effect appears to dominate when moving from portfolios *b*) and *c*) to portfolio *d*). The The modeling framework in the main paper is an appropriate tool to quantify the portfolio risk of different collections of loans, and to determine the incremental (systematic) risk contribution of additional loan exposures.

Appendix C: Simulated maximum likelihood estimation

We consider available data as in (13),

$$\mathcal{Y}_t = \left(\hat{F}'_{g,t}, \hat{F}'_{m,t}, y'_t, \hat{\theta}_t^{\text{EDF}}\right)', \quad t = 1, \dots, T,$$

where each row $\mathcal{Y}_k = (\mathcal{Y}_{k1}, \dots, \mathcal{Y}_{kT})$, $k = 1, \dots, K$, relates to time series data from a potentially different family of parametric distributions. All observations depend on a set of m dynamic factors which we collect into the $m \times 1$ vector f_t and assume a stationary vector autoregressive process, $f_{t+1} = \Phi f_t + \eta_t$ with $\eta_t \sim \text{N}(0, \Sigma_\eta)$, for $t = 1, \dots, T$, and initial condition $f_1 \sim \text{N}(0, \Sigma_f)$.

Conditional on a factor path $\mathcal{F}_t = \{ f_1, f_2, \dots, f_t\}$, the observation $\mathcal{Y}_{k,t}$ of the kth variable at time t is assumed to come from a certain density given by

$$\mathcal{Y}_{k,t} \mid \mathcal{F}_t \sim p_k\big(\mathcal{Y}_{k,t}; \theta_{k,t}, \psi\big), \quad k = 1, \dots, K, \tag{C.1}$$

with a time-varying parameter defined by $\theta_{k,t} = \lambda_k + \vartheta'_k f_t$, where λ_k is an unknown constant and ϑ_k is a loading vector with unknown coefficients. The conditional (on $\mathcal{F}_t$) independence of all observations implies that the density of the $K \times 1$ observation vector $\mathcal{Y}_t = (\mathcal{Y}_{1,t}, \dots, \mathcal{Y}_{K,t})'$ is given by

$$p(\mathcal{Y}_t \mid \mathcal{F}_t, \psi) = \prod_{k=1}^{K} p_k\big(\mathcal{Y}_{k,t} \mid \mathcal{F}_t, \psi\big).$$

The parameter vector ψ contains all unknown coefficients in the model specification including those in Φ, λ_k and ϑ_k for $k = 1, \dots, K$. To enable the identification of all entries in ψ, we assume standardized factors which we enforce by the restriction $\Sigma_f = \text{I}$ in the expression for the unconditional covariance matrix $\Sigma_f = \Phi \Sigma_f \Phi' + \Sigma_\eta$, implying that $\Sigma_\eta = \text{I} - \Phi\Phi'$.

The estimation of the parameter vector ψ and risk factors f_t via maximum likelihood is non-standard because an analytical expression for the maximum likelihood (ML) estimate of parameter vector ψ for the MM-DFM is not available. Let $\mathcal{Y} = (\mathcal{Y}'_1, \dots, \mathcal{Y}'_T)'$ and $f = (f'_1, \dots, f'_T)'$ denote the vector of all the observations and factors, respectively. Let $p(\mathcal{Y} \mid f; \psi)$ be the density of $\mathcal{Y}$ conditional on f and let $p(f; \psi)$ be the density of f. The log-likelihood function is only available in the form of an integral

$$p(\mathcal{Y}; \psi) = \int p(\mathcal{Y}, f; \psi)\, \text{d}f = \int p(\mathcal{Y} \mid f; \psi) p(f; \psi)\, \text{d}f, \tag{C.2}$$

where f is integrated out. A feasible approach to computing this integral is provided by importance sampling; see, e.g. (Kloek and van Dijk 1978), (Geweke 1989) and (Durbin and Koopman 2012). Upon computing the integral, the maximum likelihood estimator of ψ is obtained by direct maximization of the likelihood function using Newton-Raphson methods.

Inference on the latent factors can also be based on importance sampling. In particular, it can be shown that

$$\mathrm{E}(f|\mathcal{Y};\psi) = \int f \cdot p(f|\mathcal{Y};\psi)\mathrm{d}f = \frac{\int f \cdot w(\mathcal{Y}, f;\psi) g(f|\mathcal{Y};\psi)\mathrm{d}f}{\int w(\mathcal{Y}, f;\psi) g(f|\mathcal{Y};\psi)\mathrm{d}f},$$

where $w(\mathcal{Y}, f;\psi) = p(\mathcal{Y}|f;\psi)/g(\mathcal{Y}|f;\psi)$ is the importance sampling weight. The estimation of $\mathrm{E}(f|\mathcal{Y};\psi)$ via importance sampling can be achieved by

$$\tilde{f} = \sum_{k=1}^{M} w_k \cdot f^{(k)} \Bigg/ \sum_{k=1}^{M} w_k,$$

with $w_k = p(\mathcal{Y}|f^{(k)};\psi)/g(\mathcal{Y}|f^{(k)};\psi)$ and where $f^{(k)} \sim g(f|\mathcal{Y};\psi)$ is obtained by simulation smoothing. The standard error of $\tilde{f}_i$, the ith element of $\tilde{f}$, is denoted by s_i and is computed by

$$s_i^2 = \left(\sum_{k=1}^{M} w_k \cdot (f_i^{(k)})^2 \Bigg/ \sum_{k=1}^{M} w_k \right) - \tilde{f}_i^2,$$

where $f_i^{(k)}$ is the ith element of $f^{(k)}$.

Appendix D: Subsample analysis 2000Q1 – 2014Q4

This section reports parameter estimates and risk decomposition results when quarterly mixed-measurement data (13) between 2000Q1 – 2014Q4 is used for parameter and risk factor estimation. Considering a shorter sample decreases the chance of a structural break in time-invariant loading parameters, while preserving most of the default and exposure data from non-U.S. countries.

Table D.1 reports the new model parameter estimates. The new parameter estimates are similar to the full sample estimates in the main paper, based on data between 1980Q1 – 2014Q4. All sets of risk factors – macro, frailty, as well as industry-specific – continue to contribute towards explaining corporate default clustering within and across countries.

The main differences with respect to the full sample estimates are that (i) the industry-specific factors become hard to estimate; loading parameters are more uncertain. Defaults for firms in the energy & transportation sector are particularly rare in the reduced sample. (ii) There is even less evidence for a regional frailty factor for the U.K. Instead, U.K. firms load on the global (U.S.) frailty factor. (iii) Then statistical significance of most parameter estimates decreases, leading to higher p-values.

Table D.2 presents the new estimated risk shares. The risk shares from the reduced sample are approximately similar to the full sample estimates in the main paper, except that the industry-specific variation in the energy & transportation sector appears to become less important for the 2000Q1 – 2014Q4 sample. Conversely, the industry-specific variation for financial sector firms becomes slightly more pronounced.

Table D.1: **Parameter estimates**

We report the maximum likelihood estimates of selected coefficients in the specification of the log-odds ratio θ_{jt} in (5). We use an additive parametrization for $\lambda_{r,j}$ and $\alpha_{r,j}$. Coefficients $\lambda_{r,j}$ determine baseline default rates. Factor loadings refer to global macro factors f_t^g, region-specific macro factors f_t^m, one global frailty factor that is common to all firms f_t^c, country/region-specific default-specific (frailty) factors f_t^d, and six industry-specific factors f_t^i. The global macro factors are common to all macro and default data and across all four regions. The global and regional frailty factors do not load on macro data. Industry mnemonics are financials (fin), transportation & energy (tre), industrials (ind), technology (tec), retail & distribution (red), and consumer goods (con). **Estimation sample is 2000Q1 to 2014Q4.**

Baseline hazard terms

$\lambda_{r,j} = \bar{\lambda}_0 + \bar{\lambda}_{1,j} + \bar{\lambda}_{2,s} + \bar{\lambda}_{3,r}$

par	val	p-val
$\bar{\lambda}_0$	-4.69	0.00
$\bar{\lambda}_{1,fin}$	0.11	0.74
$\bar{\lambda}_{1,tre}$	-0.32	0.39
$\bar{\lambda}_{1,tec}$	-0.29	0.56
$\bar{\lambda}_{1,ret}$	0.09	0.83
$\bar{\lambda}_{1,con}$	-0.13	0.80
$\bar{\lambda}_{2,IG}$	-3.30	0.00
$\bar{\lambda}_{3,UK}$	0.14	0.36
$\bar{\lambda}_{3,EA}$	-0.27	0.02
$\bar{\lambda}_{3,AP}$	-0.39	0.01

Global macro f_t^g (ctd)

$\alpha_{k,r,j} = \bar{\alpha}_{k,0} + \bar{\alpha}_{k,1,r}$

par	val	p-val
ϕ_4^g	0.76	0.00
$\bar{\alpha}_{4,0}$	0.08	0.05
$\bar{\alpha}_{4,1,UK}$	-0.08	0.02
$\bar{\alpha}_{4,1,EA}$	-0.03	0.35
$\bar{\alpha}_{4,1,AP}$	-0.03	0.60
ϕ_5^g	0.81	0.00
$\bar{\alpha}_{5,0}$	-0.07	0.16
$\bar{\alpha}_{5,1,UK}$	0.11	0.01
$\bar{\alpha}_{5,1,EA}$	0.08	0.06
$\bar{\alpha}_{5,1,AP}$	0.05	0.44

Global frailty f_t^c

par	val	p-val
ϕ^c	0.95	0.00
$\bar{\gamma}_0$	0.47	0.01
$\bar{\gamma}_{1,UK}$	0.05	0.47
$\bar{\gamma}_{1,EA}$	0.11	0.16
$\bar{\gamma}_{1,AP}$	-0.17	0.13

Regional frailty f_t^d

par	val	p-val
ϕ_{US}^d	0.94	0.00
$\bar{\delta}_{0,US}$	0.34	0.01
ϕ_{UK}^d	0.96	0.64
$\bar{\delta}_{0,UK}$	0.00	0.98
ϕ_{EA}^d	0.98	0.00
$\bar{\delta}_{0,EA}$	0.38	0.10
ϕ_{AP}^d	0.90	0.00
$\bar{\delta}_{0,AP}$	0.34	0.00

Global macro f_t^g

$\alpha_{k,r,j} = \bar{\alpha}_{k,0} + \bar{\alpha}_{k,1,r}$

par	val	p-val
ϕ_1^g	0.90	0.00
$\bar{\alpha}_{1,0}$	0.25	0.00
$\bar{\alpha}_{1,1,UK}$	0.06	0.21
$\bar{\alpha}_{1,1,EA}$	-0.01	0.87
$\bar{\alpha}_{1,1,AP}$	0.04	0.54
ϕ_2^g	0.86	0.00
$\bar{\alpha}_{2,0}$	-0.10	0.08
$\bar{\alpha}_{2,1,UK}$	-0.06	0.22
$\bar{\alpha}_{2,1,EA}$	-0.09	0.07
$\bar{\alpha}_{2,1,AP}$	0.02	0.79
ϕ_3^g	0.84	0.00
$\bar{\alpha}_{3,0}$	-0.04	0.53
$\bar{\alpha}_{3,1,UK}$	0.12	0.01
$\bar{\alpha}_{3,1,EA}$	0.09	0.07
$\bar{\alpha}_{3,1,AP}$	0.09	0.17

Regional macros f_t^m

par	val	p-val
ϕ_{US}^m	0.65	0.00
$\bar{\beta}_{0,US}$	-0.02	0.41
ϕ_{UK}^m	0.82	0.00
$\bar{\beta}_{0,UK}$	-0.05	0.23
ϕ_{EA}^m	0.66	0.00
$\bar{\beta}_{0,EA}$	0.04	0.11
ϕ_{AP}^m	0.78	0.00
$\bar{\beta}_{0,AP}$	-0.08	0.11

Industry factors f_t^i

par	val	p-val
ϕ_{fin}^i	0.97	0.00
$\bar{\epsilon}_{fin}$	0.58	0.01
ϕ_{tre}^i	0.60	0.00
$\bar{\epsilon}_{tre}$	0.57	0.61
ϕ_{ind}^i	0.91	0.00
$\bar{\epsilon}_{ind}$	0.77	0.01
ϕ_{tec}^i	0.92	0.00
$\bar{\epsilon}_{tec}$	0.80	0.18
ϕ_{ret}^i	0.88	0.00
$\bar{\epsilon}_{ret}$	0.54	0.05
ϕ_{con}^i	0.89	0.00
$\bar{\epsilon}_{con}$	0.85	0.10

Table D.2: **Systematic risk and risk decomposition**

We report systematic risk variation estimates for six industry sectors across four economic regions. Systematic default risk is further decomposed into variation due to subsets of systematic risk drivers. For factor mnemonics see Table D.1. Risk shares refer to global and region-specific macro factors f_t^g and f_t^m, one global frailty factor that is common to all firms f_t^c, country/region-specific default-specific (frailty) factors f_t^d, and six industry-specific factors f_t^i. Industry sectors are financials (fin), transportation and energy (tre), industrial firms (ind), technology (tec), retail and distribution (red), and consumer goods (con). We refer to the financial framework in Section 3 of the main paper for a discussion of firm's systematic versus idiosyncratic risk components. **Estimation sample is 2000Q1 to 2014Q4.**

Reg.	Ind.	f_t^g, f_t^m $[a_i'a_i]$	f_t^c $[c_i'c_i]$	f_t^d $[d_i'd_i]$	f_t^i $[e_i'e_i]$	$\text{Var}[V_{it}\|\varepsilon_{it}, f_t^i]$ $a_i'a_i + \ldots + d_i'd_i$	$\text{Var}[V_{it}\|\varepsilon_{it}]$ $= w_i'w_i$
US	fin	4.9%	12.7%	6.7%	18.9%	24.2%	43.1%
US	tre	4.9%	12.7%	6.7%	18.7%	24.3%	43.0%
US	ind	4.3%	11.1%	5.9%	29.2%	21.2%	50.3%
US	tec	4.1%	10.7%	5.7%	31.2%	20.6%	51.8%
US	red	5.0%	13.0%	6.9%	16.8%	24.9%	41.6%
US	con	4.0%	10.3%	5.5%	33.7%	19.8%	53.5%
UK	fin	7.5%	15.8%	0.0%	19.1%	23.3%	42.4%
UK	tre	7.6%	15.8%	0.0%	18.9%	23.4%	42.3%
UK	ind	6.6%	13.8%	0.0%	29.5%	20.3%	49.8%
UK	tec	6.4%	13.4%	0.0%	31.5%	19.8%	51.3%
UK	red	7.7%	16.2%	0.0%	17.0%	24.0%	40.9%
UK	con	6.1%	12.9%	0.0%	34.0%	19.0%	53.0%
EA	fin	5.3%	17.6%	7.4%	17.4%	30.3%	47.7%
EA	tre	5.3%	17.7%	7.4%	17.2%	30.4%	47.6%
EA	ind	4.6%	15.5%	6.5%	27.1%	26.7%	53.8%
EA	tec	4.5%	15.1%	6.4%	29.1%	26.0%	55.1%
EA	red	5.4%	18.0%	7.6%	15.4%	31.0%	46.4%
EA	con	4.4%	14.6%	6.2%	31.5%	25.1%	56.6%
AP	fin	6.5%	5.6%	7.1%	20.1%	19.2%	39.3%
AP	tre	6.5%	5.7%	7.1%	20.0%	19.2%	39.2%
AP	ind	5.6%	4.9%	6.1%	30.9%	16.6%	47.5%
AP	tec	5.4%	4.7%	5.9%	33.0%	16.1%	49.1%
AP	red	6.6%	5.8%	7.3%	17.9%	19.7%	37.6%
AP	con	5.2%	4.6%	5.7%	35.5%	15.5%	51.0%

Appendix E: Sensitivity analysis

In our parameter-driven dynamic factor model, both the information content from continuous EDF data (via $\hat{\theta}_t^{\text{EDF}}$), as well as the integer default counts (via the Binomial specification), contribute to empirically identifying the time variation in the log-odds $\theta_{j,t}$ and in the default probabilities π_{jt}. This section studies to what extent our estimates are sensitive to the information content in either set of input data. Does the EDF data dominate the risk factor estimates, or are these determined more by the Binomial part of the model?

To investigate this issue we vary the amount of information available in the default count data. EDF and macro data remain unchanged. In a first experiment, we multiply our default and exposure counts (y_{jt}, k_{jt}) with draws from a uniform distribution $u_{jt} \sim \text{iid U}[0, 0.5]$, and then round to the nearest integer. In expectation, this cuts the count data available for estimation by 75%, while keeping the observed default fractions (y_{jt}/k_{jt}) at the same value as in the original data. In a second experiment, we multiply the same u_{jt}'s by 400. As a result, the default and exposure counts in the second dataset are much more informative about the prevailing point-in-time default risks.

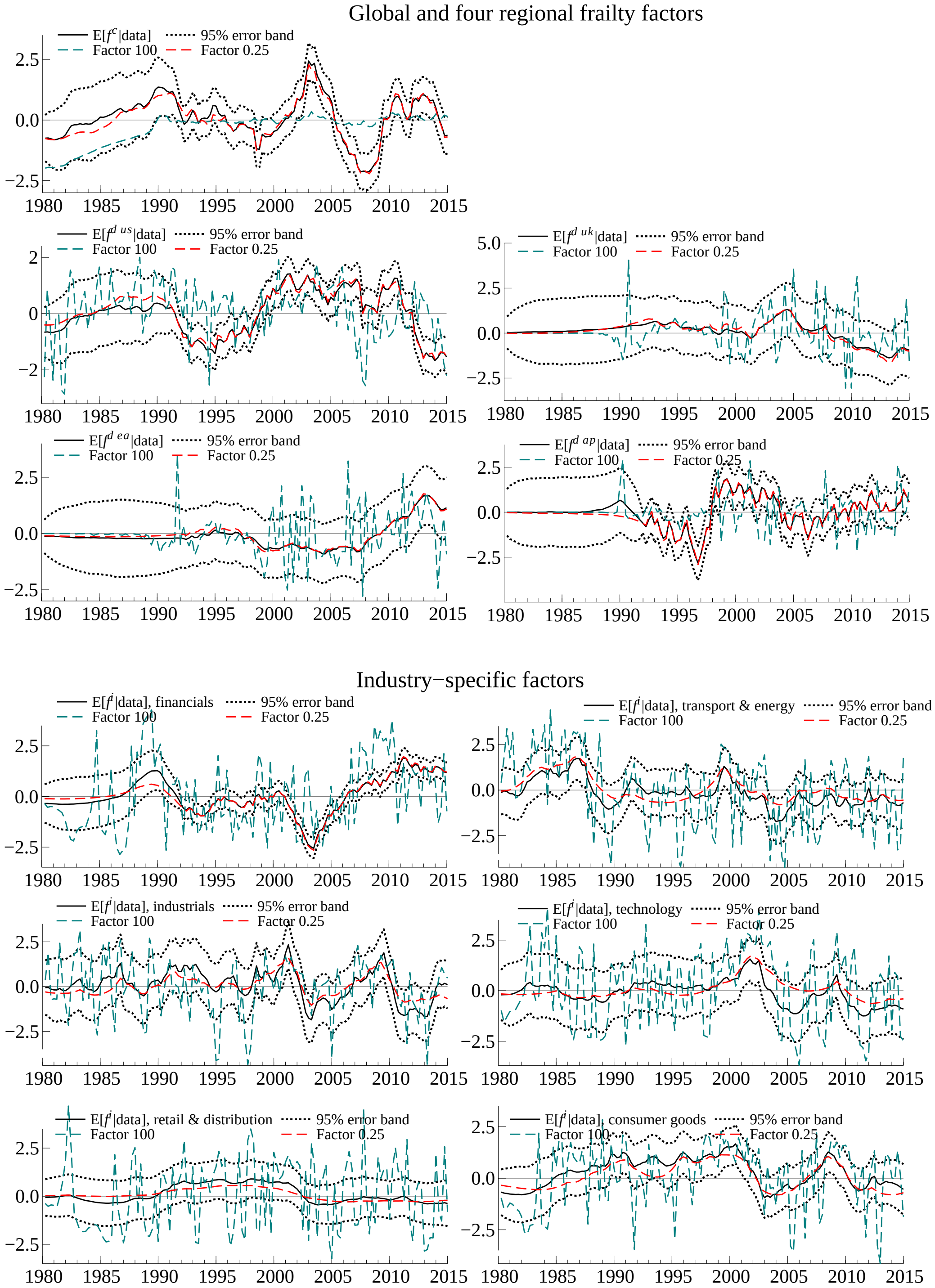
Figure E.1 plots the conditional mean estimate of the global frailty factor, four region-specific frailty factors, and six industry-specific factors. The figure suggests that the EDF data help mainly to "smooth out" the risk factor dynamics, and to inform the factor dynamics in case few count data are available.

Importantly, varying the information content in the Binomial data does not systematically pull the factor estimates in one or the other direction. This is intuitive. There should be little conflict between Moody's EDF-implied default risk conditions and the default fractions obtained from Moody's default and recovery database, as the former are calibrated non-parametrically to the latter, see (Crosbie and Bohn 2003). If the Binomial data are given little weight, the EDF data dominates the risk factor estimates, and risk factor estimates change to some extent. Conversely, if the Binomial data are given a lot of weight, the factor dynamics become quite erratic. They nevertheless remain centered around the full sample estimates, and mostly inside the respective 95% standard error bands.

Finally, the outcomes are in line with the intuition that the more actual default and exposure data is available for a given region and industry sector, the less does the EDF data contribute to risk factor inference. For example, the differences between the full-sample estimates of the regional frailty factors (black solid line) and the respective estimates from the first artificial dataset (red dashed line) appear to be more pronounced for U.S. firms than non-U.S. firms, and more pronounced in the pre-1990s.

Figure E.1: Sensitivity analysis

The top panel reports the conditional mean estimates of the global frailty factor and four region-specific frailty factors. The bottom panels plot the conditional mean estimates of the six industry-specific factors. Each panel plots the full sample estimate, as well as the conditional mean estimate obtained by scaling up and down the contribution of the Binomial part.



References

Crosbie, P. and J. R. Bohn (2003). Modeling default risk. *Moodys KMV Company, available online.*.

Durbin, J. and S. J. Koopman (2012). *Time series analysis by state space methods, 2nd edition.* Oxford: Oxford University Press.

Geweke, J. (1989). Bayesian inference in econometric models using Monte Carlo integration. *Econometrica 57*, 1317–39.

Kloek, T. and H. K. van Dijk (1978). Bayesian estimates of equation system parameters: an application of integration by Monte Carlo. *Econometrica 46*, 1–20.

McNeil, A. J., R. Frey, and P. Embrechts (2005). *Quantitative risk management: Concepts, techniques and tools.* Princeton, N.J.: Princeton University Press.