What drives global credit risk conditions?

Inference on world, regional, and industry factors*

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

[(a)] European Central Bank, Financial Research

[(b)] Tinbergen Institute and VU University Amsterdam

[(c)] Duisenberg school of finance

October 15, 2014

*Author information: Siem Jan Koopman, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands, s.j.koopman@feweb.vu.nl; Andre Lucas, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands, a.lucas@vu.nl; Bernd Schwaab, European Central Bank, Kaiserstrasse 29, 60311 Frankfurt, Germany, bernd.schwaab@ecb.int. We thank seminar and conference participants at the Bundesbank, European Central Bank, and the NUS RMI 2014 credit risk conference in Singapore. André Lucas thanks the Dutch National Science Foundation (NWO) for financial support. The views expressed in this paper are those of the authors and they do not necessarily reflect the views or policies of the European Central Bank or the European System of Central Banks.

What drives global credit risk conditions? Inference on world, regional, and industry factors

S. J. Koopman, A. Lucas, B. Schwaab

Abstract

This paper investigates the common dynamic properties of systematic default risk conditions across countries, regions, and the world. We use a high-dimensional non-linear non-Gaussian state space model to estimate common components in firm defaults in a 22-country sample covering six broad industry sectors and four economic regions in the world. The results indicate that common world factors are a first order source of default risk volatility and default clustering, and that shared exposure to world factors limit the scope of cross-border credit risk diversification for global lenders. Deviations of credit risk conditions from macro fundamentals are correlated with bank lending standards in all regions, suggesting that credit supply and systematic default risk conditions are inversely related. Contrary to intuition, we demonstrate that cross-border credit risk diversification can increase portfolio risk.

Keywords: systematic default risk; credit portfolio models; frailty-correlated defaults; international credit risk cycles; state-space methods.

JEL classification: *G21, C33*

1 Introduction

Is there a world credit risk cycle? Recent studies provide evidence that there are many cross-country links and common global dynamics in macroeconomic fluctuations, inflation dynamics, and financial asset returns. For example, Kose, Otrok and Whiteman (2003, 2008), and Kose, Otrok, and Prasad (2012) document the presence of a world business cycle, and analyse its statistical properties as well as its economic determinants. Ciccarelli and Mojon (2010) and Neely and Rapach (2011) find pronounced global common dynamics in international inflation rates, with international influences explaining more than half of the country variances on average. Yet other research points to global common movement in international stock returns (Bekaert, Hodrick, and Zhang (2009)), government bond yields Jotikasthira, Le, and Lundblad (2011), and term structure dynamics (Diebold, Li, and Yue (2008)). Given that many macro-financial observations are best thought of as global phenomena, we ask whether the same is true for corporate default rates. Thus, is there a world default risk cycle, and if so, what are its statistical properties? How different is the world default risk cycle from world business cycle dynamics that also affects international default rates? Why may the two be decoupled? Finally, what are the implications of global risk factors for the risk bearing capacity of globally active lenders such as Citigroup, Deutsche Bank, or HSBC?

Just as an important first strand of literature investigates the extent of co-movement across global macroeconomic and financial market variables, a second strand of literature investigates why corporate defaults cluster so much over time in developed market economies such as the United States. In general, the accurate point-in-time measurement of default hazard rates is a complicated task since not all processes that determine corporate default are easily observed. Recent research indicates that readily available macro-financial variables and firm-level information may not be sufficient to capture the large degree of default clustering present in corporate default data. This point is most forcefully made by Das, Duffie, Kapadia, and Saita (2007), who apply a battery of statistical tests, and almost always reject the joint

hypothesis that their default intensities are well specified in terms of (i) easily observed firm-specific and macro-financial information and (ii) the doubly stochastic default times, or conditional independence, assumption. In particular, there is substantial evidence for an additional dynamic unobserved 'frailty' risk factor as well as contagion dynamics, see McNeil and Wendin (2007), Koopman, Lucas, and Monteiro (2008), Koopman and Lucas (2008), Duffie, Eckner, Horel, and Saita (2009), Azizpour, Giesecke, and Schwenkler (2010), Lando and Nielsen (2010), Koopman, Lucas, and Schwaab (2011, 2012), and Creal, Schwaab, Koopman, and Lucas (2013). 'Frailty' and contagion risk cause default dependence above and beyond what is implied by observed covariates alone. Whether such excess clustering is an issue also for non-U.S. data is currently an open question. In this paper we show that this is the case.

The main objective of this paper is to quantify the share of systematic default risk that can be attributed to a world credit risk cycle, to infer its statistical properties and location over time, and to assess to which extent the world credit cycle is decoupled from the world business cycle. Compared to the above credit risk studies, our current paper takes an explicit international perspective on default clustering rather than a U.S.-only perspective. Data sparsity (in particular for the non-U.S. data) as well as econometric difficulties (due to the combination of non-Gaussian data and unobserved risk factors, but also due to 'big data' computational issues) have heretofore limited attention to single countries. To our knowledge there has not been a detailed study of whether fluctuations in international systematic default risk are associated with worldwide, regional, or industry-specific risk drivers. We address this and related issues by employing a high-dimensional mixed measurement dynamic factor modeling framework to disentangle common components in both international macro-financial variables and international default data. We investigate a 22-country sample covering four economic regions of the world - the United States, the United Kingdom, 17 countries from the euro area (including Germany, France, Spain, and Italy as the largest countries in terms of population), and three countries from the Asia-Pacific region (Japan, South Korea, Australia), spanning six broad industry sectors - financial firms, transportation

& energy, manufacturing, technology, retail & distribution, and consumer goods, while at the same time taking into account firm specific information on headquarter location, industry sector, and current rating category.

The econometric methodology we employ allows us to examine the effect of multiple sets of unobserved common factors. Koopman, Lucas, and Schwaab (2012) and Creal, Schwaab, Koopman, and Lucas (2013) develop mixed measurement dynamic factor models to study the systematic and idiosyncratic determinants of corporate default using macro-financial and credit risk data from the U.S.. From a methodological viewpoint, we here extend this work to a substantially larger cross-sectional dimension of data, in particular for modeling default and exposure data as well as macro-financial covariates across economic regions. Since default and exposure data is rare for some economic regions, we augment this data by in addition considering data on expected default frequencies (EDF) provided by Moody's Analytics. One of the major advantages of our current econometric framework over those used in earlier studies is that our method allows us to, first, collapse some parts of our data using the projection techniques from Jungbacker, Koopman, and van der Wel (2011) and Brauning and Koopman (2013), and second, to combine the transformed data in one integrated framework, where risk factors and parameters are estimated simultaneously.

We are able to examine region and industry-specific default risk factors simultaneously with global default risk factors. The importance of studying all three simultaneously is that studying a country (or subset of countries) in isolation can lead one to believe that observed co-movement is particular to that country, when it is in fact global or common to a much larger group of countries. For example, our findings indicate that there is a distinct world credit cycle that is related to but different from the world business cycle. We find that shared exposure to global macroeconomic factors explains 4-11% of total (systematic and idiosyncratic) default risk variation across economic regions considered in this paper. A global frailty factor accounts for an additional 9-31% of total risk. Industry-specific variation accounts for 20-36% of total risk, while residual macro factors (0-1%) and regional frailty factors (1-13%) also relatively less important. The global frailty and industry-specific factors

are quite persistent. This means that credit risk conditions can decouple substantially from what is implied by macro-economic fundamentals, and do so for an extended period of time, before eventually returning to their long run means.

We trace back the decoupling of default risk conditions from macro fundamentals to variation in bank lending standards. Default risks that lower than expected based on macro fundamentals coincide with net falling lending standards. (Vice versa, a net tightening in lending standards leads to higher than expected systematic default risk conditions. The ease of credit access can materially impact physical measures of corporate default risk for example through its effect on rollover risk (He and Xiong (2012)). As a result, local credit risk conditions are also a reflection of global trends, such as world-wide variations in the ease of credit access and (globally correlated) lending standards, see Bruno and Shin (2012).

Interestingly, and perhaps counter-intuitively, more credit risk diversification across border does not necessarily decrease portfolio risk, even if marginal risks (such as ratings) are held fixed. Two effects work in opposite directions. First, expanding the portfolio across borders decreases risk dependence if regional (macro or frailty) risk factors are important and imperfectly correlated. On the other hand, however, the expansion of the portfolio could involve loans to firms that load relatively more on the global risk factors, such as global macro and global frailty factors. We empirically document that this trade-off is a relevant concern.

Understanding the sources of international credit risk variation is important for developing reliable portfolio credit risk models at internationally active financial institutions. It also matters in the context of risk model validation and the effective supervision of global banks by the appropriate authorities. In addition, our joint modeling framework of macro-financial data and default risk variation has obvious applications to the stress testing of global bond portfolios. Finally, understanding the properties and sources of international credit risk variation matters for regulatory policy: A global credit risk cycle naturally limits the scope for cross-border credit risk diversification in loan portfolios held by globally active financial firms. While global lenders could have a superior "risk bearing capacity" (Eufinger

and Richter, 2013) compared to more local banks due to cross border diversification benefits, this advantage may not be substantial, or not be present at all. Rather, shared exposure to common factors across economic regions limits the scope of diversification benefits, and thus also the positive economies of scale ascribed to large financial sector firms, see Wheelock and Wilson (2012).

The remainder of this paper is organized as follows. Section 2 introduces our global default risk and macro data, and gives initial evidence on cross border risk clustering. Section 3 formulates a financial framework in which default dependence is driven by global, regional, and industry-specific risk factors. Section 4 introduces our empirical framework and discusses parameter and risk factor estimation. Section 5 discusses model selection as well as our key empirical results. We present a variance decomposition of global default risk variation into latent risk drivers, and discuss limits to global credit risk diversification. Section 6 concludes.

2 International default data and default clusters

This section introduces our data and provides some first evidence of common movement in macroeconomic and default risk data across borders. We use quarterly data from three sources. First, a panel of macroeconomic and financial time series data is taken from Datastream with the aim to capture international business cycle and financial market conditions. A second dataset is constructed from default and exposure data from Moody's Default and Recovery (DRD) database, covering financial and non-financial firms from several broad industry sectors and four economic regions - the U.S., U.K., euro area, and the Asia-Pacific region. Finally, we consider Expected Default Frequencies (EDF) indexes which are also taken from Moody's.

2.1 International macro data and principal components

For our stacked panel of macro-economic covariates, we select data series that are usually stressed in a supervisory macro stress test, see for example Tarullo (2010). The macro panel consists of coincident, leading, and lagging business cycle indicators. Coincident indicators are the real GDP growth rate, industrial production growth, a survey-based indicator (such as the ISM purchasing managers index), and the yoy change in the unemployment rate. Two leading indicators are the term structure spread (-5Q) and the change in a broad equity market index (-1Q)). The lagging indicators are the change in 10 year government bond yields (+1Q), change in residential property prices (+2Q), and the unemployment rate (+5Q). The lead and lag relationships are determined based on the respective cross-correlation coefficients viz-a-viz the real GDP growth rate, and are in line with e.g. Stock and Watson (1989). These nine variables are considered in the macro panel for the U.S., the U.K, the euro area, and the Asia Pacific region/Japan. This yields 9x4=36 macro variables in total, at a quarterly frequency from 1980Q1 to 2013Q3.

2.2 International default data

As a second large scale panel data set, we consider default and exposure counts from Moody's default and recovery database. The database contains rating transitions and default dates for all rated firms, worldwide, from 1980Q1 to 2013Q3. From these data, we construct quarterly values for $y_{r,j,t}$ and $k_{r,j,t}$ in (3). Moody's broad industry classification allows us to pool firms into six broad industry sectors: banks and other financial institutions such as insurers, trusts, and real estate (fin); transportation, utilities, and energy & environment (tre); capital industries and manufacturing (ind); technology firms (tec); retail & distribution (ret); and, finally, consumer industry firms (con). When counting exposures $k_{r,j,t}$ and corresponding defaults $y_{r,j,t}$, a previous rating withdrawal is ignored if it is followed by a later default. In this way, we limit the impact of strategic rating withdrawals. If there are multiple defaults per firm, we consider only the first event.

Figure 1: Historical default and firm counts

The top panel plots time series data of (a) the total default counts $\sum_j y_{r,j,t}$ aggregated to a univariate series, (b) the total number of firms at risk $\sum_j k_{r,j,t}$, and (c) aggregate default fractions $\sum_j y_{r,j,t}$ / $\sum_j k_{r,j,t}$ over time. The bottom panel plots observed default fractions at the industry level for four different economic regions. Each panel distinguishes firms headquartered in the United States, the United Kingdom, the euro area, and Asia-Pacific (Japan, Korea and Australia) region. Light-shaded areas are NBER recession times.

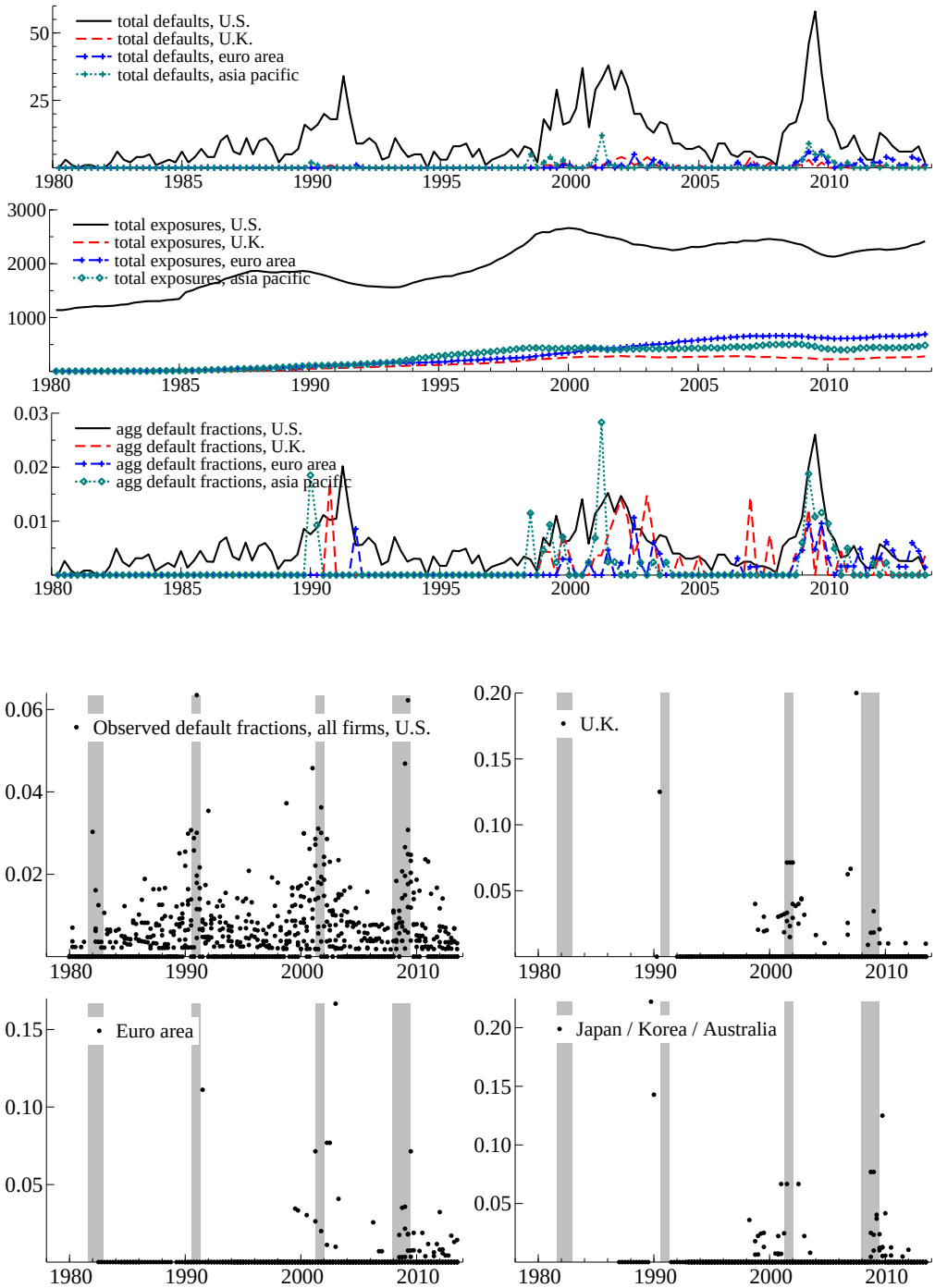


Figure 2: Aggregate EDF data for global financial and non-financial firms
We plot EDF panel data for financial and non-financial firms from Moody's KMV from the U.S., the U.K., the euro area, and the Asia-Pacific region. The series are weighted cross-sectional averages, with weights according to firms' total assets. The data sample is from 1992Q1 to 2013Q3.

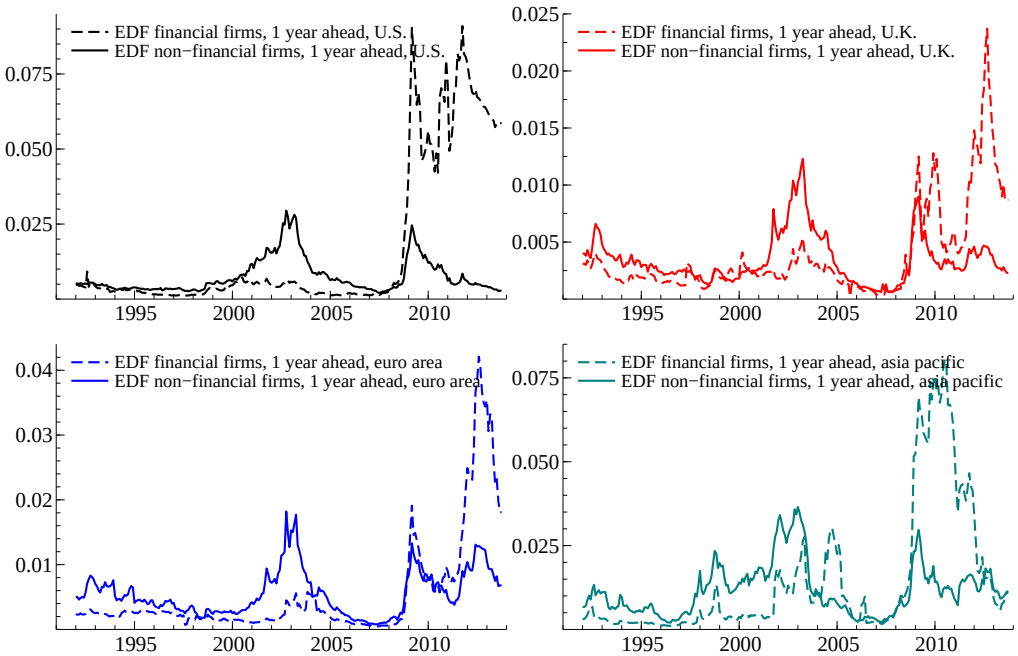


Figure 1 plots aggregate default counts, exposures, and observed fractions over time for each economic region. The bottom panel in particular suggests that defaults tend to cluster across economic regions. Put simply, bad times (meaning a cluster of corporate defaults) tend to be bad regardless of the considered region. The highest default fractions are observed during (U.S.) recession years such as 1990-91, 2001-02, and 2007-2009.

2.3 Expected default frequencies

As a third and final dataset, we consider expected default frequencies from Moody's Analytics (formerly Moody's KMV). Expected default frequencies are based on a proprietary firm value model that takes equity values and balance sheet information as inputs. We use EDF data from 1992Q1 to 2013Q3 to augment our relatively sparse data on actual defaults $y_{r,j,t}$ for U.S. but in particular non-U.S. corporates. Figure 2 reports EDF index aggregates that are

obtained as cross sectional averages weighted by firms' respective total assets. The panels distinguish financial and non-financial firms in the U.S., U.K., euro area, and the Asia-Pacific region (here: Japan). Similarly to the clustering of actual defaults from Section 2.2, there is a clearly visible coincidence of high default rate estimates during 1992 (the beginning of the sample), 2001-02, and 2007-09. Our worldwide credit risk analysis would be very hard, or impossible, to do without the additional information from the EDF measures as defaults and exposure counts are sparse for the non-U.S. part of our sample.

3 The modeling framework

3.1 A multi-factor model of default risk dependence

This section develops a simple multi-factor financial framework for dependent defaults. The financial framework is similar to the well-known CreditMetrics (2007), and introduces cross-border default dependence due to shared global macroeconomic, default-, and industry-specific factors, as well as regional macro and frailty factors. We then show that the financial framework is closely related to a mixed-measurement latent dynamic factor model. By relating the financial with the econometric model, we establish a semi-structural economic interpretation of the parameters of interest.

In the special case of a standard static one-factor credit risk model for dependent defaults the values of the borrowers' assets, V_i, are driven by a common random factor f, and an idiosyncratic disturbance ϵ_i. More specifically, the asset value of firm i, V_i, is modeled as

$$V_i = \sqrt{\rho_i}\, f + \sqrt{1-\rho_i}\, \epsilon_i,$$

where scalar $0 < \rho_i < 1$ weights the dependence of firm i on the general economic condition factor f in relation to the idiosyncratic factor ϵ_i, for $i = 1, \ldots, K$, where K is the number of firms, and where $(f\,,\,\epsilon_i)'$ has mean zero and variance matrix I_2. The conditions in this framework imply that

$$\mathrm{E}(V_i) = 0, \qquad \mathrm{Var}(V_i) = 1, \qquad \mathrm{Cov}(V_i V_j) = \sqrt{\rho_i \rho_j},$$

for $i, j = 1, \ldots, K$. In our multivariate dynamic model, the framework is extended into a more elaborate version for the asset value V_{it} of firm i at time t and is given by

$$\begin{aligned} V_{it} &= a_i' f_t^g + b_i' f_t^m + c_i' f_t^c + d_i' f_t^d + e_i' f_t^i + \sqrt{1 - a_i' a_i - b_i' b_i - c_i' c_i - d_i' d_i - e_i' e_i}\ \epsilon_{it} \quad (1) \\ &= w_i' f_t + \sqrt{1 - w_i' w_i}\ \epsilon_{it}, \qquad\qquad t = 1, \ldots, T, \end{aligned}$$

where global macro factors f_t^g, region-specific macro factors f_t^m, common global default-specific (frailty) factors f_t^c, region-specific frailty factors f_t^d, as well as global industry-specific factors f_t^i are stacked in $f_t = (f_t^{g\prime}, f_t^{m\prime}, f_t^{c\prime}, f_t^{d\prime}, f_t^{i\prime})'$, and the stacked weight vector $w_i = (a_i', b_i', c_i', d_i', e_i')'$ satisfies the condition $w_i' w_i \leq 1$. The idiosyncratic disturbance ϵ_{it} is serially uncorrelated for $t = 1, \ldots, T$.

In our framework, macroeconomic risk factors can be common to all countries (f_t^g), or region-specific (f_t^m). Analogously, the frailty factor can be common to firms from all regions (f_t^c), or region-specific (f_t^d). Taken together, the frailty factors represent credit cycle conditions after controlling for macroeconomic developments. In other words, frailty factors capture deviations of the default cycle from systematic macroeconomic and financial conditions. Industry-specific risk factors f_t^i are common to firms from the same industry sector, regardless of their geographical location. Without loss of generality we assume that all risk factors have zero mean and unit unconditional variance. Furthermore, we assume that the risk factors in f_t are uncorrelated with each other at all times. These assumptions imply that $\mathrm{E}[V_{it}] = 0$ and $\mathrm{Var}[V_{it}] = 1$ for many distributional assumptions with respect to the idiosyncratic noise component ϵ_{it} for $i = 1, \ldots, I$, such as the Gaussian or Logistic distribution.

In a firm value model, firm i defaults at time t if its asset value V_{it} drops below some default threshold $\bar{\lambda}_i$, see Merton (1974) and Black and Cox (1976). Intuitively, if the total value of the firm's assets is below the value of its debt to be repaid, equity holders with limited liability have an incentive to walk away and to declare bankruptcy. In our framework, V_{it} in (1) is driven by multiple latent systematic factors, while idiosyncratic risk is captured by ϵ_{it}. The threshold $\bar{\lambda}_i$ may depend on headquarter location, the firm's current rating category,

and possibly its industry sector. For firms which have not defaulted yet, a default occurs when $V_{it} < \bar{\lambda}_i$ or, as implied by (1), when

$$\epsilon_{it} < \frac{\bar{\lambda}_i - w_i' f_t}{\sqrt{1 - w_i' w_i}}.$$

The conditional default probability is given by

$$\pi_{it} = \Pr\left(\epsilon_{it} < \frac{\bar{\lambda}_i - w_i' f_t}{\sqrt{1 - w_i' w_i}}\right). \tag{2}$$

Favorable credit cycle conditions are associated with a high value of $w_i' f_t$ and therefore with a low default probability π_{it} for firm i. Since only firms are considered at time t that have not defaulted yet, π_{it} can also be referred to as a discrete time default hazard rate, or default intensity under the historical probability measure, see Lando (2003, Chapter3).

3.2 Binomial mixture model representation

Our empirical analysis considers a setting where the firms ($i = 1, \ldots, I$) are pooled into groups ($j = 1, \ldots, J$) according to geography (headquarter location), industry sector, and current rating class. We assume that the same risk factor loadings apply to each firm in the same group. In this case, (1) and (2) imply that, conditional on f_t, the counts y_{jt} are generated as sums over independent 0-1 binary trials (no default - default). In addition, the default counts can be modelled as a binomial sequence, where y_{jt} is the total number of default 'successes' from k_{jt} independent bernoulli trials with time-varying default probability π_{jt}. In our case, k_{jt} denotes the number of firms in cell j that are active at the beginning of period t.

$$\begin{aligned} y_{jt}|f_t &\sim \quad \text{Binomial}(k_{jt}, \pi_{jt}), && (3)\\ \pi_{jt} &= [1 + \exp(-\theta_{jt})]^{-1}, && (4)\\ \theta_{jt} &= \chi_j + \alpha_j' f_t^g + \beta_j' f_t^m + \gamma_j' f_t^c + \delta_j' f_t^d + \varepsilon_j' f_t^i, && (5) \end{aligned}$$

where χ_j and $\lambda_j = \left(\alpha_j', \beta_j', \gamma_j', \delta_j', \varepsilon'\right)'$ are loading parameters to be estimated, and θ_{jt} is the log-odds ratio of the default probability π_{jt}. For more details on binomial mixture models,

see Lando (2003, Chapter 9),McNeil, Frey, and Embrechts (2005, Chapter 8), and Koopman, Lucas, and Schwaab (2011, 2012).

Interestingly for our purposes, there is a one-on-one correspondence between the model parameters in (1) and the reduced form coefficients in (5). If ϵ_{it} is logistically distributed, equation (5) is a special case of (11), with θ_{jt} denoting the canonical parameter of the binomial distribution in the exponential family. It is easily checked that for firm i that belongs to group j,

$$\begin{aligned} \bar{\lambda}_i &= \chi_j\sqrt{1-\kappa_j}, & a_i &= -\alpha_j\sqrt{1-\kappa_j}, \\ b_i &= -\beta_j\sqrt{1-\kappa_j}, & c_i &= -\gamma_j\sqrt{1-\kappa_j}, \\ d_i &= -\delta_j\sqrt{1-\kappa_j}, & e_i &= -\varepsilon_j\sqrt{1-\kappa_j}, \end{aligned}$$

where $\kappa_j = \bar{\omega}_j/(1+\bar{\omega}_j)$, and $\bar{\omega}_j = \alpha_j'\alpha_j + \beta_j'\beta_j + \gamma_j'\gamma_j + \delta_j'\delta_j + \varepsilon_j'\varepsilon_j$. We use this correspondence between parameters when assessing the systematic default risk of firms from different regions and industry sectors.

3.3 Quantifying firms' systematic default risk

The firm value model specification (1) allows us to rank the systematic default risk of firms from different industry sectors and economic regions, while controlling for other information such as the firm's current rating category (marginal risk). As concrete examples, we anticipate that a loan to a Japanese technology firm may contribute more to the risk of a globally diversified loan or bond portfolio than the same loan to a consumer goods firm from the euro area, which in turn may be more systematically risky than a loan to a U.K. retailer. We also anticipate that the precise ranking is sensitive to whether industry-specific risk is treated as systematic risk or as idiosyncratic risk that averages out in a diversified portfolio. Clearly, a loan with a lower share of systematic risk is to be preferred from a portfolio risk perspective, assuming that the marginal risks and interest rates are similar. We define the systematic risk of firm i as the variance of its systematic risk component,

$$\mathrm{Var}[V_{it}|\epsilon_{it}] = w_i' w_i, \tag{6}$$

where $w_i = (a_i', b_i', c_i', d_i', e_i')'$, see (1). Since $\mathrm{Var}[V_{it}] = 1$, (6) also denotes the share of total risk that is non-diversifiable (or imperfectly diversifiable) as the systematic risk drivers in f_t also affect many other (or all other) firms in the global portfolio.

4 The econometric framework

This section introduces our mixed-measurement dynamic factor model (MM-DFM) for the joint analysis of default counts and macroeconomic measurements, see also Koopman et al. (2012) and Creal et al. (2013). We also explain our approach to handling the vast dimensions of our data panels. We consider available data

$$\mathcal{Y}_t = (\mathcal{Y}_{1t}, \ldots, \mathcal{Y}_{Nt})', \quad t = 1, \ldots, T, \tag{7}$$

where each row $\mathcal{Y}_i = (\mathcal{Y}_{i1}, \ldots, \mathcal{Y}_{iT})$, $i = 1, \ldots, N$, is univariate time series data from a different family of densities. In particular, some time series may be discrete (Binomial) count data, whereas others are continuous (Gaussian) measurements of business cycle conditions and expected default frequencies.

The observations depend on a set of m dynamic latent factors. These latent factors are assumed to be generated from a dynamic Gaussian process. We collect the factors into the $m \times 1$ vector f_t and assume a stationary vector autoregressive process for the factors,

$$f_{t+1} = \mu_f + \Phi f_t + \eta_t, \qquad \eta_t \sim \mathrm{N}(0, \Sigma_\eta), \qquad t = 1, 2, \ldots, \tag{8}$$

with the initial condition $f_1 \sim \mathrm{N}(\mu, \Sigma_f)$. The $m \times 1$ mean vector μ_f, the $m \times m$ coefficient matrix Φ and the $m \times m$ variance matrix Σ_η are assumed fixed and unknown with the m roots of the equation $|I - \Phi z| = 0$ outside the unit circle and Σ_η positive definite. The $m \times 1$ disturbance vectors η_t are serially uncorrelated. The process for f_t is initialized by

$f_1 \sim \mathrm{N}(0, \Sigma_f)$ where $m \times m$ variance matrix Σ_f is a function of Φ and Σ_η or, more specifically, Σ_f is the solution of $\Sigma_f = \Phi \Sigma_f \Phi' + \Sigma_\eta$.

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

$$\mathcal{Y}_{i,t} | \mathcal{F}_t \sim p_i(\mathcal{Y}_{i,t}; \mathcal{F}_t, \psi), \quad i = 1, \ldots, N. \tag{9}$$

In our case, all observations $\mathcal{Y}_{i,t}$ come from the exponential family of densities,

$$p_i(\mathcal{Y}_{i,t}; \mathcal{F}_t, \psi) = \exp\{\bar{a}_i(\psi)^{-1} \left[\mathcal{Y}_{i,t}\, \theta_{i,t} - \bar{b}_{i,t}(\theta_{i,t}; \psi)\right] + \bar{c}_{i,t}(\mathcal{Y}_{i,t})\}, \tag{10}$$

with the signal defined by

$$\theta_{i,t} = \chi_i + \sum_{j=0}^{p} \lambda'_{i,j} f_{t-j}, \tag{11}$$

where χ_i is an unknown constant and $\lambda_{i,j}$ is the $m \times 1$ loading vector with unknown coefficients for $j = 0, 1, \ldots, p$. The so-called link function in (10) $\bar{b}_{i,t}(\theta_{i,t}; \psi)$ is assumed to be twice differentiable while $\bar{c}_{i,t}(\mathcal{Y}_{i,t})$ is a function of the data only. The parameter vector ψ contains all unknown coefficients in the model specification including those in Φ, χ_i and $\lambda_{i,j}$ for $i = 1, \ldots, N$ and $j = 0, 1, \ldots, p$. To enable the identification of all entries in ψ, we assume standardized factors in (8) which we enforce by the restrictions $\mu_f = 0$ and $\Sigma_f = I$ implying that $\Sigma_\eta = I - \Phi\Phi'$. In principle, the exponential family setup (10) allows for the joint modeling of binary, binomial, Poisson, exponential, negative binomial, multinomial, Gamma, Gaussian, inverse Gaussian, and Weibull multivariate time series observations.

Conditional on $\mathcal{F}_t$, the observations at time t are independent of each other. It implies that the density of the $N \times 1$ observation vector $\mathcal{Y}_t = (\mathcal{Y}_{1,t}, \ldots, \mathcal{Y}_{N,t})'$ is given by

$$p(\mathcal{Y}_t | \mathcal{F}_t, \psi) = \prod_{i=1}^{N} p_i(\mathcal{Y}_{i,t} | \mathcal{F}_t, \psi).$$

The MM-DFM model is defined by the equations (8), (9) and (11).

4.1 Dimensionality reduction

This section explains how our three high-dimensional data sets described in Section 2 are transformed and combined to obtain data of tractable dimensions. We initially consider mixed measurement panel data $\bar{\mathcal{Y}}_t = (x_t', y_t', z_t')'$ which follows a tri-part data structure

$$x_t = (x_{1,1,t}, \ldots, x_{1,N_1,t}, \ldots, x_{R,1,t}, \ldots, x_{1,N_R,t})', \tag{12}$$

$$y_t = (y_{1,1,t}, \ldots, y_{1,J_1,t}, \ldots, y_{R,1t,}, \ldots, y_{R,J_R,t})', \tag{13}$$

$$z_t = (z_{1,1,t}, \ldots, z_{1,S_1,t}, \ldots, z_{R,1,t}, \ldots, z_{R,S_R,t})', \tag{14}$$

where $x_{r,n,t}$ represents the nth, $n = 1, \ldots, N_r$, macroeconomic or financial markets variable for region $r = 1, \ldots, R$ measured at time $t = 1, \ldots, T$; $y_{r,j,t}$ is the number of defaults between times t and $t+1$ for economic region r and cross section $j = 1, \ldots, J_r$; and $z_{r,s,t}$ is the expected default frequency (EDF) of financial firm $s = 1, \ldots, S_r$ in economic region r at time t. The cross section j represents different categories of firms. The model thus includes more 'standard', possibly normally distributed macro and financial markets variables x_t, but also count variables y_t and variables z_t that are bounded to the [0,1] interval. The panel (x_t, y_t, z_t) is typically unbalanced, such that variables may not be observed at all times.

The cross-sectional dimension of the tri-part data (12) is prohibitively large for our multi-country credit risk model. For this reason we collapse the linear Gaussian part of our mixed measurement panel data

$$\begin{pmatrix} \hat{P}_x x_t \\ \hat{P}_{z,t} z_t \end{pmatrix} = \begin{pmatrix} \hat{F}_{x,t} \\ \hat{F}_{z,t} \end{pmatrix}, \tag{15}$$

such that data x_t and z_t are projected into panel data sets of much smaller cross-sectional dimensions. $\hat{F}_{x,t}$ are the first r principal components of macro panel data x_t. As such, $\hat{F}_{x,t}$ contains the global common macro factors. For the global macro factors, the projection matrix $\hat{P}_x = U = (U_1, \ldots, U_r)$, where U_r is the eigenvector corresponding to the r largest ordered eigenvalues of $X'X$, where $X' = (x_1, \ldots, x_T)$. In addition to global macro factors, $\hat{F}_{x,t} = \left(\hat{F}_t^{g\prime}, \hat{F}_t^{m\prime}\right)'$ also contain regional macro factors. The regional factors are obtained as the first principal component of the remaining variation in the region-specific subset of

macroeconomic data after projecting off of the space spanned by the global macro factors.

EDF factors $\hat{F}_{z,t} = \hat{P}_{z,t} z_t$ in (15) are weighted averages of EDF-based quarterly default hazard rates. Matrix $\hat{P}_{z,t}$ weighs the EDFs according to balance sheet information (a firm's total assets); we take this information from Moody's as well. Weighted averages exploit the fact that hazard rates are additive, see Lando (2003).

After collapsing the Gaussian data based on the method of principal components, the transformed data is given by

$$\mathcal{Y}_t = \left(\hat{F}'_{x,t}, y'_t, \hat{F}'_{z,t}\right)'. \tag{16}$$

While the cross-sectional dimension of the original data (12) — (14)is prohibitively large, the cross-sectional dimension of collapsed data (16) is tractable and is used for parameter estimation and risk factor extraction below.

4.2 Estimation

The estimation of parameters and risk factors 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 $Y = (Y'_1, \ldots, Y'_T)'$ and $f = (f'_1, \ldots, f'_T)'$ denote the vector of all the observations and factors, respectively. Let $p(Y|f;\psi)$ be the density of 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

$$p(Y;\psi) = \int p(Y, f;\psi)\,\mathrm{d}f = \int p(Y|f;\psi)p(f;\psi)\,\mathrm{d}f, \tag{17}$$

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 (2001). Upon computing the integral, the maximum likelihood estimator of ψ is obtained by direct maximization of the likelihood function using Newton-Raphson methods.

We refer to the Appendix A1 for the main estimation details, and here only put forward a final remark. For our empirical analysis in Section 5, we need to integrate out 18 latent factors from their joint density with the mixed measurement observations, for each

evaluation of the log-likelihood. This causes the estimation setup put forward in Koopman, Lucas, and Schwaab (2011, 2012) to break down. In particular, the importance sampling weights then do not appear to have a finite variance in the large dimensional case considered in this paper. We overcome this challenge by using four antithetic variables for location and scale as suggested in Durbin and Koopman (1997). We leave the suggestion that employing antithetic variables can stabilise the importance sampling weights substantially if large dimensional vectors of stacked factors need to be integrated out for likelihood evaluation based on simulation methods.

5 Main empirical results

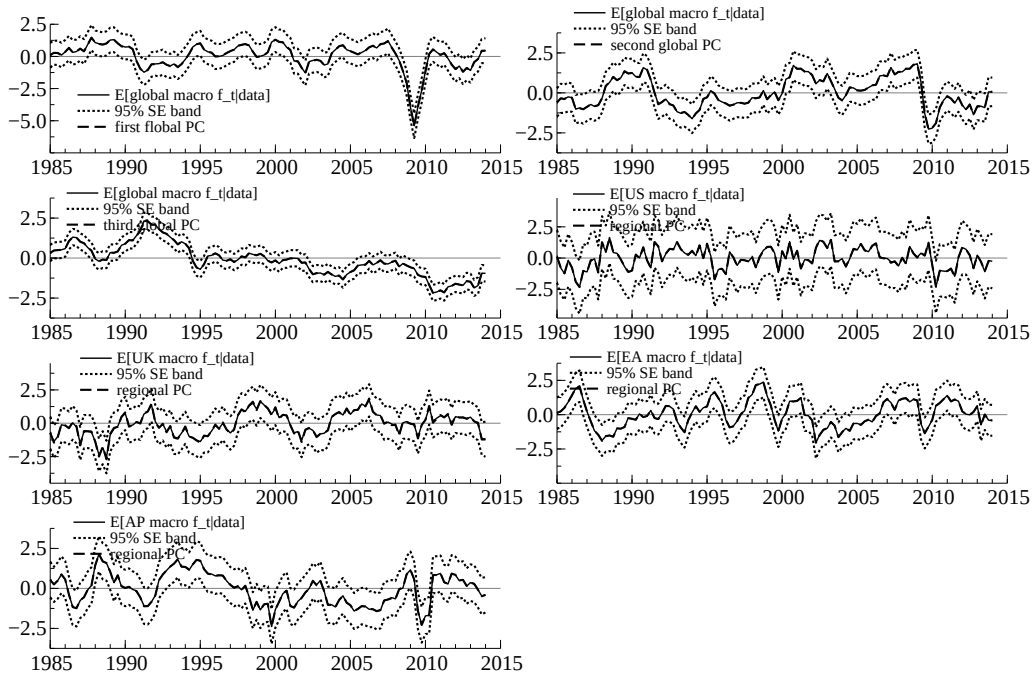
5.1 Model specification

This section discusses issues related to model specification such as the number of factors and pooling over parameters. For the selection of the number of factors we rely on likelihood-based information criteria (IC). Based on standard criteria such as the likelihood, the panel information criteria of Bai and Ng (2002), and the summary statistics reported in Section 2, we select three global macro-financial factors f_t^g, and four region-specific macro factors (one for each region). The global macro factors are common to all macro-financial covariates and all default data, capturing the positive correlation between U.S. and non-U.S. business cycle conditions. These factors are stacked in f_t^g and f_t^m, respectively.

Allowing for one common frailty factor f_t^c is standard in the literature, see for example Duffie et al. (2009) and Azizpour et al. (2010). Indeed, one (global) frailty factor is often sufficient to capture pronounced deviations of systematic default risk from macro conditions. We further allow for four region-specific frailty factors f_t^d, one for each region. Finally, we allow for six additional industry-specific factors f_t^i. These factors load only on firms from the respective industry sector, regardless of the economic region. International default data loads on global macro factors f_t^g and the global frailty factor f_t^c with region-specific factor

Figure 3: Global and regional macro factors

Conditional mean and principal components estimates for the global and regional macro factors. There are seven factors: three global macro factors, and four region-specific macro factors (one for each region).



loading coefficients.

For all risk factors, we pool risk factor loadings across rating classes. This allows us to focus on differences in systematic risk across corporates from different industries and different regions. While somewhat restrictive, this specification remains sufficiently flexible to accommodate most of the heterogeneity observed in the cross section and allows us to test the key economic hypotheses at hand.

5.2 Parameter and risk factor estimates

This section discusses our results on parameter and risk factor estimation. Table 1 reports model parameter estimates. The table indicates that all five sets of risk factors - global and regional macro, global and regional frailty, as well as industry-specific - are important for explaining corporate default clusters in each country and across countries.

The estimates of $\alpha_{k,r,j}$ and $\beta_{k,r,j}$ correspond to global and regional macro-financial factors,

Table 1: **Parameter estimates**

We report the maximum likelihood estimates of selected coefficients in the specification of the log-odds ratio (5) with an additive parametrization for $\lambda_{r,j}$ and $\alpha_{r,j}$. Coefficients $\lambda_{r,j}$ combine to fixed effects, or baseline default rates. Factor loadings $\alpha_{r,j}$, $\beta_{r,j}$, γ_j, $\delta_{r,j}$, and ϵ_j refer to three global macro factors f_t^g, four region-specific macro factors f_t^m, one global frailty factor f_t^c, four region-specific frailty factors f_t^d, and six industry-specific factors f_t^i, respectively. The global macro factors are common to all macro and default data and across all four regions. The global and regional frailty factors load on financial and non-financial firms' defaults in the respective region. Industry mnemonics are financials (fin), transportation and energy (tre), industrial firms (ind), technology (tec), retail and distribution (red), and consumer goods (con). Estimation sample is 1980Q1 to 2013Q3.

Intercept terms		
$\lambda_{r,j}=\bar{\lambda}_0+\bar{\lambda}_{1,j}+\bar{\lambda}_{2,s}+\bar{\lambda}_{3,r}$		
par	val	t-val
$\bar{\lambda}_0$	-4.66	22.9
$\bar{\lambda}_{1,fin}$	-0.54	2.51
$\bar{\lambda}_{1,tre}$	-0.19	0.72
$\bar{\lambda}_{1,tec}$	-0.39	1.13
$\bar{\lambda}_{1,ret}$	-0.06	0.17
$\bar{\lambda}_{1,con}$	-0.27	0.74
$\bar{\lambda}_{2,IG}$	-4.02	26.6
$\bar{\lambda}_{3,UK}$	0.07	0.48
$\bar{\lambda}_{3,EA}$	-0.41	3.36
$\bar{\lambda}_{3,AP}$	-0.45	3.56
Global macro factors f_t^g		
$\alpha_{k,r,j}=\bar{\alpha}_{k,0}+\bar{\alpha}_{k,1,r}; k=1,2$		
ϕ_1^g	0.88	14.1
$\bar{\alpha}_{1,0}$	-0.26	5.35
$\bar{\alpha}_{1,1,UK}$	-0.16	2.96
$\bar{\alpha}_{1,1,EA}$	-0.08	1.70
$\bar{\alpha}_{1,1,AP}$	-0.10	1.20
ϕ_2^g	0.89	15.1
$\bar{\alpha}_{2,0}$	-0.11	2.09
$\bar{\alpha}_{2,1,UK}$	-0.06	0.94
$\bar{\alpha}_{2,1,EA}$	-0.18	3.27
$\bar{\alpha}_{2,1,AP}$	0.13	1.08

Global macro factors f_t^g (continued)		
ϕ_3^g	0.97	20.9
$\bar{\alpha}_{3,0}$	-0.07	0.69
$\bar{\alpha}_{3,1,UK}$	0.03	0.27
$\bar{\alpha}_{3,1,EA}$	-0.11	1.08
$\bar{\alpha}_{3,1,AP}$	0.06	0.37
Regional macro factors f_t^m		
par	val	t-val
ϕ_{US}^m	0.42	1.15
$\bar{\beta}_{0,US}$	0.02	0.90
ϕ_{UK}^m	0.79	7.32
$\bar{\beta}_{0,UK}$	0.11	1.99
ϕ_{EA}^m	0.84	11.0
$\bar{\beta}_{0,EA}$	-0.00	0.07
ϕ_{AP}^m	0.83	10.9
$\bar{\beta}_{0,AP}$	-0.16	1.42
Global frailty factor f_t^c		
ϕ^c	0.95	3.79
$\bar{\gamma}_0$	0.45	2.90
$\bar{\gamma}_{1,UK}$	0.42	2.90
$\bar{\gamma}_{1,EA}$	0.22	2.26
$\bar{\gamma}_{1,AP}$	0.10	0.73

Regional frailty factors f_t^d		
ϕ_{US}^d	0.94	3.82
$\bar{\delta}_{0,US}$	0.42	3.24
ϕ_{UK}^d	0.98	-
$\bar{\delta}_{0,UK}$	0.00	0.01
ϕ_{EA}^d	0.93	2.77
$\bar{\delta}_{0,EA}$	0.20	2.88
ϕ_{AP}^d	0.88	6.19
$\bar{\delta}_{0,AP}$	0.56	3.94
Industry factors f_t^i		
par	val	t-val
ϕ_{fin}^i	0.98	-
$\bar{\epsilon}_{fin}$	0.81	9.16
ϕ_{tre}^i	0.79	3.21
$\bar{\epsilon}_{tre}$	0.82	5.28
ϕ_{ind}^i	0.81	3.64
$\bar{\epsilon}_{ind}$	0.76	3.37
ϕ_{tec}^i	0.95	2.86
$\bar{\epsilon}_{tec}$	0.69	2.12
ϕ_{ret}^i	0.91	3.02
$\bar{\epsilon}_{ret}$	0.80	2.38
ϕ_{con}^i	0.94	2.96
$\bar{\epsilon}_{con}$	0.91	2.90

Figure 4: Global frailty and industry factors

Top panel: location estimates for the global common frailty and three region-specific frailty factors. Bottom panel: five industry-specific factors. EDF data is available from 1990Q1 onwards.

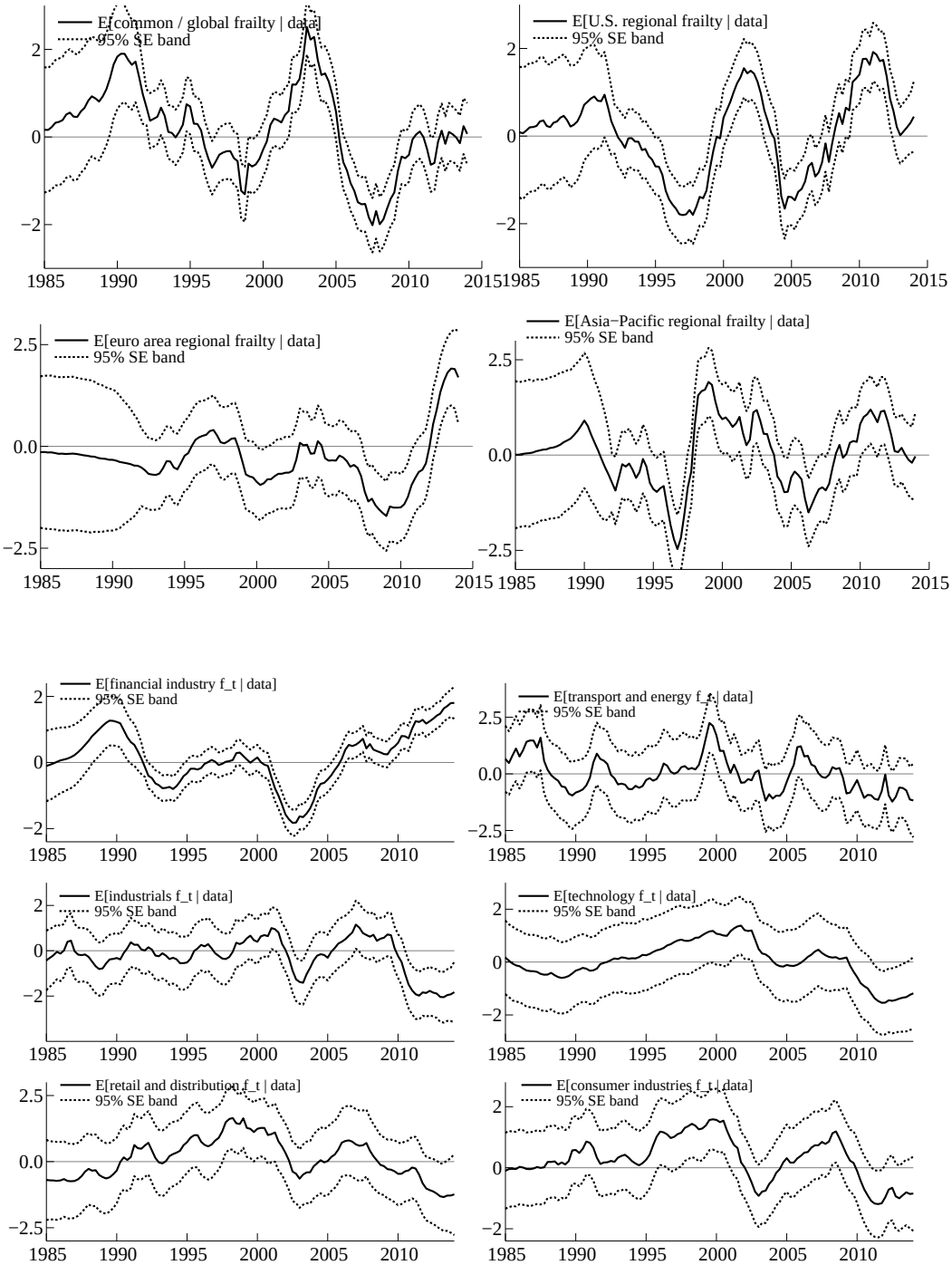


Figure 5: Industry-level default hazard rates

Each panel plots the model-implied default hazard rate, or time-varying quarterly default probability, for a specific industry sector. Each panel reports estimates for four different regions. The reported sample is from 1985Q1 to 2013Q3.

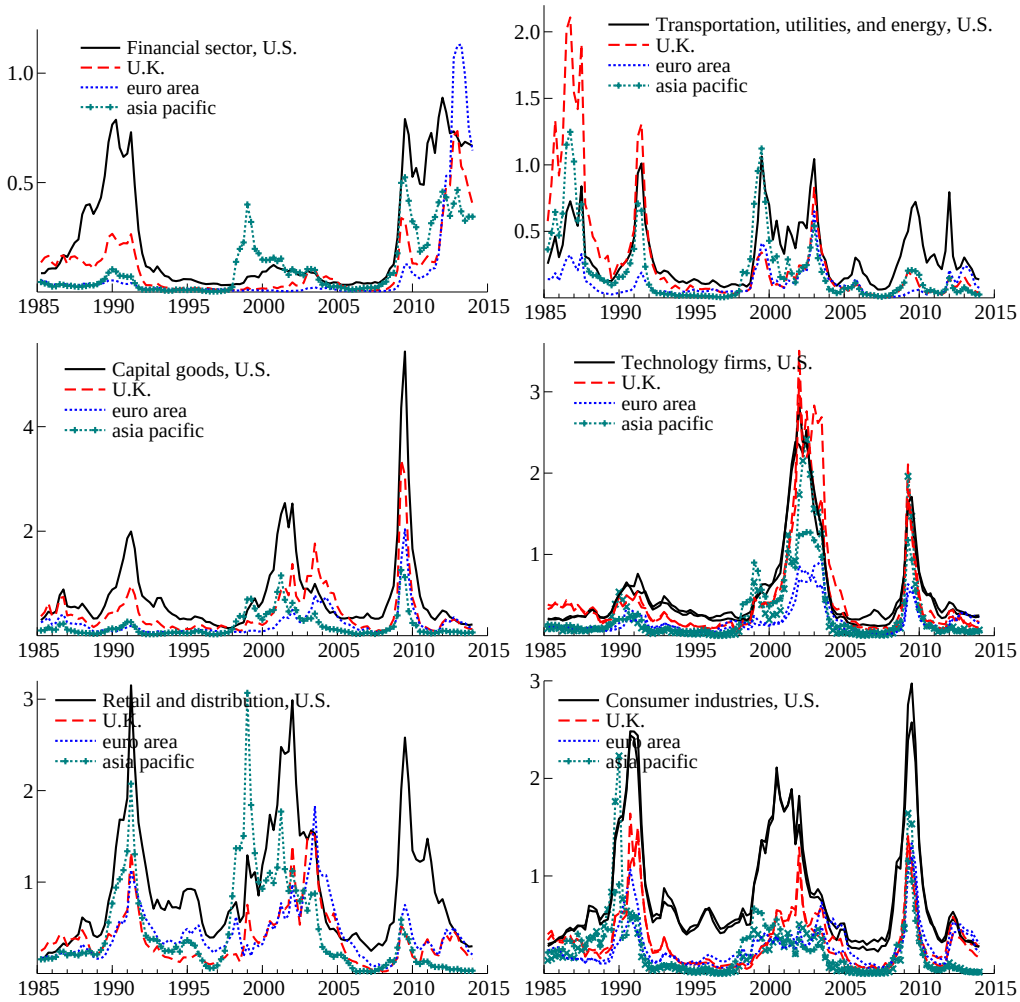
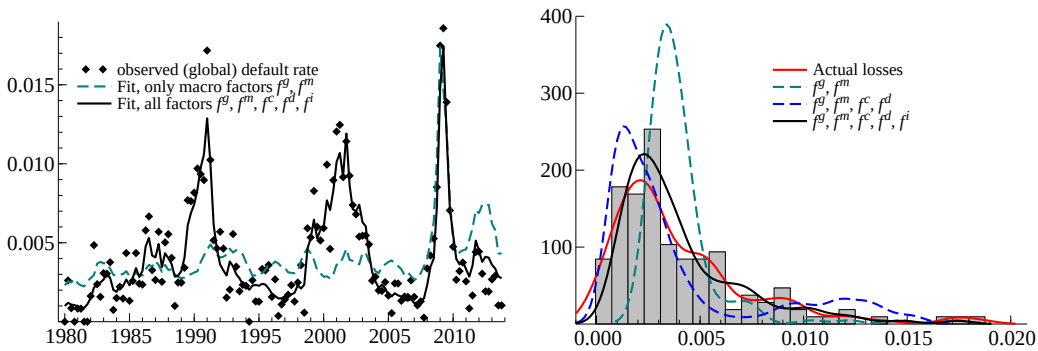


Figure 6: Model fit and loss rates

The left panel reports the in-sample fitted values to the observed global quarterly default fractions. The right panel plots the respective unconditional loss densities for the observed default fractions viz-a-viz the loss densities as implied by the empirical model.



respectively. Defaults from all regions and industries load on macro factors. However, the region-specific macro factors are relatively less relevant than the global macro factors. This finding already implies some degree of cross-border default clustering. Figure 3 plots the principal components estimates of global and regional macro-financial factors. The first three global factors explain 24.9%, 14.7%, and 11.0% of the total data variance, respectively. The regional macro factors explain 25.5%, 27.2%, 24.8%, and 24.3% of the residual variation in the U.S., U.K., euro area, and Asia-Pacific macro data, respectively.

The common variation of defaults with macro data is not sufficient. The estimates of $\gamma_{k,r}$ and $\bar{\delta}_{k,r,j}$ correspond to global and regional macro-financial factors, respectively. The global frailty factor is also found to be important for defaults in all regions. Interestingly, global frailty loads more strongly on non-U.S. data than on U.S. data. The estimates of ϵ_j correspond to global industry-specific factors. Industry-specific dynamics are significant for all defaults. Figure 4 reports location estimates of the global and regional frailty factors, as well as the industry factors. All factors, together with the estimated risk factor loadings, combine into time-varying default hazard rate estimates across industry sectors and regions. Figure 5 plots the respective estimates of default rates for six industry sectors and four regions. Hazard rates vary widely over time and across sectors, often by a factor of ten.

Figure 6 reports the model in-sample fit (left panel) and the unconditional density of

loss rates (right panels). We conclude that the full empirical model specification gives an acceptable fit to the observed aggregate default fractions.

5.3 Variance decomposition

This section uses the framework introduced in Section 3 to decompose the systematic default risk variation of firms from different industry sectors and countries into its underlying risk sources. We present three main empirical findings, and refer to Table 2 for the respective estimates of risk shares.

First, conditional on global business cycle developments, the remaining regional macro factors are less important. This is the case for all combinations of industry sector and geographic location in our sample. Firms from different countries tend to default jointly across borders because of shared exposure to global business cycle dynamics, or alternatively, because the respective macroeconomic conditions are correlated.

Second, global and regional frailty factors explain more default variation than the macro factors. This is true for both U.S. firms nut also the non-U.S. firms from the U.K., euro area and the Asia-Pacific region (as proxied by Japan, South Korea, and Australia). This finding implies significant additional default clustering across borders above and beyond that which is implied by the correlation in observed macro-financial covariates. In short, the excess default clustering as documented in Das et al. (2007) is not a U.S.-specific, but an international phenomenon. Systematic credit risk conditions (the credit risk cycle) can significantly and persistently be decoupled from macro-financial fundamentals (the business cycle), in many countries other than the U.S.

Finally, industry-specific variation is a significant additional source of default clustering. Industry-specific dynamics are the most relevant for the transportation and energy-related (tre) sectors, for which such industry-specific variation is the most important determinant of default risk variation. In addition, the consumer goods (con) and technology (tec) sectors also exhibit strong industry sector dynamics.

Table 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. 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 for a discussion of firm's systematic versus idiosyncratic risk components. Sample is 1980Q1 to 2013Q3.

Reg.	Ind.	f_t^g $[a_i'a_i]$	f_t^m $[b_i'b_i]$	f_t^c $[c_i'c_i]$	f_t^d $[d_i'd_i]$	f_t^i $[e_i'e_i]$	$\mathrm{Var}[V_{it}\|\varepsilon_{it}, f_t^i]$	$\mathrm{Var}[V_{it}\|\varepsilon_{it}]$ $= w_i'w_i$
US	fin	4.0%	0.0%	9.5%	8.4%	31.0%	22.1%	53.0%
US	tre	4.0%	0.0%	9.5%	8.4%	31.5%	21.9%	53.4%
US	ind	4.2%	0.0%	10.0%	8.8%	28.0%	23.0%	51.0%
US	tec	4.4%	0.0%	10.4%	9.2%	24.7%	24.1%	48.7%
US	red	4.1%	0.0%	9.6%	8.5%	30.5%	22.2%	52.7%
US	con	3.8%	0.0%	8.9%	7.9%	35.9%	20.5%	56.4%
UK	fin	7.9%	0.5%	28.7%	0.0%	25.0%	37.1%	62.1%
UK	tre	7.8%	0.5%	28.5%	0.0%	25.5%	36.9%	62.3%
UK	ind	8.2%	0.5%	29.7%	0.0%	22.4%	38.4%	60.8%
UK	tec	8.5%	0.5%	30.8%	0.0%	19.6%	39.8%	59.3%
UK	red	7.9%	0.5%	28.9%	0.0%	24.6%	37.3%	61.9%
UK	con	7.4%	0.5%	27.1%	0.0%	29.4%	34.9%	64.3%
EA	fin	9.7%	0.0%	19.0%	1.7%	27.7%	30.4%	58.1%
EA	tre	9.7%	0.0%	18.8%	1.7%	28.2%	30.2%	58.4%
EA	ind	10.1%	0.0%	19.7%	1.8%	24.9%	31.6%	56.5%
EA	tec	10.5%	0.0%	20.5%	1.8%	21.8%	32.9%	54.7%
EA	red	9.8%	0.0%	19.1%	1.7%	27.2%	30.6%	57.8%
EA	con	9.1%	0.0%	17.8%	1.6%	32.3%	28.5%	60.7%
AP	fin	5.3%	1.0%	12.6%	13.0%	27.0%	32.0%	59.0%
AP	tre	5.3%	1.0%	12.5%	12.9%	27.5%	31.8%	59.3%
AP	ind	5.5%	1.1%	13.1%	13.5%	24.3%	33.2%	57.5%
AP	tec	5.8%	1.1%	13.6%	14.0%	21.3%	34.5%	55.8%
AP	red	5.4%	1.1%	12.7%	13.1%	26.6%	32.2%	58.8%
AP	con	5.0%	1.0%	11.8%	12.2%	31.6%	30.0%	61.6%

5.4 What explains global credit risk decoupling from macro fundamentals?

This section demonstrates that deviations of systematic default risk conditions from macro fundamentals can be traced back to a significant degree to the behaviour of financial intermediaries, and in particular to variation in international bank lending standards.

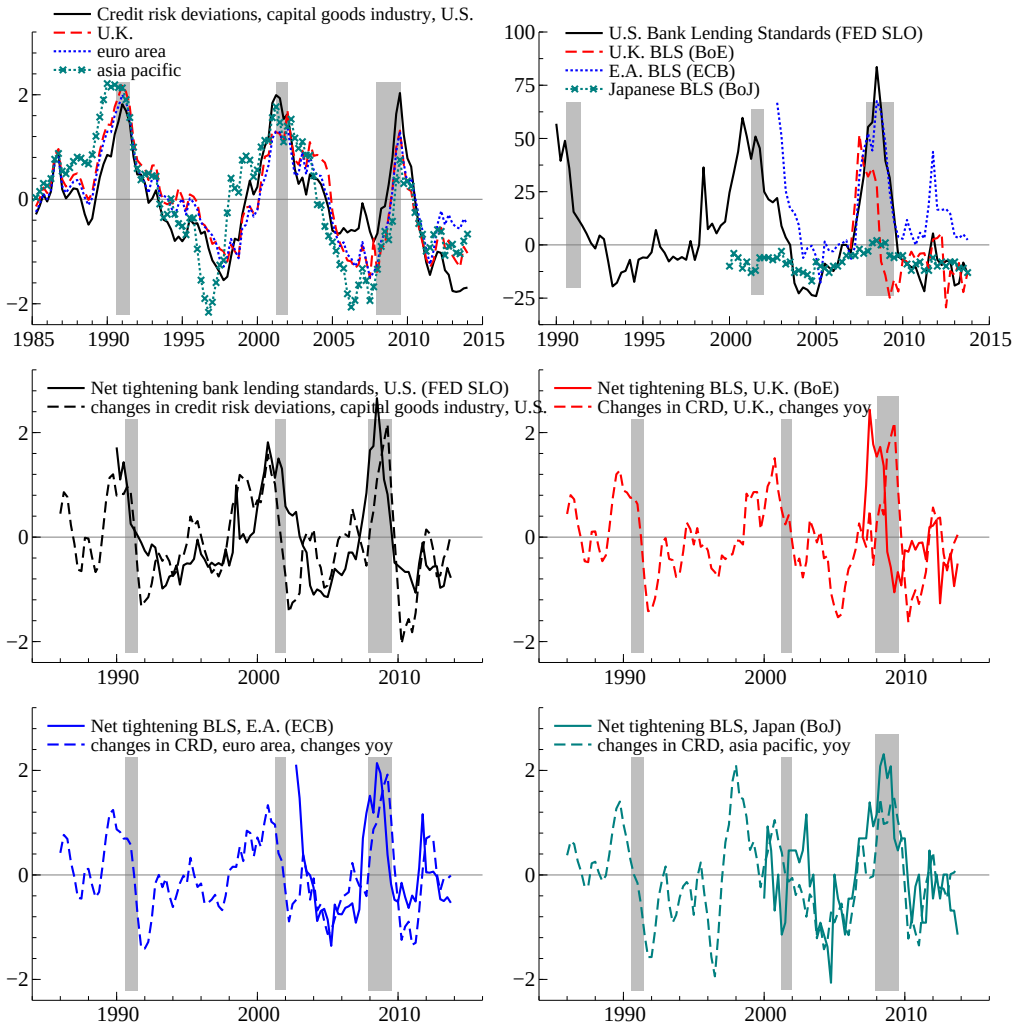
The top left panel in Figure 7 reports deviations of systematic default risk conditions from macroeconomic fundamentals, here taken for firms from the capital goods industry. For example, there is a particularly large and persistent decoupling of risk conditions from fundamentals for non-financial corporates preceding the financial crisis of 2007-2009. Risk conditions were then significantly and persistently below what was suggested by fundamentals. Such a development may indicate a lending bubble, in particular if credit quantity growth is unusually high as well and bank lending standards are generous (which has been the case). The respective top right panel reports bank lending standards for the four economic regions, based on bank surveys undertaken by the respective central banks: The Federal Reserve, the Bank of England, the European Central Bank, and the Bank of Japan.

The bottom four panels in Figure 7 demonstrate that our risk deviation estimates are highly correlated with ex post reported lending standards. Two observations follow. First, physical credit risks and credit quantities are related: In a credit boom, even bad risks have ample access to credit, and can thus postpone default. Therefore, in such a credit boom, bad risks default less frequently than what could be expected conditional on the state of the business cycle. A significantly "too low" default rate is then not a sign of economic strength, and to be welcomed, but could instead be indicative of a boom and thus a warning signal of impending weakness. The reverse holds in a credit crunch. In a credit crunch, even financially sound corporates find it hard to roll over debt, which raises their default risk due to illiquidity concerns. As a result, they default more often than what is expected conditional on the macroeconomic environment.[1] Second, the correlation between bank

[1]To the best of our knowledge, the connection between ease of credit access and systematic credit risk conditions (under the historical measure) was first argued informally in Das, Duffie, Kapadia, and Saita

Figure 7: Credit risk deviations vs. bank lending standards

The top right panel plots deviations of systematic default risk conditions from macro-financial fundamentals for firms from the capital goods industry. Shaded areas pertain to NBER recession dates. The top right panel plots bank lending standards based on survey responses undertaken by the U.S. Fed (for the U.S.), the Bank of England (for the U.K.), the European Central Bank (for the euro area), and the Bank of Japan (for Japan). The bottom four panels correlate changes (yoy) in our credit risk deviation measures with the net tightening reported in the respective bank lending surveys. The sample is from 1980Q1 to 2013Q3.



lending standards and physical credit risks suggest that second round effects through bank lending standards may need to be taken into account in macro-prudential stress tests. While macro-financial stress is naturally mapped into higher default probabilities (and therefore higher bank losses and lower capital ratios), the ensuing financial sector deleveraging and tightening of lending standards may subsequently raise credit risks even further. Figure 7 suggests that such second round effects are substantial.

5.5 The benefits from cross border credit risk diversification

In this section we take the point of view of a risk management officer at a global lending institution, such as, for example, Citigroup, Deutsche Bank, or HSBC. In particular, we are interested in the effects of credit risk diversification across economic regions and industry sectors, as implied by historical data and captured by our empirical models. Specifically, we compare portfolio credit risk measures from a concentrated credit portfolio with the risk of holding alternative portfolios that are more diversified across industries and countries.

Table 3 considers portfolio risk measures for four loan portfolios. Figure 8 plots the historical loss rates and the expected shortfall for each portfolio. The baseline loan portfolio (a) contains 1000 loans of 1$ each to 1000 U.S. manufacturing firms (split equally between investment and speculative grade). As a result, it is concentrated both geographically and in terms of one industry sector. For simplicity, the loans are short term loans (one quarter), and extended to the firms that are active in the beginning of the respective quarter. We compare this baseline case with three, successively more diversified alternatives, holding the overall size of the portfolio constant. Portfolio (b) contains 200 loans of 1$ each to 200 U.S. firms from our five non-financial industry sectors (cf. Section 2.2). As a result, portfolio

(2007) and Duffie, Eckner, Horel, and Saita (2009). The link is explored more formally in a firm value model by He and Xiong (2012), while Koopman, Lucas, and Schwaab (2012) are the first to empirically tie unobserved credit risk deviations to U.S. bank lending standards. These results are much in line with a literature on portfolio credit risk that concludes that easily observed macro-financial covariates and firm-specific information, while helpful, are not sufficient to fully explain time-varying systematic credit risk conditions, see for example Das et al. (2007), Koopman, Lucas, and Monteiro (2008), Duffie et al. (2009), Koopman, Kräussl, Lucas, and Monteiro (2009), Azizpour, Giesecke, and Schwenkler (2010), and Koopman, Lucas, and Schwaab (2011).

Figure 8: Unconditional portfolio loss densities

The figure reports unconditional (historical) loss densities for portfolios (a) to (d) as discussed in the main text. The data sample is from 1980Q1 to 2013Q3.

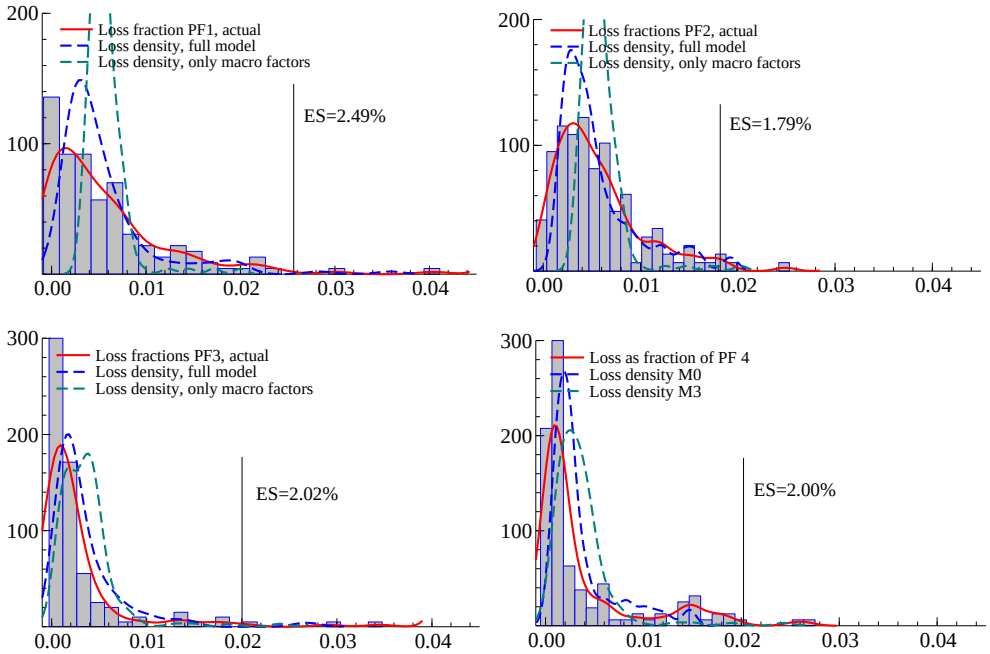


Table 3: **Portfolio losses and risk measures**

The table reports portfolio risk measures and summary statistics for losses from four different loan portfolios; see description in the main text. Sample is 1980Q1 to 2013Q3.

Portfolio	Loss density	mean loss	Var 95%	ES 95%	max loss	min loss
PF1 [most concentrated]	historical data	0.57%	1.90%	2.49%	4.04%	0%
	model-implied $\{f_t^g, f_t^m\}$	0.57%	0.80%	1.22%	2.00%	0.57%
	model with all factors	0.56%	1.76%	2.22%	3.58%	0.13%
PF2 [country concentration]	historical data	0.56%	1.52%	1.79%	2.50%	0%
	model-implied $\{f_t^g, f_t^m\}$	0.58%	0.81%	1.24%	2.03%	0.36%
	model with all factors	0.57%	1.49%	1.69%	1.94%	0.15%
PF3 [industry concentration]	historical data	0.28%	1.30%	2.02%	3.45%	0%
	model-implied $\{f_t^g, f_t^m\}$	0.37%	0.77%	1.29%	2.40%	0.09%
	model with all factors	0.36%	1.02%	1.62%	2.76%	0.05%
PF4 [fully diversified]	historical data	0.39%	1.69%	2.00%	2.66%	0%
	model-implied $\{f_t^g, f_t^m\}$	0.36%	0.72%	1.29%	2.44%	0.10%
	model with all factors	0.36%	1.14%	1.34%	1.50%	0.00%

(b) continues to be concentrated in the U.S., but is now diversified across industry sectors. Portfolio (c) contains 250 loans of 1$ each to firms from the capital goods industry in each region. Thus, it is diversified geographically, but still concentrated in its exposure to firms from one industry. Finally, portfolio (d) contains 50 loans a 1$ to firms in each of the four regions and five non-financial industry sectors. This relaxes both constraints and means that the portfolio is diversified across both industry sectors and regions

Interestingly, and perhaps counter-intuitively, more credit risk diversification across borders does not necessarily decrease portfolio default risk, even when holding marginal default probabilities - the ratings - constant. Two effects work in opposite directions. First, expanding the credit portfolio across borders decreases portfolio risk if regional (macro or frailty) factors are imperfectly correlated (or independent). Second, however, the expansion of the portfolio could involve loans to firms that load *relatively more* on the global risk factors, such as global macro, global frailty, and the industry-specific risk factors. Figure 8 suggests that the first effect dominates when moving from portfolio (a) to portfolios (b), (c), and (d). However, the second effect dominates when moving from portfolio (b) (which is diversified across industries within the U.S.) to portfolio (d) (which is in addition diversified across borders). We conclude that cross-border risk diversification is not necessarily beneficial, in particular if the initial portfolio is already somewhat diversified. A careful analysis of systematic risks is required.

6 Conclusion

We investigated the common dynamic properties of systematic default risk conditions across countries, regions, and the world. To this purpose we developed a high-dimensional, partly non-linear non-Gaussian state space model to estimate common components in firm defaults in a 22-country sample covering six broad industry groups and four economic regions in the world. The results indicate that common world factors are a first order source of default risk variation and observed default clustering, thus providing evidence for a world credit

risk cycle. Global macro and frailty dynamics limit the scope for cross-border credit risk diversification, which in turn limits the economies of scale that can be achieved by globally active lending institutions.

References

Azizpour, S., K. Giesecke, and G. Schwenkler (2010). Exploring the sources of default clustering. *Stanford University working paper*.

Bai, J. and S. Ng (2002). Determining the number of factors in approximate factor models. *Econometica 70(1)*, 191–221.

Bekaert, G., R. J. Hodrick, and X. Zhang (2009). International Stock Return Comovements. *Journal of Finance 64(6)*, 2591–2626.

Black, F. and J. C. Cox (1976). Valuing corporate securities: Some effects of bond indenture provisions. *The Journal of Finance 31*(2), 351–367.

Brauning, F. and S. J. Koopman (2013). Forecasting Macroeconomic Variables using Collapsed Dynamic Factor Analysis. *International Journal for Forecasting 29*.

Bruno, V. and H. S. Shin (2012). Capital flows, cross-border banking and global liquidity. *NBER working paper w19038*.

Ciccarelli, M. and B. Mojon (2010). Global Inflation. *Review of Economics and Statistics 92(3)*, 524–535.

Creal, D., B. Schwaab, S. J. Koopman, and A. Lucas (2013). An observation driven mixed measurement dynamic factor model with application to credit risk. *The Review of Economics and Statistics 35*.

CreditMetrics (2007). CreditMetrics (TM) - Technical Document, RiskMetrics Group. www.riskmetrics.com/pdf/dnldtechdoc/CMTD1.pdf.

Das, S., D. Duffie, N. Kapadia, and L. Saita (2007). Common Failings: How Corporate Defaults Are Correlated. *The Journal of Finance 62(1)*, 93–117(25).

Diebold, F. X., C. Li, and V. Yue (2008). Global yield curve dynamics and interactions: a dynamic Nelson-Siegel approach. *Journal of Econometrics 146(2)*, 405–43.

Duffie, D., A. Eckner, G. Horel, and L. Saita (2009). Frailty Correlated Default. *Journal of Finance 64(5)*, 2089–2123.

Durbin, J. and S. J. Koopman (1997). Monte Carlo Maximum Likelihood estimation for non-Gaussian State Space Models. *Biometrika 84(3)*, 669–684.

Durbin, J. and S. J. Koopman (2001). *Time Series Analysis by State Space Methods*. Oxford: Oxford University Press.

Durbin, J. and S. J. Koopman (2002). A simple and efficient simulation smoother for state space time series analysis. *Biometrika 89(3)*, 603–616.

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

Harvey, A. (1989). *Forecasting, Structural Time Series Models, and the Kalman Filter*. Cambridge University Press.

He, Z. and W. Xiong (2012). Rollover risk and credit risk. *Journal of Finance 67*, 391–429.

Jotikasthira, C., A. Le, and C. Lundblad (2011). Why Do Term Structures in Different Currencies Comove? *mimeo*.

Jungbacker, B., S. J. Koopman, and M. van der Wel (2011). Maximum likelihood estimation for dynamic factor models with missing data. *Journal of Economic Dynamics and Control 35*, 1358 – 1368.

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.

Koopman, S., R. Kräussl, A. Lucas, and A. Monteiro (2009). Credit cycles and macro fundamentals. *Journal of Empirical Finance 16*, 42–54.

Koopman, S. J. and A. Lucas (2008). A Non-Gaussian Panel Time Series Model for Estimating and Decomposing Default Risk. *Journal of Business and Economic Statistics 26 (4)*, 510–525.

Koopman, S. J., A. Lucas, and A. Monteiro (2008). The Multi-Stage Latent Factor Intensity Model for Credit Rating Transitions. *Journal of Econometrics 142(1)*, 399–424.

Koopman, S. J., A. Lucas, and B. Schwaab (2011). Modeling frailty correlated defaults using many macroeconomic covariates. *Journal of Econometrics 162(2)*, 312–325.

Koopman, S. J., A. Lucas, and B. Schwaab (2012). Dynamic factor models with macro, frailty, and industry effects for u.s. default counts: the credit crisis of 2008. *Journal of Business and Economic Statistics 30(4)*, 521–532.

Kose, M. A., C. Otrok, and E. Prasad (2012). Global Business Cycles: Convergence Or Decoupling? *International Economic Review 53(2)*, 511–538.

Kose, M. A., C. Otrok, and C. H. Whiteman (2003). International Business Cycles: World, Region, and Country-Specific Factors. *American Economic Review 93(4)*, 1216–1239.

Kose, M. A., C. Otrok, and C. H. Whiteman (2008). Understanding the evolution of world business cycles. *Journal of International Economics 75(1)*, 110130.

Lando, D. (2003). *Credit Risk Modelling - Theory and Applications*. Princeton University Press.

Lando, D. and M. S. Nielsen (2010). Correlation in corporate defaults: Contagion or conditional independence. *Journal of Financial Intermediation 19(3)*, 355–372.

McNeil, A. and J. Wendin (2007). Bayesian inference for generalized linear mixed models of portfolio credit risk. *Journal of Empirical Finance 14(2)*, 131–149.

McNeil, A. J., R. Frey, and P. Embrechts (2005). *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press.

Merton, R. (1974). On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *Journal of Finance 29(2)*, 449–470.

Neely, C. J. and D. E. Rapach (2011). International comovements in inflation rates and country characteristics. *Journal of International Money and Finance 30(7)*, 14711490.

Schweppe, F. (1965). Evaluation of likelihood functions for Gaussian signals. *IEEE Transactions on Information Theory 11*, 61–70.

Shephard, N. and M. K. Pitt (1997). Likelihood analysis of non-Gaussian measurement time series. *Biometrika 84*, 653–67.

Stock, J. H. and M. W. Watson (1989). New Indexes of Coincident and Leading Economic Indicators. *NBER Macroeconomics Annual*, 351–393.

Tarullo, D. K. (2010). Equipping Financial Regulators with the Tools Necessary to Monitor Systemic Risk. *Statement before the Subcommittee on Security and International Trade and Finance, Committee on Banking, Housing, and Urban Affairs, U.S. Senate, 12 February 2010*.

Wheelock, A. C. and P. W. Wilson (2012). Do Large Banks Have Lower Costs? New Estimates of Returns to Scale for U.S. Banks. *Journal of Money, Credit and Banking 44(1)*, 171199.

A1. Maximum likelihood estimation based on importance sampling

Importance sampling proceeds by finding a proposal distribution $g(f|y;\psi)$, called the importance density, which closely approximates $p(f|y;\psi)$ but has heavier tails. Assume that the conditions underlying the application of importance sampling hold, in particular that $g(f|y;\psi)$ is sufficiently close to $p(f|y;\psi)$ and simulation from $g(f|y;\psi)$ is feasible. Then a Monte Carlo estimate of the likelihood $p(y;\psi)$ can be obtained as

$$\tilde{p}(y;\psi) = g(y;\psi)\, M^{-1} \sum_{k=1}^{M} \frac{p(y|f^{(k)};\psi)}{g(y|f^{(k)};\psi)}, \qquad f^{(k)} \sim g(f|y;\psi), \tag{A.18}$$

where M is a large number of draws. Density $g(y;\psi)$ is the likelihood of an approximating model which is employed to obtain the samples $f^{(k)} \sim g(f|y;\psi)$, see below. A derivation of (A.18) is provided in the appendix A1.

For a practical implementation, the importance density $g(f|y;\psi)$ can be based on the linear Gaussian state space model

$$\tilde{y}_t = \theta_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathrm{N}(0, \tilde{H}_t), \tag{A.19}$$

where the transition equation for θ_t is the same as in the original model of interest. The pseudo-observations $\tilde{y}_t$ and covariance matrices $\tilde{H}_t$ are chosen in such a way that the distribution $g(f|y;\psi)$ implied by the approximating state space model is sufficiently close to the distribution $p(f|y;\psi)$ from the original non-Gaussian model. Shephard and Pitt (1997) and Durbin and Koopman (1997) argue that $\tilde{y}_t$ and $\tilde{H}_t$ can be uniquely chosen such that the mode and curvature at the mode of $g(f|y;\psi)$ match the mode and curvature of $p(f|y;\psi)$ for a given value of ψ.

To simulate values from the importance density $g(f|y;\psi)$, the simulation smoothing method of Durbin and Koopman (2002) can be applied to the approximating model (A.19). For a set of M draws of $g(f|y;\psi)$,

the evaluation of (A.18) relies on the computation of $p(y|f;\psi)$, $g(y|f;\psi)$ and $g(y;\psi)$. Density $p(y|f;\psi)$ is based on (9), density $g(y|f;\psi)$ is based on the Gaussian density for $y_{i,t} - \mu_{it} - \theta_{i,t} \sim \mathrm{N}(0, \sigma^2_{i,t})$ (A.19) and $g(y;\psi)$ can be computed by the Kalman filter applied to (A.19), see Schweppe (1965) and Harvey (1989).

Once an ML estimator is available for ψ, the estimation of the location of f can be based on importance sampling. It can be shown that

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

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

$$\tilde{f} = \sum_{k=1}^{M} w_k \cdot f^{(k)} \Bigg/ \sum_{k=1}^{M} w_k, \tag{A.20}$$

with $w_k = p(y|f^{(k)};\psi)/g(y|f^{(k)};\psi)$, and $f^{(k)} \sim g(f|y;\psi)$. Similarly, the standard errors s_t of $\tilde{f}_t$ can be estimated by

$$s_t^2 = \left(\sum_{k=1}^{M} w_k \cdot (f_t^{(k)})^2 \Bigg/ \sum_{k=1}^{M} w_k \right) - \tilde{f}_t^2, \tag{A.21}$$

with $\tilde{f}_t$ the tth elements of $\tilde{f}$. The availability of conditional variance estimates allows us to construct estimated standard error bands around the conditional mean of the factors.

A2. Additional results

Figure 9 reports several diagnostic checks that illustrate that the assumptions underlying the application of importance sampling are likely to be satisfied in the present application.

Figure 9: Diagnostic checks importance sampling weights

The figure reports several visual diagnostics for the importance sampling weights. The top four panels report recursive sample moments. The bottom left panel reports Hill tail index estimates for different numbers of tail observations. The bottom right panel reports the contribution of each weight (ordered) to the total sum.