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

*Author information: Bernd Schwaab, European Central Bank, Sonnemannstraße 29, 60314 Frankfurt, Germany, Email: bernd.schwaab@ecb.int. Siem Jan Koopman, Department of Econometrics, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands, Email: s.j.koopman@vu.nl. André Lucas, Department of Finance. VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands, Email: a.lucas@vu.nl. We thank the seminar and conference participants at the Bundesbank, European Central Bank, and RMI International Risk Management Conference 2014 at NUS Singapore. Lucas acknowledges support from the Dutch National Science Foundation (NWO; grant VICI453-09-005). Koopman acknowledges support from CREATES, Aarhus University, Denmark, as funded by the Danish National Research Foundation, (DNRF78). Further, Koopman and Lucas acknowledge support from the European Union Seventh Framework Programme (FP7-SSH/2007-2013, grant agreement 320270 - SYRTO). 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.

Global Credit Risk: World, Country and Industry Factors

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

Abstract

This paper investigates the dynamic properties of systematic default risk conditions for firms in different countries, from different industries and rating groups. We use a high-dimensional nonlinear non-Gaussian state space model to estimate common components in corporate defaults in a 41 country sample between 1980Q1-2014Q4, covering both the global financial crisis and euro area sovereign debt crises. We find that macro and default-specific world factors are a primary source of default clustering across countries. Defaults cluster more than what shared exposures to macro factors imply, indicating that other factors also play a significant role. For all firms, deviations of systematic default risk from macro fundamentals are correlated with net tightening bank lending standards, implying that bank credit supply and systematic default risk are inversely related.

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

JEL classification: *G21, C33*

1 Introduction

Recent studies provide evidence of many cross-country links and common global dynamics in macroeconomic fluctuations and financial asset returns.[1] Simultaneously, these common movements in macroeconomic fluctuations and asset returns are known to influence the time variation in corporate default rates; see, for example, Pesaran, Schuermann, Treutler, and Weiner (2006), Koopman, Kräussl, Lucas, and Monteiro (2009), and Giesecke, Longstaff, Schaefer, and Strebulaev (2011). Given that many macro-financial phenomena are pictured best in a global perspective, we raise the question whether the same holds true for corporate default rates. In particular, we ask ourselves whether there is a world default risk cycle? And if so, what are its statistical properties? To what extent is the world default risk cycle different from the world business cycle, which also affects defaults? When can the default risk cycle and business cycle decouple? Is such decoupling only specific to the U.S., or is it an international phenomenon? And finally, what are the implications, if any, of world default risk factors for the risk bearing capacity of internationally active financial intermediaries?

The main objective of this paper is to quantify the share of systematic default risk that can be attributed to *world* business cycle factors and default-specific factors, to infer their statistical properties, to estimate their location over time, and to assess to what extent the world default risk cycles can be decoupled from world macroeconomic conditions. Compared to previous credit risk studies that focus on the U.S. perspective, our paper provides an international perspective on default clustering. We investigate the credit experience of more than 20.000 firms from a 41-country sample covering four economic regions in the world over 35 years from 1980Q1 to 2014Q4.

In addition to the literature investigating the extent of co-movement across global macroe-

[1] For example, Kose, Otrok and Whiteman (2003, 2008), and Kose, Otrok, and Prasad (2012) document the presence of a world business cycle in macroeconomic variables, 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 factors explaining more than half of the country variances on average. Yet other research points to global common movement in international stock returns (see e.g. Bekaert, Hodrick, and Zhang (2009)), government bond yields (see e.g. Jotikasthira, Le, and Lundblad (2011)), and term structure dynamics (see e.g. Diebold, Li, and Yue (2008)).

conomic and financial market variables as mentioned above, a second strand of literature investigates why corporate defaults cluster so much over time within certain economies. For example, quarterly default probabilities for U.S. industrial firms can be an order of magnitude higher in a bust than they are in an economic boom; see Das, Duffie, Kapadia, and Saita (2007) and Koopman, Lucas, and Schwaab (2011). In general, the accurate measurement of point-in-time default hazard rates is a complicated task since not all processes that determine corporate defaults can easily be 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 compellingly made by Das et al. (2007), who apply a multitude 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 assumption, also known as the conditional independence assumption. In particular, there is substantial evidence for an additional dynamic unobserved 'frailty' risk factor and/or contagion dynamics; see McNeil and Wendin (2007), Koopman, Lucas, and Monteiro (2008), Koopman and Lucas (2008), Duffie et al. (2009), Lando and Nielsen (2010), Koopman, Lucas, and Schwaab (2011, 2012), Azizpour, Giesecke, and Schwenkler (2014), and Creal, Schwaab, Koopman, and Lucas (2014).

Understanding the sources of international default risk variation is crucial for developing robust risk models at internationally active financial intermediaries as well as for effective supervision by the appropriate authorities. In addition, studying a country (or region) in isolation can lead one to erroneously believe that the observed co-movement is specific to that country, say the U.S., when it is in fact common to a much larger group of countries. Unfortunately, data sparsity (in particular for non-U.S. firms) as well as econometric challenges have so far limited attention to single countries. These challenges include the combination of having non-Gaussian default data on the left-hand side and unobserved risk factors on the right hand side, as well as computational challenges when jointly modeling different sets of macro and default risk data from a larger number of countries. Only few

studies provide a more international perspective. For example, Claessens, Kose, and Terrones (2012) study the interactions between macroeconomic-business and financial cycles in an international context. Pesaran, Schuermann, Treutler, and Weiner (2006) study credit risk conditions in multiple countries in a unified (GVAR) framework. Finally, we refer to the RMI credit risk initiative as a noteworthy attempt to build a world-wide credit risk map from the bottom up; see, for instance, Duan and van Laere (2012).[2] All these authors, however, do not distinguish between different sets of latent default risk drivers, such as world, country, and industry factors, and do not provide a variance decomposition of default data with respect to these components. We address these issues by employing a high-dimensional dynamic factor modeling framework to disentangle common components in both international macro-financial variables and international default risk data.

From a methodological point of view, we combine the estimation frameworks of Koopman, Lucas, and Schwaab (2012) and Bräuning and Koopman (2014) to model a large cross-section of mixed-measurement observations subject to a substantial number of latent factors. Specifically, we explain how the computational difficulties can be overcome through dimensionality reduction in a preliminary first step, as in Bräuning and Koopman (2014), and the use of antithetic variables in Monte Carlo maximum likelihood evaluation for a parameter-driven mixed-measurement dynamic factor model, as introduced in Koopman et al. (2012). As an additional contribution, we show that a one-on-one correspondence exists between our empirical mixed-measurement dynamic factor model and a CreditMetrics (2007)-type multi-factor credit risk model of default dependence. This is convenient, as it allows us to establish an economic interpretation of the empirical model parameters, and to define the shares of systematic default risk variation. In our econometric framework, non-Gaussian (integer) default counts are modelled jointly with (continuous, Gaussian) macro-financial covariates and expected default frequencies (EDF) data. Considering risk data based on EDFs in addition to actual defaults is crucial, since defaults are rare for most economic regions outside the U.S.

[2]RMI is the Risk Management Institute of the National University of Singapore, `http://www.rmicri.org`.

EDF data are standard default risk measurements that are routinely used in the financial industry and credit risk literature; see for example Lando (2003), Duffie et al. (2007) and Duffie et al. (2009).

We obtain the following four main empirical findings. First, our results indicate that there is a distinct world default risk cycle that is related to, but distinct from world macro-financial cycles. We find that between 18-26% of global default risk variation is systematic, while the remainder is idiosyncratic. The share of systematic default risk is higher (39-51%) if industry-specific variation is counted as systematic. Shared exposure to global and regional macroeconomic factors explains 2-4% of total (i.e. systematic plus idiosyncratic) default risk variation across the economic regions and industry sectors considered in this paper. The remainder of systematic global default risk variation is accounted for by global default-specific (frailty) risk factors (7-18%) and regional frailty factors (1-11%). The latter is an important source of default risk clustering in some regions, but not others. Finally, industry-specific variation (17-31%) represents a significant additional source of default clustering. Industry dynamics are most pronounced for the transportation and energy, consumer goods, and retail and distribution industries.

Second, all risk factors tend to be highly persistent, with most autoregressive parameters well above 0.8 at the quarterly frequency. The frailty and industry-specific factors are particularly persistent, with autoregressive coefficients of up to 0.98. Such values imply a half-life of a shock to default risk of approximately 5 to 25 quarters. As a result, default risk conditions can decouple substantially and for an extended period of time compared to what macroeconomic and financial markets data imply, before eventually returning to their long run means.

Third, we show that the decoupling of systematic default risk from macro fundamentals is strongly related to variation in bank lending standards in all four regions. This supports economic models that have provided empirical evidence of the importance of financial intermediary behavior as a determinant of economy-wide corporate default risk; see Aoki and Nikolov (2015), Boissay, Collard, and Smets (2015), and Clerc et al. (2015). In our sample,

unusually low physical default risk conditions almost always coincide with net falling bank lending standards. This finding is intuitive: when bad risks receive ample and easy access to credit, they can avoid, or at least delay, default. Vice versa, net tightening bank lending standards coincide with higher systematic default risk. This phenomenon is also intuitive: when credit access is tight, even solvent firms have a higher risk of becoming illiquid; compare Acharya, Davydenko, and Strebulaev (2012) and He and Xiong (2012). Our global frailty factor is consistent with bank lending standards that are strongly correlated across borders, in line with correlated monetary policy cycles and 'global liquidity' conditions; see Bruno and Shin (2012) and Hoffmann, Eickmeier, and Gambacorta (2014).

Finally, given the key importance of global factors, we conclude that – perhaps counter-intuitively – more risk diversification across borders does not necessarily decrease portfolio default risk through a reduced dependence across firms. Two effects work in opposite directions. On the one hand, expanding the portfolio across borders *decreases* risk dependence if regional macro and regional default-specific factors are imperfectly correlated. On the other hand, portfolio diversification across borders can *increase* risk dependence if it involves new credit to firms that load more heavily on the world factors. Our empirical results demonstrate that this trade-off is a relevant concern. As a corollary, these results suggest that the risk bearing capacity of global lenders is not necessarily more superior to that of more regionally constrained lenders, as long as the credit portfolios of the latter are already well diversified across industry sectors and rating groups; see Wheelock and Wilson (2012) and Richter and Euffinger (2014).

The remainder of this paper is organized as follows. Section 2 introduces our global data and provides preliminary evidence for default clustering across borders. Section 3 formulates a financial framework in which default dependence is driven by multiple global, regional, and industry-specific risk factors. It also introduces our estimation methodology. Section 4 presents our key empirical results. Section 5 concludes. Technical details concerning estimation are presented in the Appendix.

2 Risk data and international default clustering

This section describes our global data and provides preliminary evidence of pronounced default clustering across borders. We consider quarterly data from three sources. First, we construct default and 'firms at risk' count data for firms from 41 countries. Second, we briefly discuss EDF-based risk indices at the region/country level. Finally, we select macroeconomic and financial time series data with the aim to capture business cycle conditions.

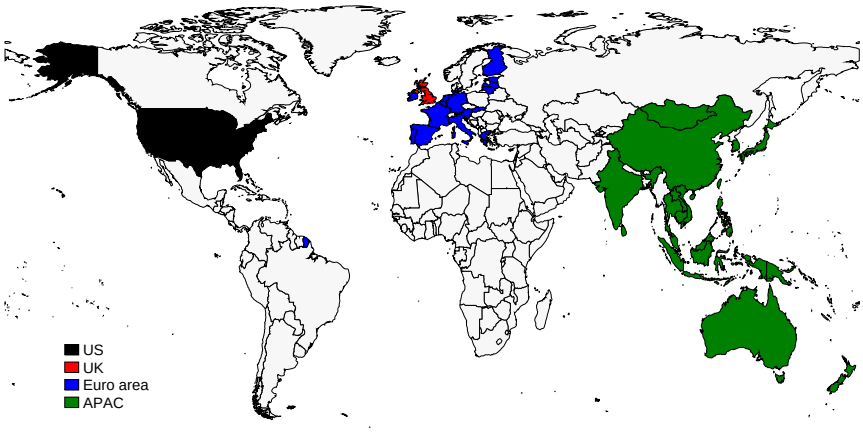
2.1 International default data

As a first panel data set, we consider default and firms-at-risk count data from Moody's extensive default and recovery database (DRD). The database contains all rating transitions and default dates for all Moody's-rated firms worldwide. Figure 1 visualizes the main regions we distinguish.[3] We focus on 35 years of quarterly data from 1980Q1 to 2014Q4. We use Moody's broad industry classification to allocate firms to six broad industry sectors: banks & other financial institutions (fin); transportation, utilities, energy & environment (tre); capital goods & manufacturing (ind); technology firms (tec); retail & distribution (ret); and, finally, consumer goods (con). When counting firms at risk and the corresponding defaults, we ensure that a firm's rating withdrawal is ignored if it is later followed by a default event. In this way, we limit the impact of strategic rating withdrawals preceding a default. We apply other standard filters; for example, we consider only the first default event when there are multiple defaults for the same firm.

We take into account data from 16.360 rated firms in the U.S., 903 firms in the U.K., 2.087 firms in euro area countries, and 1517 firms in the Asia-Pacific region. In total, we consider 20.867 firms, worldwide. Most of these firms are only active during a subset of time from 1980Q1 to 2014Q4. The corresponding number of defaults are 1660, 62, 106, and 72,

[3] Our country selection and grouping is in part motivated by the availability of EDF data to augment the count data from Moody's extensive default and recovery database (DRD). 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 1: Country sample and grouping



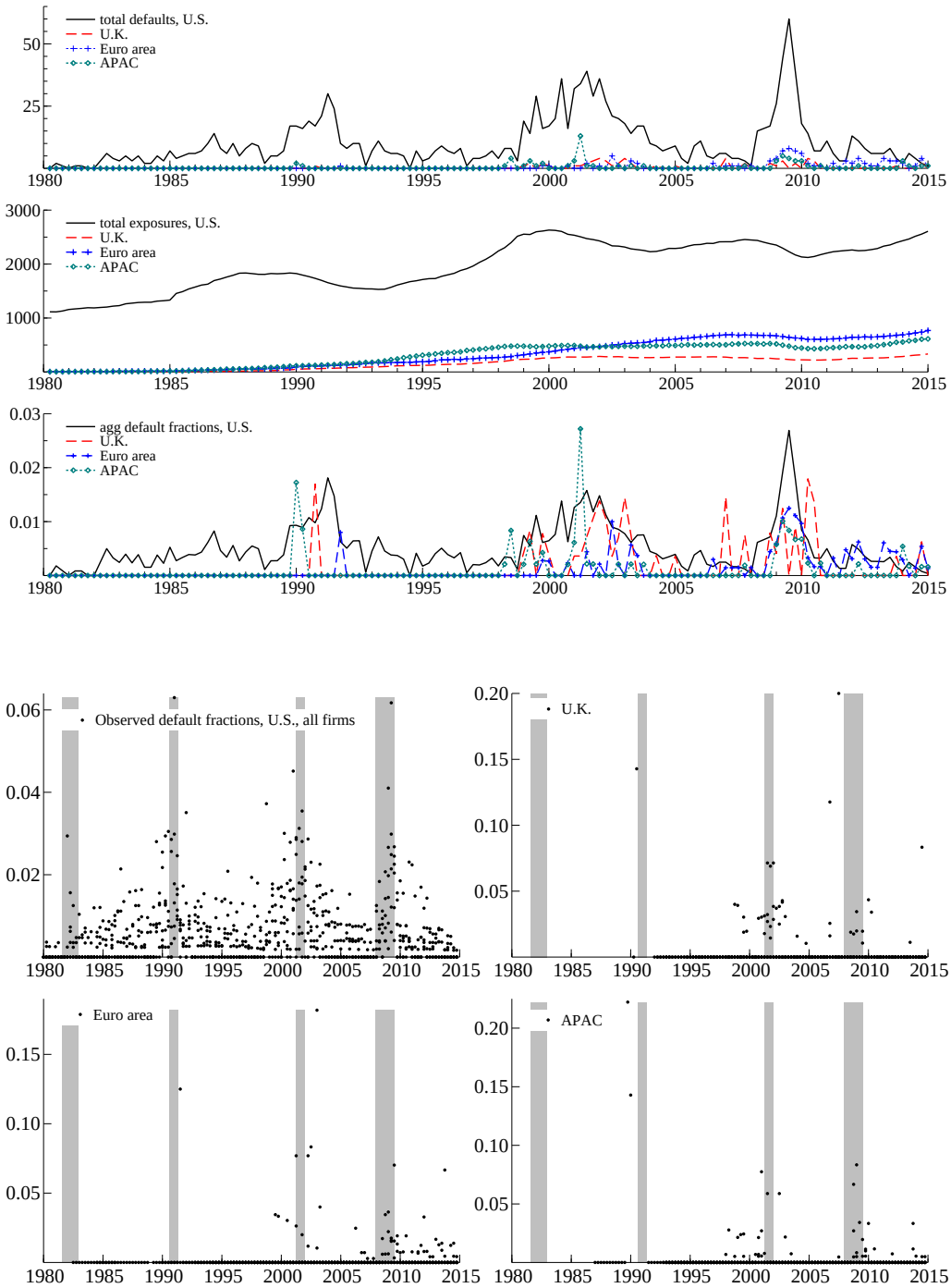
Region 1: U.S.		Region 3: Euro area		Region 4: Asia-Pacific	
U.S.A.		Austria	Latvia	Australia	Mongolia
U.S. Territories		Belgium	Lithuania	Cambodia	New Zealand
		Cyprus	Luxembourg	China	Papua New Guinea
		Estonia	Malta	Hong Kong	Philippines
		Finland	Netherlands	India	Singapore
Region 2: U.K.		France	Portugal	Indonesia	South Korea
United Kingdom		Germany	Slovakia	Japan	Sri Lanka
British Virgin Islands		Greece	Slovenia	Laos	Taiwan
Isle of Man		Ireland	Spain	Macau	Thailand
		Italy		Malaysia	Vietnam

respectively, totaling 1902 default events.

We visualize the data in Figure 2. The top panel reports the total number of defaults, the total number of firms at risk, and the respective default fractions. The bottom panels present aggregate default and firm counts, as well as observed default fractions over time for each economic region. The third panel and bottom panels in Figure 2 strongly suggest that for our global default data high default losses in one region tend to coincide with high default losses in any of the other regions. In addition to the pronounced cross-country correlation, the top panel in the figure shows that defaults also strongly cluster in the time dimension. 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.

Figure 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.



2.2 EDF risk indices

As a second set of data we consider expected default frequencies from Moody's Analytics (formerly Moody's KMV). EDFs are proprietary point-in-time forecasts of default rates, and are based on a proprietary firm value model that takes firm equity values and balance sheet information as inputs. We use one-year ahead EDF-based risk indices to augment our sparse data on actual defaults. This is crucial, as our worldwide credit risk analysis would be hard or even impossible to do without the additional information from the EDF measures, particularly when considering the systematic default rate variation for firms outside the U.S.

Figure 3 plots the EDF-based risk indices that are used in our empirical analysis below. The figure distinguishes risk data for financial (left panel) and non-financial firms (right panel) located in the U.S., U.K., euro area, and Japan. For non-financial firms, we use risk indices that are constructed as weighted averages across a large number of firms, with a firm's total assets used as weights. For financial firms, we use a risk index based on median EDF values instead, due to robustness considerations and the particularly high concentration of total assets in that industry. Finally, we use EDF-based risk indices for Japan to approximate default risk conditions for the Asia-Pacific region as a whole because data for the rest of the region is not available to us.

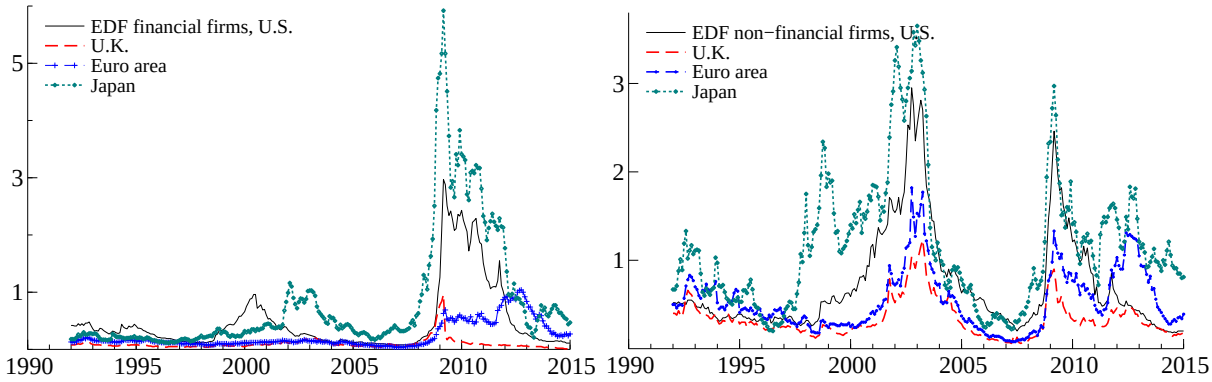
Similar to what is visible in both panels of Figure 2, the EDFs in Figure 3 also reveal a striking extent of shared variation across countries, industries, and time. Expected default rates for financial firms are high between 2001-2003, and in particular the global financial crisis between 2007-2010. Financial EDFs continue to be elevated in the euro area between 2010-2013 during the euro area sovereign debt crisis. Expected default rates for non-financial corporates are peaking, in all regions, in 1992, between 2002-2004, and from 2008-2010.

2.3 Macro-financial data

Finally, we consider macro-financial time series data that are commonly considered in the empirical credit risk literature. All macroeconomic and financial time series data are taken

Figure 3: EDF data for financial and non-financial firms

EDF-based risk indices for financial (left) and non-financial firms (right). The aggregate risk measures cover firms from the U.S., the U.K., the euro area, and Japan. The sample is from 1992Q1 to 2014Q4.



from Thomson Reuters/Datastream. Macroeconomic time series data are routinely used as conditioning variables in supervisory stress tests; see for example Tarullo (2010). Important macroeconomic covariates are real GDP growth, industrial production growth, and the unemployment rate. Key financial market variables are residential property prices, broad equity indices, and bond yields.

For the modeling of the macro data, we distinguish leading, lagging, and coincident indicators of the business cycle. We collect nine macro-financial covariates for each region/country. Two macro variables tend to lead the business cycle: the term structure spread (-5Q) and the change in a broad equity market index (-1Q).[4] Four coincident business cycle indicators include the real GDP growth rate, industrial production growth, the ISM[5] purchasing managers index (and a similar alternative for non-U.S. data), and the yearly change in the unemployment rate. Three lagging indicators include the change in 10 year government bond yields (+1Q), the change in residential property prices (+2Q), and the unemployment rate (+5Q). Stacking these macro-financial time series for each region yields a total of $9 \times 4 = 36$ macro-financial time series.

[4]The lead and lag relationships are based on the respective cross-correlation coefficients viz-a-viz the real GDP growth rate and are approximately in line with those reported in Stock and Watson (1989).

[5]ISM is the Institute for Supply Management in the U.S.

Table 1: **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%	max loss
PF1 [most concentrated]	0.58%	1.74%	2.47%	4.09%
PF2 [country concentration]	0.57%	1.45%	1.75%	2.32%
PF3 [industry concentration]	0.27%	0.97%	1.73%	3.49%
PF4 [fully diversified]	0.39%	1.51%	2.00%	2.78%

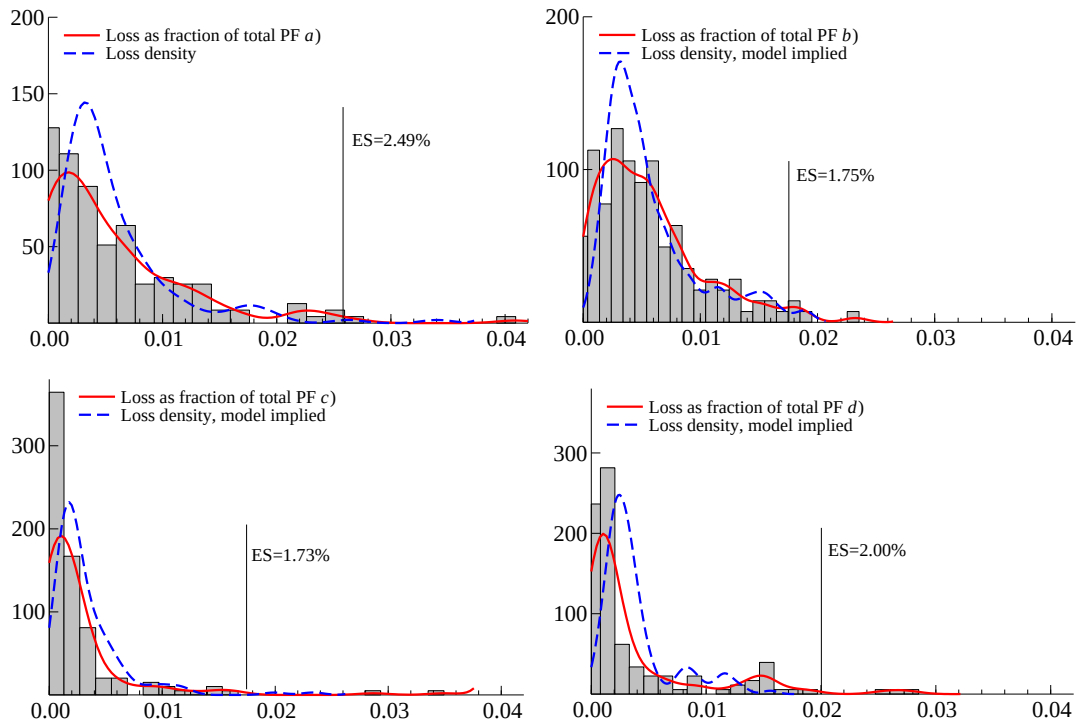
The macroeconomic variables tend to be highly correlated across countries as is well documented in the large literature on global business cycles; see, e.g., Kose, Otrok, and Whiteman (2003). A principal components analysis suggest that the first six principal components (global macro factors) account for 26%, 11%, 9.0%, 8%, 7%, and 5% of the total macro data variance, and therefore collectively explain 66% of the total variation in the macro panel.

2.4 Observed default clustering in international credit portfolios

We analyze historical default and firm count data as described in Section 2.1 to study the benefits and limits of credit risk diversification across borders. Figures 2 and 3 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 when holding marginal risks (ratings) constant.

Figure 4: 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 Section 4. The 95% expected shortfall is indicated in each panel as a vertical line.



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.

The results show that, perhaps counter-intuitively, more credit risk diversification across borders does not necessarily decrease portfolio risk. Figure 4 plots the quarterly historical loss rates for portfolios *a*) to *d*). Table 1 reports the respective portfolio risk measures for the four loan portfolios. An estimate of the 95% expected shortfall (ES) is indicated in each panel of Figure 4 as a vertical line. 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 (first moving downwards, then upwards) when portfolio risk is alternatively measured as the 95% Value-at-Risk or as the mean portfolio loss.

When global factors are the main source of international default clustering, two effects work in opposite directions. First, expanding the credit portfolio across borders *decreases* dependence across firms, and therefore decreases portfolio credit risk, if country-specific (regional) macro-financial and other factors are imperfectly correlated. 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 4 suggests that the first effect dominates when moving from portfolio *a*) to portfolios *b*), *c*), and *d*). The second effect dominates when moving from portfolios *b*) and *c*) to portfolio *d*). We conclude 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 or across countries.

While the above tradeoff is universal, the benchmark country does matter. Our empirical estimates in Section 4 suggest such an asymmetry. More specifically, it may be less beneficial for U.S. lenders to diversify exposures across borders to rated firms in Europe than it is for European (euro area and U.K.) lenders to diversify across rated firms in the U.S. In the

latter case, lenders benefit in two ways. First, regional factors are imperfectly correlated. Second, U.S. firms tend to load less heavily on global factors. We discuss these issues in more detail in our empirical study.

3 The modeling framework

3.1 A multi-factor model of default risk dependence

This section presents a multi-factor financial framework for dependent defaults. Our framework is simple but sufficiently flexible to allow us to disentangle, quantify, and test which share of cross-country default dependence is due to world, country, and industry-specific risk factors. Our framework is similar to the well-known CreditMetrics (2007) model, which is a standard in the financial industry. Importantly, the financial framework presented here is closely related to a latent dynamic factor model, which we fit to the data in Section 4. The close relationship between the two models allows us to establish an economic interpretation of the model parameters and systemic default risk shares by mapping the parameters of the econometric model back to those of the financial model.

In the special case of a standard static one-factor credit risk model for dependent defaults, as in Lando (2003), 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 $0 < \rho_i < 1$ weighs 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 2×2 unit matrix. 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 to include a more elaborate version for the asset value V_{it} of firm i at time t, and is given by

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

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

Macroeconomic risk factors are either world factors and common to all countries (f_t^g), or country/region-specific (f_t^m) and thus common only to firms in a particular country. Analogously, the frailty factors are either world factors and common to all firms (f_t^c), or country/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 risk cycle from systematic macro-financial conditions. Industry-specific risk factors f_t^i are common to firms from the same industry, 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 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 $i = 1, \ldots, K$, 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 exogenous default threshold τ_i, see Merton (1974) and Longstaff and Schwartz (1995). 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 systematic risk factors, while idiosyncratic

(firm-specific) risk is captured by ϵ_{it}. The default threshold τ_i may depend on the firm's current rating, headquarter location, and industry sector. For firms which have not defaulted yet, a default occurs when $V_{it} < \tau_i$ or, as implied by (1), when

$$\epsilon_{it} < \frac{\tau_i - w_i' f_t}{\sqrt{1 - w_i' w_i}}.$$

The conditional default probability is given by

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

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).

Our empirical analysis considers a setting where the firms ($i = 1, \ldots, K$) are pooled into groups ($j = 1, \ldots, J$) according to headquarter location, industry sector, and current rating. We assume that the firms in each group are sufficiently similar (homogenous) such that the same risk factors and risk factor loadings apply. 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 modeled 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. Our final model reads

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

where λ_j and $\vartheta_j = \left(\alpha_j', \beta_j', \gamma_j', \delta_j', \varepsilon_j'\right)'$ 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).

3.2 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 group.

Interestingly, and useful for our purposes, there is a one-to-one correspondence between the model parameters in (1) and the reduced form coefficients in (5). If ϵ_{it} is logistically distributed, then the log-odds ratio $\theta_{jt} = \log(\pi_{jt}) - \log(1 - \pi_{jt})$ from (5) also denotes the canonical parameter of the binomial distribution. It can be easily verified that for any firm i that belongs to group j, we have

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

where $\kappa_j = \tilde{\omega}_j/(1+\tilde{\omega}_j)$, and $\tilde{\omega}_j = \alpha_j'\alpha_j + \beta_j'\beta_j + \gamma_j'\gamma_j + \delta_j'\delta_j + \varepsilon_j'\varepsilon_j$. A related simpler expression is derived in Koopman and Lucas (2008) in the context of a univariate risk factor. By contrast, the current formulation allows for multiple groups of vector-valued risk factors. The restriction $w_i'w_i \leq 1$ from the firm value model (1) is always satisfied (see the expression for κ_j) and does not need to be imposed during the estimation stage.

We use the above correspondence between the firm-value model and statistical model parameters when assessing the systematic default risk of firms from different regions and industry sectors. Specifically, we define the systematic risk of firm i as the variance of its systematic risk component,

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

where $w_i = (a_i', b_i', c_i', d_i', e_i')'$ is introduced and discussed in and below equation (1). Since $\text{Var}[V_{it}] = 1$, (6) also denotes the *share of total default risk* that is systematic, or non-diversifiable. In our empirical study below, we also report

$$\text{Var}[V_{it} \mid \epsilon_{it}, f_t^i] = a_i' a_i + \ldots + d_i' d_i, \tag{7}$$

which treats industry-specific variations as idiosyncratic effects that can be diversified.

3.3 Data structure and combination

This section explains how our high-dimensional data are combined in a mixed-measurement dynamic factor model. Our initial high-dimensional mixed-measurement data vector $(x_t', y_t', z_t')'$ has three parts

$$x_t = (x_{1,1,t}, \ldots, x_{1,N,t}, \ldots, x_{R,1,t}, \ldots, x_{R,N,t})', \tag{8}$$

$$y_t = (y_{1,1,t}, \ldots, y_{1,J,t}, \ldots, y_{R,1,t}, \ldots, y_{R,J,t})', \tag{9}$$

$$z_t = (z_{1,t}, \ldots, z_{S,t})', \tag{10}$$

where $x_{r,n,t}$ represents the nth, $n = 1, \ldots, 9$, macroeconomic or financial markets variable for region $r = 1, \ldots, 4$; $y_{r,j,t}$ is the number of defaults between times t and $t+1$ for economic region r and cross-sectional group $j = 1, \ldots, J$; and $z_{s,t}$ is the expected default frequency (EDF) at a one-year ahead horizon for some set of firms $s = 1, \ldots, S$, all measured at time $t = 1, \ldots, T$. The EDF sets are constructed on a somewhat ad-hoc basis depending on the availability of the EDF data as discussed in Section 2.2. The sector and region can be identified from each EDF set. In other words, the index j can be uniquely determined from the index s, and vice-versa.

As a result, the model includes various 'standard' macro and EDF variables that we will consider to be conditionally normally distributed. However, the model also includes (integer) default count variables in vector y_t. The data panel consisting of (x_t, y_t, z_t), for $t = 1, \ldots, T$,

is typically unbalanced. It implies that variables may not be observed for all time indices t. For example, the EDF data z_t starts to become available only from 1992Q1 onwards.

The cross-sectional dimension of the data vector implied by (8) to (10) is prohibitively large for any worldwide credit risk model. For example, 36 macro data series and 5 common macro factors would already imply 180 coefficients that need to be estimated numerically by the method of maximum likelihood. For this practical reason, we first collapse our macro panel data to smaller dimensions, and consider EDF-based risk indices at the country/region level instead of firm-specific input data.

We proceed in three steps. First, we assume that the standard approximate factor analysis as used in Stock and Watson (2002)) can also be adopted for our macro data as well. The static factor analysis can be based on the multivariate model representation

$$x_t = \Lambda_g F_{g,t} + u_t, \qquad t = 1, \ldots, T,$$

where $F_{g,t}$ are global macro factors, Λ_g are the respective factor loadings, and u_t are residual terms. The dimension of the vector $F_{g,t}$ represents the number of factors r. From the results in, for example, Lawley and Maxwell (1971), the estimated factors can be computed as

$$\hat{F}_{g,t} = \hat{\Lambda}_g' x_t, \quad \text{and} \quad \hat{u}_t = x_t - \hat{\Lambda}_g \hat{F}_{g,t}, \tag{11}$$

where $\hat{F}_{g,t}$ are the first r principal components of all macro panel data x_t. The columns of the "estimated" loading matrix $\hat{\Lambda}_g$ consists of the first r eigenvectors that correspond to the r largest ordered eigenvalues of $X'X$, where $X' = (x_1, \ldots, x_T)$.

Second, we obtain estimates of regional macro factors from the residual variation in $\hat{u}_t$. We use the same method based on principal components as described above. We extract four regional macro factors, one for each region, from the four subsets of residuals $\hat{u}_t = (\hat{u}_{1,t}', \ldots, \hat{u}_{4,t}')'$, that is

$$\hat{u}_{r,t} = \Lambda_r F_{r,t} + v_{r,t}, \qquad r = 1, \ldots, 4, \tag{12}$$

where $F_{r,t}$ is then interpreted as the region-specific macro factor. The principal components from this analysis are given by $\hat{F}_{r,t} = \hat{\Lambda}_r' u_{r,t}$ for $r = 1, \ldots, 4$. The four regional factors are stacked into $\hat{F}_{m,t} = \left(\hat{F}_{1,t}, \ldots, \hat{F}_{4,t}\right)'$. This method of estimating regional macro factors may well be overly simple and suffer from the problem that we attribute some of the regional macro variation to the global macro factors; see Moench, Ng, and Potter (2013). For this reason we do not distinguish between world and regional macro variation when reporting the systematic default risk shares further below.

Finally, we obtain one-quarter ahead expected default probabilities from annual EDF data as $\hat{z}_{r,s,t} = 1-(1-z_{r,s,t})^{1/4}$. Quarterly log-odds ratios are calculated as $\hat{\theta}_{s,t}^{\text{EDF}} = \log(\hat{z}_{s,t}) - \log(1-\hat{z}_{s,t})$ and are collected as $\hat{\theta}_t^{\text{EDF}} = \left(\hat{\theta}_{1,t}^{\text{EDF}}, \ldots, \hat{\theta}_{S,t}^{\text{EDF}}\right)$.

The transformed and collapsed data vector is given by

$$\mathcal{Y}_t = \left(\hat{F}_{g,t}', \hat{F}_{m,t}', y_t', \hat{\theta}_t^{\text{EDF}}\right)', \qquad t = 1, \ldots, T. \tag{13}$$

While the cross-sectional dimension of the original data (8) — (10) is prohibitively large, the cross-sectional dimension of collapsed data (13) is tractable.

The new measurement equation for our lower dimensional model is given by

$$\begin{aligned}
\hat{F}_{g,t} &= f_t^g + e_t^g, \qquad e_t^g \sim \text{N}(0, \Sigma_g), \\
\hat{F}_{m,t} &= f_t^m + e_t^m, \qquad e_t^m \sim \text{N}(0, \Sigma_m), \\
p(y_{j,t} \mid f_t) &\sim \text{Bin}(\theta_{j,t}, k_{j,t}), \\
\hat{\theta}_{s,t}^{\text{EDF}} &= \mu_s + \theta_{j,t} + e_{s,t}^z, \qquad e_{s,t}^z \sim \text{N}(0, \Sigma_z),
\end{aligned} \tag{14}$$

where the index j in the last model equation can be uniquely determined from the index s. The log-odds ratio $\theta_{j,t} = \log(\pi_{j,t}) - \log(1-\pi_{j,t})$ is also the canonical parameter of the Binomial distribution (see McCullagh and Nelder (1989)), $k_{j,t}$ is the number of firms at risk at the beginning of period t, and μ_s is the vector of unconditional means of the respective quarterly log-odds of default from EDF measures. We collect the risk factors, including

f_t^g, f_t^m and $\theta_{j,t}$, into the $m \times 1$ vector f_t and assume it is subject to the stationary vector autoregressive process

$$f_{t+1} = \Phi f_t + \eta_t, \qquad \eta_t \sim \mathrm{N}(0, \Sigma_\eta), \qquad t = 1, \ldots, T, \tag{15}$$

with the initial condition $f_1 \sim \mathrm{N}(0, \Sigma_f)$. The coefficient matrix Φ and the variance matrix Σ_η are assumed fixed and unknown. The disturbance vectors η_t are serially uncorrelated. Stationarity implies that the roots of the equation $|\mathrm{I} - \Phi z| = 0$ are outside the unit circle. Furthermore, the unconditional variance matrix Σ_f is implied by the dynamic process and is a function of Φ and Σ_η.

We stress that the measurement equation (14) treats the macro factors from (11) and (12) and the EDF forecasts as noisy estimates from a preliminary first step. As a result, both macro factors and EDFs are subject to measurement error; see also Bräuning and Koopman (2014). The parameters of the diagonal measurement error variance matrices Σ_g, Σ_m, Σ_z are estimated simultaneously with all other parameters. This feature is novel with respect to the modeling frameworks presented in Koopman, Lucas, and Schwaab (2011, 2012) and Creal et al. (2014). In the present setup, both the information content from (continuous) EDF data – via $\hat{\theta}_t^{\mathrm{EDF}}$ – 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 default probabilities π_{jt}.

3.4 Parameter and risk factor estimation

The joint modeling of (discrete) default count data on the one hand and (continuous) macro-financial and EDF data on the other hand implies that a parameter-driven mixed-measurement dynamic factor model (MM-DFM) is appropriate. All estimation details are relegated to the Appendix.

For each evaluation of the log-likelihood, we need to integrate out many latent factors from their joint density with the mixed-measurement observations. The estimation approach

put forward in Koopman, Lucas, and Schwaab (2011, 2012) is challenging within this high-dimensional setting. For example, the importance sampling weights may not have a finite variance in our empirical application, which is a necessary condition for the methodology to work and to obtain consistent and asymptotically normal parameter estimates. However, this challenge can be partly overcome by including more antithetic variables to balance the simulations for location and scale as suggested by Durbin and Koopman (2000) and further explored in detail by Durbin and Koopman (2012, p. 265-266). This solution has made our procedure feasible, even in this high-dimensional setting.

4 Main empirical results

In our empirical study, we analyze the credit exposures of more than 20.000 firms from 41 countries in four economic regions of the world during a time period of 35 years, from 1980Q1 to 2014Q4. The headquarter location of the company determines the country of the firm which is further identified by its current rating category and its industry sector. The main objective is to quantify the share of systematic default risk that can be attributed to *world* business cycle factors and default-specific factors. The analysis may assess to what extent the world default risk cycles can be decoupled from world macroeconomic conditions.

4.1 Model specification

For the selection of the number of factors we rely on likelihood-based information criteria (IC). The panel information criteria of Bai and Ng (2002) suggest two or three common factors for the global macro data. We select five global macro-financial factors f_t^g to be very conservative and not to bias our results towards attributing too much variation to default-specific (frailty) factors when in fact they are due to macro factors. We also include four additional region/country-specific macro factors, one for each region, to ensure that we do not miss regional macroeconomic variation that may matter for the respective regional default rates.

Allowing for one default-specific frailty factor f_t^c is standard in the literature, see for example Duffie et al. (2009) and Azizpour et al. (2014). We further include four region-specific frailty factors f_t^d, one for each region. Finally, we select six additional industry-specific factors f_t^i that affect firms from the same industry sector. Such industry factors capture (global) industry-specific developments as well as possible contagion through up- and downstream business links, see Lang and Stulz (1992) and Acharya, Bharath, and Srinivasan (2007), and have been included in earlier models; see for example Koopman et al. (2012).

Regarding risk factor loadings, all firms load on global factors f_t^g and f_t^c with region-specific factor loadings. This means that all firms are subject to these risk factors, but to different extents. Ratings affect the baseline (unconditional) default hazard rates but not the factor loadings. While somewhat restrictive, this specification is parsimonious and remains sufficiently flexible to accommodate most of the heterogeneity observed in the cross section. In particular, it allows us to focus on the commonalities and differences in the share of systematic default risk that is explained by world, country and industry factors.

4.2 Parameter and risk factor estimates

Table 2 reports model parameter estimates. All sets of risk factors – macro, frailty, as well as industry-specific – contribute towards explaining corporate default clustering within and across countries. Importantly, defaults from all regions load on macro factors (in particular the first one). The region-specific macro factors are overall relatively less important. This is intuitive, since much of the regional macro variation is already accounted for by the global macro factors and most of the firms that request a rating are internationally active. In addition, this finding may suggest some sample selection, in that non-U.S. firms that request to be rated by Moody's also tend to be internationally active, and more so than their U.S. counterparts. Non-U.S. firms from the euro area and APAC region also differ from U.S. firms in their unconditional hazard rates $\lambda_{r,j}$, see the left column in Table 2. Again, this may reflect some sample selection, in that these non-U.S. firms are sufficiently large and of a high credit quality to access capital markets rather than refinance themselves via financial

Table 2: **Parameter estimates**

We report the maximum likelihood estimates of selected coefficients in the specification of the log-odds ratio (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 f_t^c, region-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 1980Q1 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
$\bar{\lambda}_0$	-4.82	0.00
$\bar{\lambda}_{1,fin}$	0.07	0.73
$\bar{\lambda}_{1,tre}$	0.07	0.80
$\bar{\lambda}_{1,tec}$	-0.23	0.39
$\bar{\lambda}_{1,ret}$	0.14	0.67
$\bar{\lambda}_{1,con}$	-0.13	0.72
$\bar{\lambda}_{2,IG}$	-3.70	0.00
$\bar{\lambda}_{3,UK}$	0.30	0.02
$\bar{\lambda}_{3,EA}$	-0.21	0.05
$\bar{\lambda}_{3,AP}$	-0.41	0.00

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	0.83	0.00
$\bar{\alpha}_{4,0}$	0.00	0.93
$\bar{\alpha}_{4,1,UK}$	0.03	0.37
$\bar{\alpha}_{4,1,EA}$	0.07	0.04
$\bar{\alpha}_{4,1,AP}$	0.04	0.52
ϕ_5^g	0.84	0.00
$\bar{\alpha}_{5,0}$	0.03	0.37
$\bar{\alpha}_{5,1,UK}$	-0.00	0.99
$\bar{\alpha}_{5,1,EA}$	0.04	0.32
$\bar{\alpha}_{5,1,AP}$	0.06	0.29

Global frailty f_t^c		
par	val	p-val
ϕ^c	0.95	0.00
$\bar{\gamma}_0$	0.43	0.00
$\bar{\gamma}_{1,UK}$	0.08	0.19
$\bar{\gamma}_{1,EA}$	0.12	0.08
$\bar{\gamma}_{1,AP}$	-0.05	0.58
Regional frailty f_t^d		
ϕ_{US}^d	0.96	0.00
$\bar{\delta}_{0,US}$	0.43	0.01
ϕ_{UK}^d	0.96	0.00
$\bar{\delta}_{0,UK}$	0.17	0.21
ϕ_{EA}^d	0.97	0.00
$\bar{\delta}_{0,EA}$	0.21	0.13
ϕ_{AP}^d	0.87	0.00
$\bar{\delta}_{0,AP}$	0.41	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
$\bar{\alpha}_{1,1,UK}$	-0.00	0.90
$\bar{\alpha}_{1,1,EA}$	-0.10	0.03
$\bar{\alpha}_{1,1,AP}$	-0.06	0.43
ϕ_2^g	0.84	0.00
$\bar{\alpha}_{2,0}$	-0.02	0.65
$\bar{\alpha}_{2,1,UK}$	-0.03	0.47
$\bar{\alpha}_{2,1,EA}$	-0.06	0.08
$\bar{\alpha}_{2,1,AP}$	-0.05	0.41
ϕ_3^g	0.84	0.00
$\bar{\alpha}_{3,0}$	-0.07	0.14
$\bar{\alpha}_{3,1,UK}$	0.08	0.09
$\bar{\alpha}_{3,1,EA}$	0.08	0.08
$\bar{\alpha}_{3,1,AP}$	0.10	0.20

Regional macros f_t^m		
par	val	p-val
ϕ_{US}^m	0.41	0.00
$\bar{\beta}_{0,US}$	0.04	0.01
ϕ_{UK}^m	0.79	0.00
$\bar{\beta}_{0,UK}$	-0.05	0.18
ϕ_{EA}^m	0.82	0.00
$\bar{\beta}_{0,EA}$	0.01	0.52
ϕ_{AP}^m	0.87	0.00
$\bar{\beta}_{0,AP}$	-0.12	0.13

Industry factors f_t^i		
par	val	p-val
ϕ_{fin}^i	0.97	0.00
$\bar{\epsilon}_{fin}$	0.53	0.00
ϕ_{tre}^i	0.83	0.00
$\bar{\epsilon}_{tre}$	0.78	0.00
ϕ_{ind}^i	0.84	0.00
$\bar{\epsilon}_{ind}$	0.58	0.00
ϕ_{tec}^i	0.90	0.00
$\bar{\epsilon}_{tec}$	0.55	0.00
ϕ_{ret}^i	0.93	0.00
$\bar{\epsilon}_{ret}$	0.61	0.00
ϕ_{con}^i	0.92	0.00
$\bar{\epsilon}_{con}$	0.76	0.00

intermediaries.

While the common variation in defaults implied by shared exposure to macro factors is significant and important, it is not sufficient. The global frailty factor is found to be a significant determinant of default rates in all regions. In addition, it tends to load slightly more strongly on some non-U.S. data than on U.S. data (although the statistical evidence is not high). All loadings on industry-specific risk factors are significant. Industry-specific variation is most important for firms from the transportation & energy (tre) sector, probably reflecting their shared exposure to oil price developments.

Our finding of significant frailty effects, while in line with most credit risk literature, is somewhat at odds with studies that attribute somewhat less importance to such factors, see for example Lando and Nielsen (2010), and Duan, Sun, and Wang (2012). We stress that the importance of frailty factors depends on right hand side conditioning variables, in particular firm specific information. Firm-specific covariates such as equity returns, volatility, and leverage are often found to be important predictors of default, see Vassalou and Xing (2004), Duffie et al. (2007), and Duffie et al. (2009). We acknowledge that ratings alone may not provide sufficient statistics for future default, and that our frailty factors may in part reflect this fact.[6] We argue in Section 4.4, however, that missing firm-specific effects are unlikely to provide a complete explanation.

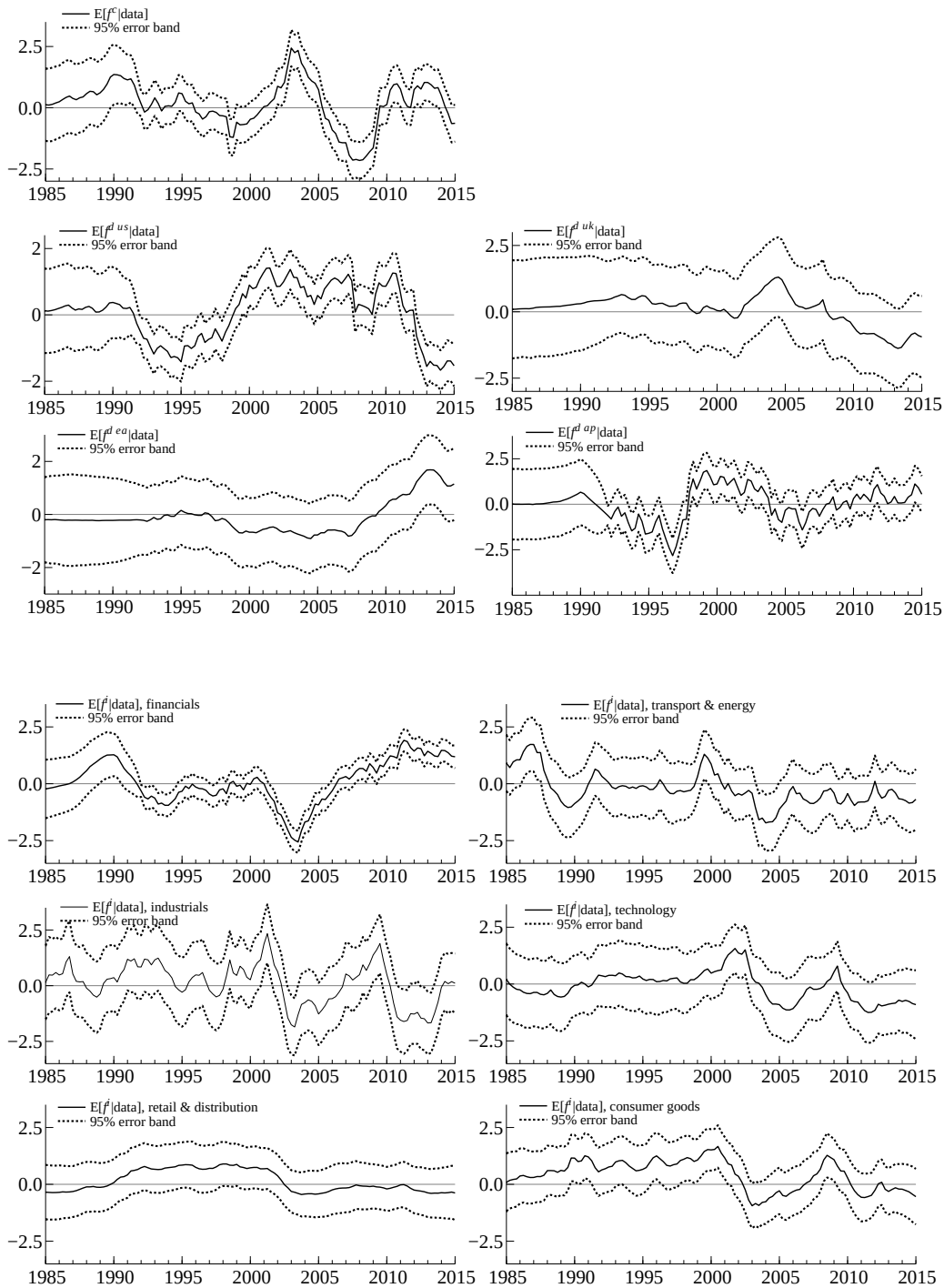
All default risk factors tend to be highly persistent. Most autoregressive parameters are well above 0.8 at the quarterly frequency. The frailty and industry-specific factors are particularly persistent, with autoregressive coefficients of up to 0.98. Such values imply a half-live of a shock to default risk of approximately 5-25 quarters.

Figure 5 plots the conditional mean estimates of the global and regional frailty factors in the top panel, as well as six industry-specific factors in the bottom panel. The evolution of the world frailty risk factor (top panel) suggests that worldwide excess default clustering

[6]Our modeling framework considers groups of homogenous firms rather than individual firms, and considers data of more than 20.000 firms, worldwide. As a result it is hard, if not impossible, to include firm-specific information beyond rating classes, geography, and industry sectors. In this regard we, again, refer to the RMI credit risk initiative as a noteworthy attempt to build a world-wide credit risk map from the bottom up, see `http://www.rmicri.org`.

Figure 5: Global frailty and industry factors

The top panel reports the conditional mean estimates of the global frailty and region-specific frailty factors. The bottom panel plots the conditional mean estimates of the six industry-specific factors.



was most pronounced during the early 1990s, as well as between 2002-03. Before the global financial crisis, default risk conditions were significantly below than what was implied by macro fundamentals during 2006-2007. This pattern already suggests that the frailty factor may be related to the behaviour of financial intermediaries: non-financial firms experience higher default stress as credit access dried up following the 1991 and 2001 economic contractions. At the same time, non-financial firms appeared to have easy access to credit in the years leading up to the global financial crisis.

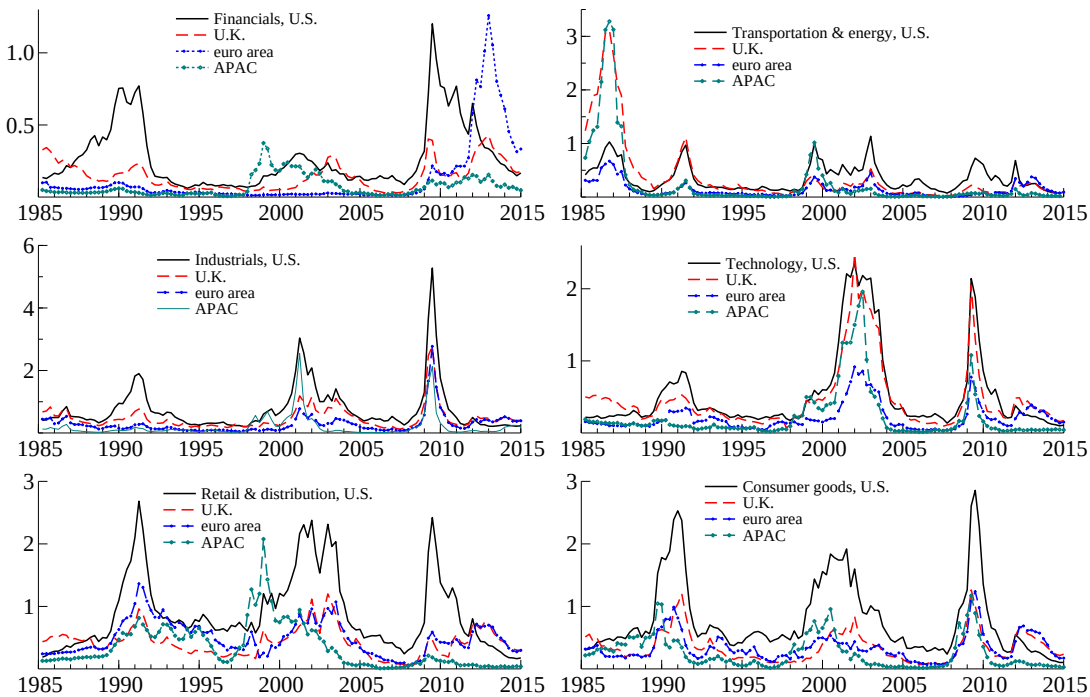
The world frailty risk factor (top panel) is quite different from the U.S. frailty factor reported in Duffie, Eckner, Horel, and Saita (2009) and Koopman, Lucas, and Schwaab (2012). Indeed, U.S. firms load significantly on their own regional frailty factor (second panel in Figure 5). World and U.S. systematic default risk are related but do not coincide.

Some evidence of additional default clustering due to regional default-specific factors is also found for firms located in the Asia-Pacific region. The loading parameters on regional frailty factors are small and insignificant for firms from the U.K. and the euro area. Instead, these firms load more heavily on the global frailty factor.

Shared exposure to volatile macroeconomic and default-specific factors implies that default hazard rates can vary substantially over time. Figure 6 plots the respective estimates of time-varying quarterly default rates for six industry sectors and four regions. Default probabilities are particularly volatile for industrial firms. U.S. financial defaults cluster in particular during the savings and loans crisis between 1986-1990 (in the U.S.), after the burst of the dot-com bubble between 2001-2002, and during the financial crisis up to the federal spending crisis between 2008-2013. In addition, financial sector firms located in the euro area suffered particularly during the euro area sovereign debt crisis between 2010-12, when default hazard rates were substantially above their historical average. Default risk for industrial sector firms is particularly high between 1990-1991, 2001-2002, and 2008-09, in line with the business cycle contractions during these periods. Not surprisingly, the bursting of the dot-com bubble is particularly visible for the implied default rates of technology sector firms between 2001-2002.

Figure 6: Global default hazard rates at the industry level

Each panel plots the model-implied time-varying default rate (one quarter ahead default probability in percent) for a specific industry sector. The panels refer to financial firms (top left), transportation & energy (top right), industrial (middle left), technology (middle right), retail & distribution (bottom left), and consumer goods (bottom right). Each panel distinguishes firms from the U.S., U.K., euro area, and Asia-Pacific region.



4.3 Variance decomposition

This section attributes the time series variation in the systematic default risk of firms from different industries and countries to different sets of systematic and idiosyncratic risk drivers, using the firm-value framework from Section 3. Table 3 presents the estimated risk shares.

We focus on four main empirical findings. First, we find that between 18-26% of global default risk variation is systematic, while the remainder is idiosyncratic. The share of systematic default risk is higher (39-51%) if we count industry-specific variation as systematic.

Second, rated firms from different countries tend to default together to some extent simply because of their shared exposure to the international business cycle and related macroeconomic conditions. The shared exposure to global macro factors, however, only explains 2-4% of total (i.e., systematic plus idiosyncratic) default risk variation.

Third, our (global and regional) frailty factors explain a large share of international default risk variation. A global default-specific (frailty) risk factor accounts for 7-18% of total default risk. Regional frailty factors (1-11%) are an important source of default risk clustering in some regions, particularly the U.S. and Asia-Pacific regions. Overall, our frailty factors account for a larger share of default risk than our (nine) macro factors, in all regions. Excess default clustering is thus also an international phenomenon.

As a corollary, default risk conditions can decouple substantially and for an extended period of time from what is implied by macroeconomic data, before eventually returning to their long run means. There is a distinct world default risk cycle that is related to but distinct from world macro-financial cycles. It is important to acknowledge the existence of such a cycle when designing a macro-prudential policy framework for the financial cycle.

Finally, industry-specific variation is a significant additional source of default clustering. Industry-specific factors explain between 17-31% of total default risk variation. Industry factors are most pronounced for the transportation & energy, consumer goods, and retail & distribution industries. Arguably, industry effects may be classified as non-systematic and diversifiable in a large portfolio that is sufficiently spread across industries.

Table 3: **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 2014Q4.

Reg.	Ind.	f_t^{g+m} $[a_i'a_i]$	f_t^c $[c_i'c_i]$	f_t^d $[d_i'd_i]$	f_t^i $[e_i'e_i]$	$\mathrm{Var}[V_{it} \mid \varepsilon_{it}, f_t^i]$ $a_i'a_i + \ldots + d_i'd_i$	$\mathrm{Var}[V_{it} \mid \varepsilon_{it}]$ $= w_i'w_i$
US	fin	4.1%	10.8%	10.7%	16.5%	25.6%	42.2%
US	tre	3.5%	9.1%	9.0%	29.8%	21.6%	51.3%
US	ind	4.0%	10.5%	10.4%	18.9%	24.9%	43.8%
US	tec	4.1%	10.7%	10.6%	17.3%	25.4%	42.7%
US	red	3.9%	10.3%	10.2%	20.3%	24.5%	44.8%
US	con	3.5%	9.3%	9.1%	28.7%	21.9%	50.6%
UK	fin	4.3%	16.0%	1.7%	17.3%	22.0%	39.3%
UK	tre	3.6%	13.4%	1.4%	31.0%	18.4%	49.4%
UK	ind	4.1%	15.6%	1.7%	19.8%	21.4%	41.1%
UK	tec	4.2%	15.9%	1.7%	18.1%	21.8%	39.9%
UK	red	4.1%	15.3%	1.6%	21.2%	21.0%	42.2%
UK	con	3.6%	13.6%	1.4%	29.9%	18.7%	48.6%
EA	fin	2.3%	18.1%	2.6%	17.1%	23.1%	40.2%
EA	tre	1.9%	15.2%	2.2%	30.6%	19.3%	49.9%
EA	ind	2.2%	17.6%	2.6%	19.5%	22.4%	41.9%
EA	tec	2.3%	18.0%	2.6%	17.8%	22.8%	40.7%
EA	red	2.2%	17.3%	2.5%	20.9%	22.0%	42.9%
EA	con	2.0%	15.4%	2.2%	29.6%	19.6%	49.2%
AP	fin	4.1%	8.7%	10.2%	17.1%	22.9%	40.1%
AP	tre	3.4%	7.3%	8.5%	30.7%	19.2%	49.9%
AP	ind	3.9%	8.5%	9.9%	19.5%	22.3%	41.8%
AP	tec	4.0%	8.6%	10.1%	17.9%	22.7%	40.6%
AP	red	3.9%	8.3%	9.7%	21.0%	21.9%	42.8%
AP	con	3.4%	7.4%	8.6%	29.6%	19.5%	49.1%

The current model with frailty components can also readily be applied to study risk diversification. As we have shown in in Figure 4, in line with our preliminary data analysis, global or industry diversification of the loan portfolio does not automatically reduce risk. The risk reduction obtained by new exposures to imperfectly correlated industry or region-specific frailty factors because of new loan commitments may be off-set by a larger exposure to the common, global frailty and macro factors due to the same new commitments. Such questions are harder, if not impossible to answer without the current model set-up and factor decomposition.

4.4 Global credit risk decoupling from macro fundamentals

We documented that global and regional frailty factors explain a large share of systematic default risk variation. This section demonstrates that deviations of systematic default risk from what is implied by the macro fundamentals can, to a substantial extent, be traced back to variations in international bank lending standards.

Figure 7: Credit risk deviations from fundamentals

Deviations of systematic default risk from macro-financial fundamentals for financial (left) and industrial (right) firms. The risk deviations are obtained according to (16). Shaded areas are NBER recession dates for the U.S.

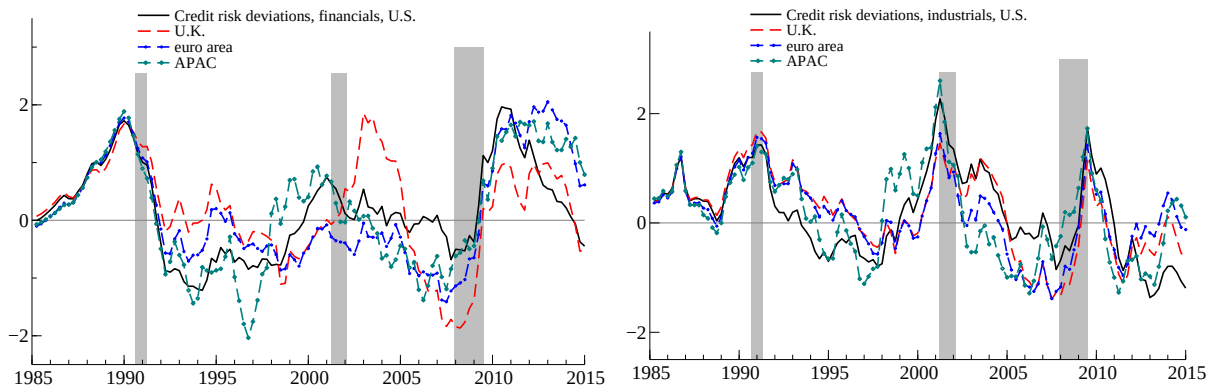


Figure 7 plots systematic default risk deviations from fundamentals. These are measured as the sum of the global, regional, and industry frailty factors times their respective loadings,

standardized by the square root of their unconditional variance,

$$\mathrm{CRD}_{r,j} = \frac{\gamma_j' f_t^c + \delta_j' f_t^d + \epsilon_j' f_t^i}{\sqrt{\gamma_j'\gamma_j + \delta_j'\delta_j + \epsilon_j'\epsilon_j}}, \tag{16}$$

and are unconditionally standard normally distributed as a result. Figure 7 focuses on financial (left) and industrial firms (right) in each region. Risk conditions can decouple significantly and persistently from the risk levels implied only by shared exposure to macro-financial fundamentals. Interestingly, there is a particularly large and persistent deviation of risk from fundamentals preceding the global financial crisis of 2007-2009 in the U.K., euro area, and Asia-Pacific regions. Risk conditions were then significantly and persistently below what had been suggested by fundamentals.

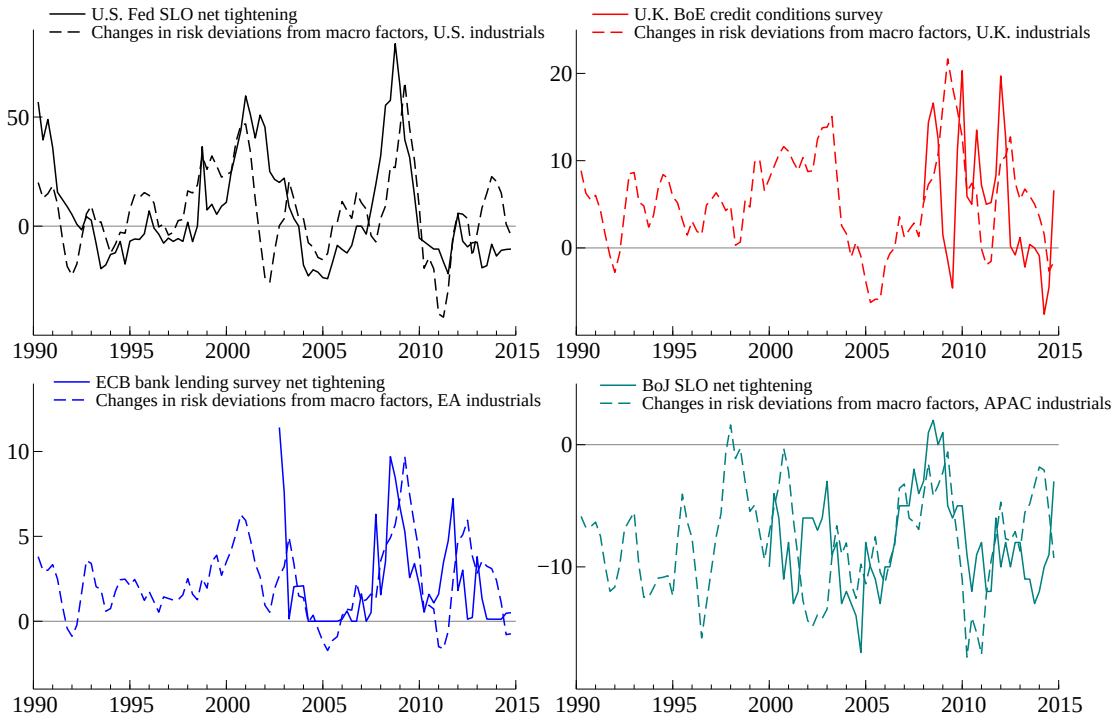
Figure 8 plots risk deviation estimates and the net tightening of bank lending standards for the four economic regions. The bank lending standards refer to commercial and investment loans to medium to large enterprises, and are taken from respective surveys conducted by the Federal Reserve, Bank of England, European Central Bank, and Bank of Japan.[7]

We draw two main conclusions. First, physical credit risks and bank lending (credit) quantities are strongly 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. The (too) low default rate is then not a sign of economic strength, but rather an indication of a financial boom and thus a warning signal for financial fragility and its subsequent potential unraveling. The reverse holds in a credit crunch. In a credit crunch, even financially sound firms 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. Our results are in line with a literature on portfolio credit risk that concludes

[7]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 (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, to our knowledge, the first to empirically tie unobserved credit risk deviations to the variation in (U.S.) bank lending standards.

Figure 8: Credit risk deviations vs. bank lending standards

Changes (year-on-year) in the credit risk deviations from fundamentals for industrial firms (see Figure 7) are associated with the net tightening of lending standards as reported in central bank surveys. Bank lending standards are from surveys by Federal Reserve for the U.S. (top left), Bank of England for the U.K. (top right), European Central Bank for the euro area (bottom left), and Bank of Japan for Japan (bottom right). Shaded areas pertain to NBER recession dates. Credit risk deviations (dashed lines) are matched in terms of means and variances to the net tightening of bank lending standards (solid lines).



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 (2014), and Koopman, Lucas, and Schwaab (2011).[8]

[8]The close association of bank lending standards with frailty factor dynamics observed in Figure 8 would suggest default risk model specifications which include bank lending standards directly as a separate observable factor, possibly even distinguishing its global (average) and local (region-specific deviations from that average) components. We do not do this in our current analysis for the main reason that our estimation sample is from 1980Q1 to 2014Q4, while bank lending standards are only available for a fraction of that time (from 2003Q2 onwards in the euro area, and from 2007Q3 onwards in the U.K., for example). We would expect that the need for unobserved residual factors in default risk models reduces significantly once survey-based information on bank lending practises are included as a conditioning variable. For a commercial

Second, the observed co-movement between bank lending standards and (excess) default risks suggest that two-way feedback effects between the health of the financial sector and the macro-economy are important. This is potentially important for predictive models that forecast default risk parameters conditional on a macro-financial stress scenario. While macro-financial stress is naturally mapped into higher default probabilities (and therefore higher bank losses and lower capital ratios), the subsequent financial tightening of bank lending standards could raise credit risk parameters even further. The tight correlation between physical default risk and financial intermediation implied by Figure 8 suggests that such second round feedback effects are substantial. It might thus be worthwhile to include bank lending standards as an additional conditioning variable in macro-prudential stress tests.

5 Conclusion

We investigated the common dynamic properties of systematic default risk conditions across countries, regions, and the world. For this purpose we developed a high-dimensional, partly non-linear non-Gaussian state space model to estimate common components in firm defaults in a 41-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 of observed default clustering, thus providing evidence for a world credit risk cycle on top of world business cycle conditions. The presence of such global macro and frailty dynamics naturally limit the scope for cross-border credit risk diversification in the financial industry and may negatively affect the risk bearing capacity of globally active lenders as a result. Such limits to risk baring capacity may particularly impede cross-border diversification of (currently) U.S. concentrated lenders, whereas lenders from other world regions typically benefit from diversification to the U.S.

application that proceeds in this way; see Benzschawel and Su (2014).

References

Acharya, V., S. Bharath, and A. Srinivasan (2007). Does industry-wide distress affect defaulted firms? Evidence from creditor recoveries. *Journal of Financial Economics 85(3)*, 787–821.

Acharya, V. V., S. A. Davydenko, and I. A. Strebulaev (2012). Cash holdings and credit risk. *Review of Financial Studies 12*, 3572–3609.

Aoki, K. and K. Nikolov (2015). Bubbles, banks and financial stability. *Journal of Monetary Economics, forthcoming*.

Azizpour, S., K. Giesecke, and G. Schwenkler (2014). Exploring the sources of default clustering. *Unpublished working paper, Stanford University*.

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.

Benzschawel, T. L. and Y. Su (2014). Predicting annual default rates. *Citi Research Mimeo*.

Boissay, F., F. Collard, and F. Smets (2015). Booms and banking crises. *Journal of Political Economy, forthcoming*.

Bräuning, F. and S. J. Koopman (2014). Forecasting macroeconomic variables using collapsed dynamic factor analysis. *International Journal for Forecasting 30*(3), 572–584.

Bruno, V. and H. S. Shin (2012). Capital flows, cross-border banking and global liquidity. *NBER working paper 19038*.

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

Claessens, S., M. A. Kose, and M. E. Terrones (2012). How do business and financial cycles interact. *Journal of International Economics 87(1)*, 178–190.

Clerc, L., A. Derviz, C. Mendicino, S. Moyen, K. Nikolov, L. Stracca, J. Suarez, and A. Vardoulakis (2015). Capital regulation in a macroeconomic model with three layers of default. *Journal of International Central Banking, forthcoming*.

Creal, D., B. Schwaab, S. J. Koopman, and A. Lucas (2014). An observation driven mixed measurement dynamic factor model with application to credit risk. *The Review of Economics and Statistics 96 (5)*, 898–915.

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.

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.

Duan, J.-C., J. Sun, and T. Wang (2012). Multiperiod corporate default prediction - A forward intensity approach. *Journal of Econometrics 170(1)*, 191–209.

Duan, J.-C. and E. van Laere (2012). A public good approach to credit ratings – from concept to reality. *Journal of Banking and Finance 36*(12), 3239–3247.

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

Duffie, D., L. Saita, and K. Wang (2007). Multi-period corporate default prediction with stochastic covariates. *Journal of Financial Economics 83(3)*, 635–665.

Durbin, J. and S. J. Koopman (2000). Time series analysis of non-Gaussian observations based on state space models from both classical and Bayesian perspectives. *Journal of the Royal Statistical Society, Series B 62(1)*, 3–56.

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.

Giesecke, K., F. A. Longstaff, S. Schaefer, and I. Strebulaev (2011). Corporate bond default risk: A 150-year perspective. *Journal of Financial Economics 102*, 233–250.

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

Hoffmann, B., S. Eickmeier, and L. Gambacorta (2014). Understanding global liquidity. *European Economic Review 68*, 1–18.

Jotikasthira, C., A. Le, and C. Lundblad (2011). Why do term structures in different currencies comove? *Unpublished working paper*.

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. J., 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)*, 110–130.

Lando, D. (2003). *Credit risk modelling – Theory and applications*. Princeton, N.J.: 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.

Lang, L. and R. Stulz (1992). Contagion and competitive intra-industry effects of bankrupcy announcements. *Journal of Financial Economics 32*, 45–60.

Lawley, D. N. and A. E. Maxwell (1971). *Factor analysis as a statistical method.* American Elsevier Publishing, New York.

Longstaff, F. and E. Schwartz (1995). A simple approach to valuing risky fixed and floating rate debt. *Journal of Finance 50*, 789–819.

McCullagh, P. and J. A. Nelder (1989). *Generalized linear models, 2nd edition.* Chapman and Hall, London.

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, N.J.: 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.

Moench, E., S. Ng, and S. Potter (2013). Dynamic hierarchical factor models. *Review of Economics and Statistics 95(5)*, 1811–1817.

Neely, C. J. and D. E. Rapach (2011). International comovements in inflation rates and country characteristics. *Journal of International Money and Finance 30(7)*, 1471–1490.

Pesaran, H., T. Schuermann, B. Treutler, and S. Weiner (2006). Macroeconomic dynamics and credit risk: A global perspective. *Journal of Money, Credit, and Banking 38, No. 5*, 1211–1261.

Richter, B. and C. Euffinger (2014). VaR(ious) leverage cycles of banks. *Unpublished working paper, Goethe University of Frankfurt*.

Stock, J. H. and M. W. Watson (1989). New indexes of coincident and leading economic indicators. *NBER Macroeconomics Annual*, 351–393.

Stock, J. H. and M. W. Watson (2002). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association 97(460)*, 1167–1179.

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*.

Vassalou, M. and Y. Xing (2004). Default risk in equity premiums. *Journal of Finance 109(2)*, 831–868.

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)*, 171–199.

Appendix: Simulated Maximum Likelihood Estimation

We consider available data as in (13),

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

where each row $\mathcal{Y}_k = (\mathcal{Y}_{k1}, \ldots, \mathcal{Y}_{kT})$, $k = 1, \ldots, K$, relates to time series data from a potentially different family of parametric distributions according to (14). 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 $\eta_t \sim \mathrm{N}(0, \Sigma_\eta)$, for $t = 1, \ldots, T$, and initial condition $f_1 \sim \mathrm{N}(0, \Sigma_f)$.

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

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

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 $\mathcal{Y}_t = (\mathcal{Y}_{1,t}, \ldots, \mathcal{Y}_{K,t})'$ 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 $k = 1, \ldots, K$. To enable the identification of all entries in ψ, we assume standardized factors in (15) which we enforce by the restriction $\Sigma_f = \mathrm{I}$ in the expression for the unconditional covariance matrix $\Sigma_f = \Phi \Sigma_f \Phi' + \Sigma_\eta$, implying that $\Sigma_\eta = \mathrm{I} - \Phi\Phi'$.

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 $\mathcal{Y} = (\mathcal{Y}_1', \ldots, \mathcal{Y}_T')'$ and $f = (f_1', \ldots, f_T')'$ 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

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

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 $f^{(k)}$.