Exploring the Sources of Default Clustering*

S. Azizpour, K. Giesecke, G. Schwenkler

June 15, 2011; this draft February 24, 2014†

Abstract

We study the sources of corporate default clustering using data on industrial and financial default timing in the U.S. between 1970 and 2012. The analysis is based on a new reduced-form model of correlated default timing, in which the event arrival rate is allowed to depend on past defaults and time-varying risk factors, some of which cannot be measured. The likelihood estimates provide strong evidence of the presence of several distinct sources of default clustering. One source is firms' joint exposure to a common macro-economic factor represented by the U.S. GDP growth rate. Another is the influence on firms of a common latent factor with strong mean-reverting behavior. A third is a contagion mechanism through which the default by one firm has a direct impact on the health of other firms. We find that the impact is governed by the debt outstanding at default and decays with time.

*This paper is a significant revision of an earlier paper by the first two authors entitled "Self-Exciting Corporate Defaults: Contagion vs. Frailty." Azizpour is at Apollo Global Management, Giesecke is at the Department of Management Science & Engineering, Stanford University, Schwenkler is at the Department of Finance, Boston University School of Management. Schwenkler is corresponding author: Phone (617) 358-6266, email: gas@bu.edu, web: http://people.bu.edu/gas/.

†We are grateful for grants from Moody's Corporation and the Global Association of Risk Professionals, and for data from Moody's. Schwenkler was supported by a Mayfield Stanford Graduate Fellowship and a Lieberman Fellowship. We thank Richard Cantor for providing access to data, Darrell Duffie, Andreas Eckner, Jean-David Fermanian, Lisa Goldberg, Francis Longstaff, Michael Ohlrogge, George Papanicolaou, Bjorgvin Sigurdsson, Ken Singleton, Justin Sirignano, Ilya Strebulaev, Stefan Weber, Liuren Wu, and seminar participants at Boston University, Columbia University, the Fifth Bachelier World Congress, Georgia Institute of Technology, the International Monetary Fund, Princeton University, Stanford University, the University of Illinois at Urbana-Champaign, and the University of Southern California for discussions and comments, and Xiaowei Ding, Baeho Kim, and Supakorn Mudchanatongsuk for excellent research assistance.

1 Introduction

The U.S. economy has repeatedly suffered significant clusters of corporate default events. Examples include the savings and loan crisis of the 1990s, the burst of the dotcom bubble in 2001, and the financial crisis of 2007-09. Figure 1 illustrates these events. An understanding of the sources and the degree of default clustering is crucial for the measurement of credit risk at financial institutions, the management of systemic financial risk, and the rating and risk analysis of securities exposed to correlated corporate default risk, such as collateralized debt obligations.

A major source of default clustering is the joint exposure of firms to common or correlated risk factors such as interest rates, stock returns, and GDP growth.[1] The movements of these factors cause correlated changes in firms' conditional default rates. For example, strong economic growth often reduces the likelihood of default across the board. However, Das, Duffie, Kapadia & Saita (2007) provide strong evidence that this channel on its own cannot explain the degree of clustering found in the data. The literature discusses two potential sources of the excess clustering. Several authors suggest information effects. It is possible that some relevant risk factors have not been identified or are simply unobservable.[2] The uncertainty regarding the current values of these latent "frailty" factors, which has an influence on the conditional default rates of the firms that depend on the same frailties, is an additional source of clustering. Others make the case for contagion, by which the default by one firm may have a direct impact on the conditional default rates of other firms.[3] Financial, legal, or business relationships between firms might act as a conduit for the spread of risk. For instance, a default by the protection seller in a credit swap could expose the buyer of protection and increase the default risk borne by the protection buyer's other counterparties.[4] A related example is the collapse of Lehman Brothers in 2008, which pushed some of Lehman's creditors, trading partners and clients into financial distress.[5] Contagion is not limited to the financial sector. For example, the default by major parts supplier Delphi in 2005 exposed General Motors, as indicated by a jump in GM's stock price and credit swap spreads.[6]

[1] Chava & Jarrow (2004), Campbell, Hilscher & Szilagyi (2008), Duffie, Saita & Wang (2007), Giesecke, Longstaff, Schaefer & Strebulaev (2011), Shumway (2001), and others analyze the significance of these and other firm-specific and macro-economic factors for U.S. default timing.

[2] Theoretical and empirical models of frailty are provided by Collin-Dufresne, Goldstein & Helwege (2009), Delloye, Fermanian & Sbai (2006), Duffie, Eckner, Horel & Saita (2009), Giesecke (2004), Koopman, Lucas & Monteiro (2008), Koopman, Lucas & Schwaab (2011), and others.

[3] Aït-Sahalia, Cacho-Diaz & Laeven (2013), Berndt, Ritchken & Sun (2010), Collin-Dufresne, Goldstein & Hugonnier (2004), Errais, Giesecke & Goldberg (2010), Jarrow & Yu (2001), Jorion & Zhang (2007), Lang & Stulz (1992) and others study theoretical and empirical models of contagion.

[4] Jorion & Zhang (2007) and Stulz (2010) study these counterparty effects.

[5] For empirical analyses of the aftermath of Lehman's bankruptcy, see Aragon & Strahan (2012), Brunnermeier (2009), Chakrabarty & Zhang (2012), Dumontaux & Pop (2013), Fernando, May & Megginson (2012), and others.

[6] Boone & Ivanov (2012) provide evidence of default spillover effects on business partners.

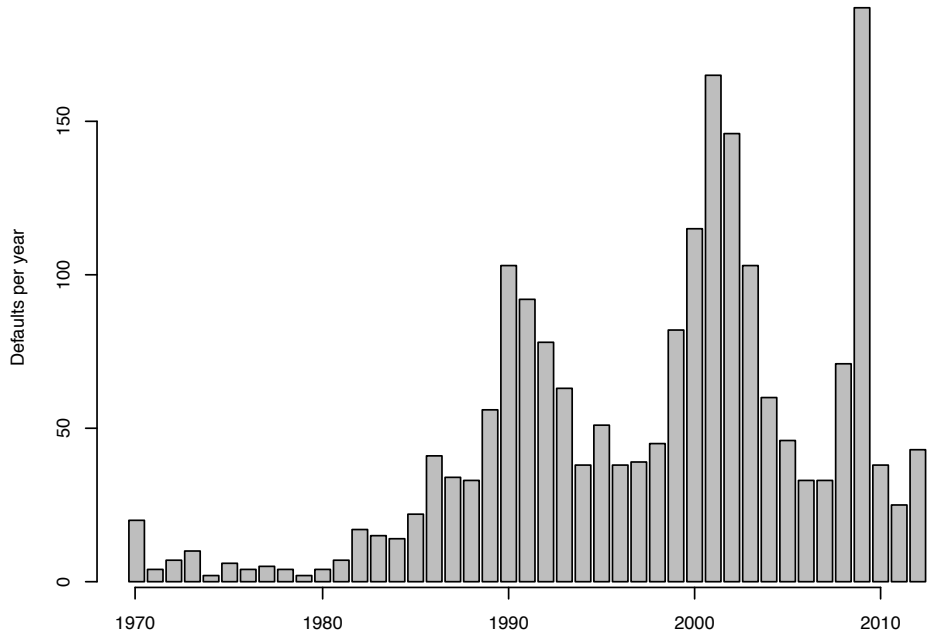


Figure 1: *Annual number of defaults of U.S. firms with Moody's rated debt.* Source: Moody's Default Risk Service.

While both frailty and contagion are plausible potential sources of default clustering, it has not been resolved whether they are indeed responsible for the excess clustering not explained by firms' exposure to common systematic risk factors. As discussed in Section 1.1, several authors uncover evidence of the presence of a latent frailty factor influencing U.S. non-financial default timing. However, they ignore the contagion channel when measuring the impact of frailty. Others analyze non-financial contagion effects, but without controlling for the influence of common observable or latent factors. Failure to account for all potential clustering sources when analyzing the data may lead to biased estimates of the role of individual sources.

Using data on industrial and financial default timing in the U.S. between 1970 and 2012, this paper evaluates the *joint significance* of all potential clustering sources discussed in the literature. We find strong evidence that firms' joint exposure to a common macroeconomic factor represented by the U.S. GDP growth rate is a major source of clustering. An additional significant source of clustering we distinguish in the data is the influence of a common latent frailty factor which exhibits strong mean-reverting behavior. We also find strong evidence of the importance of a contagion channel. Even after controlling for the influence of observable and latent risk factors, a default event is found to have a persistent impact on conditional default rates, with a half-life of about 3 months. A default of a firm with $240 million of debt outstanding at default, the average amount of debt outstanding at default in our data set, ramps up the conditional economy-wide default rate by roughly 3 defaults per year.

Our empirical results are based on a new reduced-form model of correlated default risk, which addresses firms' exposure to observable factors, a latent frailty factor, and failure events. The conditional rate of defaults is allowed to depend on time-varying factors that are observable throughout the sample period, a latent frailty factor with square-root dynamics that is not observable at all, and past failures. Our model significantly extends the standard doubly-stochastic formulation widely used in theoretical and empirical analyses of correlated default risk.[7] In a doubly-stochastic formulation, default arrivals are assumed to be conditionally Poisson given the paths of the observable risk factors, and the only source of clustering is firms' joint exposure to these factors. Our model also extends a richer formulation with observable factors and a latent frailty factor, in which arrivals are assumed to be conditionally Poisson given the paths of all factors. The doubly-stochastic assumption imposes strong restrictions on the conditional distribution of events since it precludes a direct influence of past failures on the conditional default rate. Because we do not make the doubly-stochastic assumption, we avoid such restrictions. By allowing the conditional default rate to depend on past failures, our model of correlated default risk generates a much richer set of distributions.

Without the doubly-stochastic assumption, model estimation and goodness-of-fit testing are no longer standard. We employ a variant of the filtered likelihood estimation developed by Giesecke & Schwenkler (2014) for point processes with incomplete data. To rigorously evaluate the goodness-of-fit of an estimated model, we significantly extend the time-change tests developed by Das et al. (2007) for doubly-stochastic models with observable risk factors. Our tests facilitate the statistical assessment of models that are not doubly-stochastic because of the influence of past defaults and latent frailty factors. Unlike the tests developed by Das et al. (2007), our tests allow for the comparison of models addressing different sources of default clustering.

We estimate our model and several nested alternatives, each of which addresses a different set of clustering sources. We then evaluate each of these models using in- and out-of-sample tests to understand the empirical significance of the potential clustering sources. Martingale specification tests indicate the significance of the frailty and contagion channels. We reject the null hypothesis of a correctly specified base model that addresses only the clustering implied by firms' joint exposure to the observable macro-economic risk factors that prior studies have identified as predictors of U.S. defaults. Likelihood ratio tests suggest that the contagion channel takes a somewhat more prominent role than the frailty channel for explaining the default clusters in the data. The time-change tests indicate that a model including the contagion channel outperforms, in terms of in-sample fit, a model ignoring that channel. We find evidence that a model ignoring the contagion channel tends to overstate the significance of the frailty channel.

Out-of-sample tests of forecast accuracy paint a similar picture. For each model alternative, we compute the value at risk (VaR) of the year-ahead default forecast. We then

[7] See Chava & Jarrow (2004), Duffie & Garleanu (2001), Duffie et al. (2007), Feldhütter & Nielsen (2012), Mortensen (2006), and others.

perform a series of tests and compute forecast accuracy measures for these VaRs. We reject the null hypothesis that the base model accurately predicts correlated default risk out-of-sample. The base model, which addresses only the clustering implied by firms' exposure to the observable macro-economic risk factors identified by prior work, produces sluggish forecasts that severely lag actual defaults. The models addressing additional clustering sources perform much better at forecasting defaults. There are, however, important differences between the models. A model ignoring the contagion channel generates excessively high and volatile VaR forecasts. A model ignoring the frailty channel tends to understate the VaR. Only a model addressing all three sources of clustering provides accurate and sensible out-of sample forecasts of correlated default risk in the U.S.

Our empirical findings have significant implications for the management of credit risk at financial institutions. Most importantly, our results indicate that the doubly-stochastic models of correlated default risk widely used to estimate risk capital in practice may understate the risk of large losses from defaults, especially during clustering periods such as 2001-02 and 2007-09. As a result, financial institutions may end up with capital buffers that are inadequate to withstand the large losses typically associated with default clustering episodes. Our results suggest it is necessary to address the frailty and contagion channels of clustering when measuring correlated default risk and estimating risk capital. Accounting for the effects of frailty and contagion will lead to more accurate risk assessments and more adequate capital buffers.

The rest of this paper is organized as follows. The remainder of the introduction discusses the related literature. Section 2 details the empirical data. Section 3 presents a preliminary regression analysis of the data. The results of this analysis motivate the features of our reduced-form model of default timing, which is developed in Section 4. Section 5 outlines the estimation approach and discusses the parameter estimates for several model alternatives. Section 6 tests the in-sample fit of these alternatives, while Section 7 evaluates their out-of-sample forecast performance. Section 8 concludes. There are several technical appendices.

1.1 Related literature

Using data on industrial defaults in the U.S. between 1980 and 2004, Duffie et al. (2007) identify a set of observable firm-specific and macro-economic risk factors influencing the timing of defaults in a doubly-stochastic model. Das et al. (2007) develop a time-change test to evaluate the doubly-stochastic assumption that firms' default times are correlated only as implied by the dependence of firms' conditional default rates on these factors. They reject the joint hypothesis of well-specified default rates and the doubly-stochastic assumption. Using a different set of risk factors, however, Lando & Nielsen (2010) cannot reject this hypothesis. We provide additional evidence on this issue, focusing on the

economy-wide default rate rather than firm-level rates.[8] Based on a significantly longer sample period that also covers the recent financial crisis and a larger event sample that also includes failures of financial firms, we firmly reject the hypothesis that the economy-wide default rate is influenced only by the macro-economic factors that the aforementioned papers have identified.

Duffie et al. (2009) use data on U.S. industrial defaults between 1979 and 2004 to estimate a model in which a firm-level default rate is influenced by a set of observable firm-specific and macro-economic factors and a mean-reverting frailty. Their model is doubly-stochastic with respect to observable and latent frailty risk factors. They find that the frailty has a large impact on fitted default rates. In contrast to this study, we consider the economy-wide rate of industrial and financial defaults and control for the contagion channel when measuring the impact of frailty. We find that the role of frailty is overstated when the contagion channel is ignored. Moreover, we find that the contagion channel is highly significant even in the presence of frailty. Our findings suggest that both frailty and contagion are significant clustering sources.

Lando & Nielsen (2010) use data on U.S. industrial defaults between 1982 and 2005 to fit a model in which a firm-level default rate is influenced by a set of observable firm-specific and macro-economic factors as well as past defaults. They find that the influence of past defaults is insignificant. In contrast to this study, we consider the economy-wide rate of industrial and financial defaults and control for the presence of frailty when measuring the impact of past defaults. We find strong evidence that the conditional default rate depends on past failures, regardless of whether or not frailty is present.

Self-exciting event timing models that incorporate the dependence of the conditional event rate on past events have been used by Aït-Sahalia et al. (2013) for analyzing the dynamics of asset returns with feedback jumps, by Bowsher (2007) and others for studying the dynamics of order book data, and by Jarrow & Yu (2001) and others for pricing corporate bonds and credit derivatives. However, in contrast to these prior studies, we also model and estimate the dependence of the conditional event rate on a latent frailty factor with square-root dynamics.

2 The data

The sample period is 1/1/1970 to 12/31/2012. Data on corporate (industrial and financial) default timing in the U.S. are obtained from Moody's Default Risk Service. The data cover all defaults of firms domiciled in the U.S. with Moody's rated debt. We also obtain data on the debt outstanding at default from the same source.

We adopt Moody's definition of a default event (see Hamilton (2005)). A "default" is an event in any of the following categories: (1) a missed or delayed disbursement of interest

[8] Longstaff & Rajan (2008), Giesecke et al. (2011), and Giesecke, Longstaff, Schaefer & Strebulaev (2014) also pursue an economy-wide rather than a firm-level approach to default timing.

or principal, including delayed payments made within a grace period; (2) bankruptcy (Section 77, Chapter 10, Chapter 11, Chapter 7, Prepackaged Chapter 11), administration, legal receivership, or other legal blocks to the timely payment of interest or principal; or (3) a distressed exchange occurs where: (i) the issuer offers debt holders a new security or package of securities that amount to a diminished financial obligation; or (ii) the exchange had the apparent purpose of helping the borrower avoid default. A repeated default by the same issuer is included in the set of events if it was not within a year of the initial event and the issuer's rating was raised above Caa after the initial default. This treatment of repeated defaults is consistent with that of Moody's.

We observe a total of 2001 default events during the sample period. Figure 2 shows the number of defaults by event category. The vast majority of defaults are due to missed interest payments, Chapter 11 filings, or exchanges of distressed debt that resulted in losses to bond investors. Figure 2 also indicates the number of events by sector. Most defaults occur in the capital industry, followed by the consumer industry, technology, retail and distribution, and media and publishing sectors. Figure 3 indicates the distribution of the total amount of debt outstanding at default. About 95% of the events involve debt outstanding at default of $500 million or less.

We study the influence on default timing of a number of explanatory variables (risk factors). Time series of these variables are constructed from data obtained from the Federal Reserve Banks of New York and Saint Louis; Table 1 reports summary statistics. The variables we consider include the quarterly growth rate of the U.S. industrial production calculated from final products and nonindustrial supplies (observed monthly), the quarterly growth rate of the U.S. gross domestic product (observed quarterly), the trailing 1-year return of the S&P 500 index (observed daily), the trailing 1-year volatility of the S&P 500 index (observed daily), the yield of the 3-month Treasury bill (observed daily), the spread between the 10-year Treasury note and the 1-year Treasury bill (observed daily), and the spread between the yields of BAA and AAA Moody's rated corporate bonds (observed weekly). Under different econometric assumptions and using different default data sets, Chava & Jarrow (2004), Das et al. (2007), Duffie et al. (2007), Campbell et al. (2008), Duffie et al. (2009), Lando & Nielsen (2010), Giesecke et al. (2011), and others have found some or all of these variables to be significant predictors of U.S. defaults. Unlike some of the aforementioned studies, we do not consider any firm-specific variables because we are primarily interested in understanding the mechanisms generating clustered default events rather than individual defaults. Only systematic factors are relevant in this context.

3 Regression analysis

We begin with an exploratory regression analysis of the number of defaults per month. Our objective is to highlight several salient features of the default data. The findings will

guide the formulation of a reduced-form model of event timing in Section 4.

Table 1 presents summary statistics for the time series of monthly defaults, total amount of debt outstanding at default per month, and risk factors (sampled monthly). On average, there are 3.88 defaults per month; the median is 3 and the standard deviation is 4.42. The month with the most defaults is March 2009; it witnessed 29 events. There is moderate negative correlation between monthly defaults and industrial production growth, GDP growth, trailing S&P 500 returns, and Treasury yields. There is some positive correlation between monthly defaults and S&P 500 volatility, slope of the yield curve, and corporate yield spreads.

We run Poisson regressions of the monthly defaults on the explanatory variables.[9] Table 2 reports the regression coefficients and the corresponding t-statistics. All explanatory variables are significant at the 99% level. The R^2 of the regression suggests that these variables can explain only about 40% of the variation of monthly defaults. This and our other findings discussed below are robust with respect to the choice of the aggregation time window (e.g., weekly or quarterly defaults).

Figure 4 shows that the regression residuals are excessive during clustering episodes. There is strong positive correlation of 75% between the residuals and monthly defaults. Figure 5 indicates that the residuals are significantly positively autocorrelated. Also, the constant coefficient of the regression is large and highly significant. These observations suggest the existence of *latent* explanatory variables (frailties) influencing default timing beyond the variables we have already included as regressors.

A natural candidate for an additional explanatory variable is lagged defaults. Figure 6 indicates that monthly defaults are strongly autocorrelated; the lag one autocorrelation is roughly 80%. Furthermore, the inclusion of lagged defaults in the regression allows us to address a contagion channel, by which the default by a firm may adversely affect that firm's creditors, counterparties, or business partners. Several studies provide evidence supporting the existence of contagion effects. For example, Jorion & Zhang (2007) document contagion effects using CDS data, Collin-Dufresne et al. (2009) study such effects using bond return data, and Lang & Stulz (1992) examine intra-industry contagion effects.

The regression coefficients and t-statistics, reported in Table 2, suggest that lag one monthly defaults and lag one monthly debt outstanding at default are highly significant predictors of future defaults, after controlling for the influence of the explanatory variables identified earlier. Including the lagged default data improves the R^2 to 0.56 from 0.40, an improvement that is significant at the 99.9% level according to a likelihood ratio test. Moreover, including the lagged default data reduces the correlation between regression residuals and monthly defaults from 74% to 62% (see Figure 4). It also reduces the autocorrelation of the regression residuals; the autocorrelation is now insignificant for the first order (see Figure 5). Despite these substantial improvements, the features of the residuals still suggest the presence of latent frailty variables, even if higher order lagged default

[9]We use Poisson regressions rather than standard linear regressions because the default data are integer-valued. See Cameron & Trivedi (1998) and Winkelmann (2008) for details on Poisson regressions.

data are included in the regression.

4 Reduced-form model

Guided by the results of the regression analysis, we formulate a reduced-form model of correlated default timing. Reduced-form models are widely used; see Berndt et al. (2010), Chava & Jarrow (2004), Das et al. (2007), Duffie et al. (2007), Duffie et al. (2009), Jarrow & Yu (2001), Lando & Nielsen (2010), Longstaff & Rajan (2008), and many others. Unlike a regression model, a reduced-form model addresses the exact timing of events. It allows us to capture the potentially significant effects of defaults within a month and the substantial intra-month time-series variation of some of the explanatory variables.

The defaults in the sample arrive at an intensity that measures the conditional mean arrival rate of events per year (see Appendix A for details). To address the potential sources of default clustering, we allow the intensity to depend on the explanatory variables $X_1, \ldots, X_d$ discussed in Section 2, a latent frailty variable Z, and past defaults. More precisely, if Z were observable, the intensity at time t would be of the form

$$\lambda_t = \exp\left(a_0 + \sum_{i=1}^{d} a_i X_{i,t}\right) + Y_t + Z_t, \tag{1}$$

where Y models the influence of past defaults on the intensity, and $a = (a_0, a_1, \ldots, a_d)$ is a vector of coefficients to be estimated. The intensity at t given the actually available information is the posterior mean of λ_t; details are in Appendix C.

The first term in (1) captures firms' exposure to the common risk factor $X = (X_1, \ldots, X_d)$. As in Chava & Jarrow (2004), Das et al. (2007), Duffie et al. (2007) and other articles, this term takes the classical proportional hazards form of Cox (1972). When Y and Z are positive, this formulation ensures that the intensity is positive for any value of the factor vector X. Moreover, when $Y = 0$ and $Z = 0$, the influence of the risk factor X_i is multiplicative and the parameter a_i is a proportionality factor. If X_i increases by one unit, then the intensity increases by a factor of $e^{a_i} - 1$.

The results of the regression analysis in Section 3 suggest the existence of a latent frailty risk factor influencing the timing of defaults in our sample. The term Z in (1) models this influence. We propose a standard mean-reverting Cox-Ingersoll-Ross (CIR) model for the dynamics of the latent factor Z:

$$\mathrm{d}Z_t = k(z - Z_t)\mathrm{d}t + \sigma\sqrt{Z_t}\mathrm{d}W_t, \tag{2}$$

for $z, k, \sigma \geq 0$ and $2kz \geq \sigma^2$.[10] Here, W is a standard Brownian motion independent of X, and Z is initialized at its stationary gamma distribution with shape parameter $\frac{2kz}{\sigma^2}$ and scale parameter $\frac{\sigma^2}{2k}$. The choice of a mean-reverting model is motivated by the results of

[10] We have considered alternative parameterizations of the model (2), including one with unit volatility but Z in (1) replaced by cZ. The results for these alternatives are very similar to those for (2).

Duffie et al. (2009), who find strong evidence for the existence of a mean-reverting frailty influencing U.S. non-financial default timing between 1979 and 2004.

The results of the regression analysis also highlight the dependence of default rates on past failures, consistent with a number of studies analyzing contagion effects.[11] The term Y in (1) addresses this dependence in our reduced-form formulation. We propose the classical self-exciting specification of Hawkes (1971):

$$Y_t = b \sum_{n:\, T_n \le t} e^{-\kappa(t-T_n)} f(U_n) \tag{3}$$

where $b > 0$ and $\kappa > 0$ are parameters, T_n is the time of the nth default in the sample, U_n is the debt outstanding at the nth default (measured in million USD), and the function $f(u) = \max(0, \log u)$ specifies the impact of an event on the intensity.[12] Upon a default, the intensity is ramped up proportionally to the logarithm of the total debt outstanding at default. The impact fades away with time at rate κ. This formulation captures the positive autocorrelation of monthly defaults found in Section 3, which decreases with increasing lag. It also captures the significant positive dependence between monthly defaults and lag one debt outstanding at default per month found in the regression analysis.

5 Likelihood estimation

We estimate the parameter $\theta = (a, b, \kappa, \sigma, k, z)$ governing the reduced-form model (1) by the method of maximum likelihood. Since the frailty Z influencing the intensity cannot be observed, the likelihood problem is not standard unless $z = \sigma = 0$. We implement a variant of the filtered likelihood estimation developed by Giesecke & Schwenkler (2014); see Appendix B for details. The likelihood is the posterior mean of the complete-data likelihood given the observed data, which include the default dates, the debts outstanding at defaults, and time series of the explanatory variables $X_1, \ldots, X_d$ discussed in Section 2.[13] The posterior mean is evaluated using a quadrature method. This approach eliminates the need to perform Monte Carlo simulations of the frailty path over the 43-year sample period, which would be computationally burdensome. It also avoids the loss of estimation efficiency associated with the simulated expectation-maximization (EM) algorithm often used for this type of problem.[14] The resulting likelihood estimators are consistent and asymptotically normal as the sample period grows (see Appendix B.3).

[11] See, for example, Berndt et al. (2010), Chakrabarty & Zhang (2012), Collin-Dufresne et al. (2004), Errais et al. (2010), Jarrow & Yu (2001), Jorion & Zhang (2007), and Lang & Stulz (1992).

[12] We have experimented with several alternative models of f but found the specification presented here to be the most significant one.

[13] Following the empirical default timing literature, we interpolate the discrete observations of the explanatory variables to obtain continuous time series.

[14] Duffie et al. (2009), Koopman et al. (2008), Koopman et al. (2011), and others use EM algorithms to estimate the influence of frailty on default timing.

We estimate the model (1) and three nested alternatives, each of which addresses a different set of clustering sources. Table 3 summarizes these models. The unrestricted model, for which $b, z, \sigma > 0$, addresses the clustering due to firms' exposure to observable and frailty factors as well as to the clustering generated by the contagion channel. A model for which $b = 0$ ignores the contagion channel. A model for which $z = \sigma = 0$ ignores the frailty channel. A base model for which $b = z = \sigma = 0$ ignores both the contagion channel and the frailty channel; it captures only the clustering implied by firms' joint dependence on the observable risk factor X. Table 4 reports the parameter estimates for the different models along with the corresponding asymptotic standard errors and maximum log-likelihoods. We summarize our findings below.

The self-exciting term Y is highly significant. As suggested by the regression analysis in Section 3, past defaults appear to be significant predictors of future defaults, after controlling for the impact of observable and frailty factors. A default has a persistent impact on the intensity, with a half-life of about 3 months. A default of a firm with $240 million of debt outstanding at default, the average amount of debt outstanding at default in our data set, ramps up the intensity by roughly 3 defaults per year. The fitted values of the parameters b and κ governing the term Y are very similar to the fitted values in the model ignoring the frailty channel. This observation suggests that the term Y has been appropriately identified. In particular, the role of the term Y is not overstated at the expense of the frailty term.

The latent frailty factor Z exhibits significant mean-reverting behavior. The fitted mean-reversion parameters k and z differ by orders of magnitude from those in the model ignoring contagion. In the unrestricted model, the fitted frailty quickly reverts to a very low but significant level. In the restricted model, the fitted frailty is very persistent and has a high reversion level. This indicates that a model ignoring the contagion channel may overstate the role of frailty.

Among the explanatory variables, only GDP growth has a significant influence on the intensity. The sensitivity of the intensity to this factor is negative, consistent with the intuition that strong economic growth reduces the likelihood of defaults. Moreover, the fitted value of the sensitivity parameter is similar to the values estimated for the restricted models (if these values are significant), suggesting that the impact of GDP growth on the intensity is not overstated at the expense of other risk factors or the frailty and self-exciting terms.

Unlike Duffie et al. (2007), Duffie et al. (2009), and Lando & Nielsen (2010), we do not find the S&P 500 return and the industrial production growth to be significant risk factors. This may be due to the fact that our data set covers a much longer sample period, beginning in 1970 and extending well beyond the financial crisis. Unlike the earlier studies, our event sample also includes the failures of financial firms, which were relatively numerous during the savings and loan crisis. Finally, we control for the frailty and contagion channels of clustering when measuring the impact of the explanatory variables. The aforementioned papers control for only one of these channels, or neither.

Also the corporate bond yield spread is found to be insignificant. This finding complements previous results by Collin-Dufresne, Goldstein & Martin (2001), Giesecke et al. (2011), and Gilchrist & Zakrajšek (2012) on the connection between credit spreads and the business cycle. It provides additional evidence that credit spreads contain little information about actual default probabilities.

6 In-sample analysis

We perform a series of tests to evaluate the in-sample fit of the nested model alternatives we have estimated. Our goal is to better understand the roles that the frailty and contagion channels play for the default clusters in the data.

6.1 Fitted intensities

Figure 7 compares the fitted intensity to the default data. The fitted intensity at time t is given by the fitted posterior mean of λ_t given the data observed to time t. Without frailty ($z = \sigma = 0$), the posterior mean, h_t, is equal to λ_t. With frailty, the computation of h_t is a filtering problem; see Appendix C for details.

The base model does a relatively poor job at explaining the default clusters in the data. It lags clusters and significantly overstates the arrival rates during the first quarter of the sample period and during the recent financial crisis. The richer models capture the substantial time-variation of realized arrival rates much better. The models that address the contagion channel match the low arrival rates during the early part of the sample better than the models that ignore that channel.

6.2 Martingale specification tests

Next, we formally test model specification. Let N_t be the number of defaults observed to time t. The variables $\{M_t : t \geq 0\}$ given by

$$M_t = N_t - \int_0^t h_s \mathrm{d}s, \tag{4}$$

where h_s is the posterior mean of λ_s at time s, form a martingale with respect to the observable data. We consider the sequence of increments of the fitted $\{M_t : t \geq 0\}$ between consecutive defaults. Under the null hypothesis of a correctly specified model, this sequence is a martingale difference sequence.

We run the martingale difference tests described in Charles, Darné & Kim (2011). The results of these tests are reported in Table 5. We reject the null hypothesis of a correctly specified base model. We cannot reject the null for any of the three richer models, indicating the importance of going beyond the base model and addressing the frailty and contagion channels of default clustering.

6.3 Likelihood ratio tests

Using likelihood ratio tests, we measure the improvement in fit achieved by including the frailty and the self-exciting terms. The results of these tests, reported in Table 6, provide strong evidence in favor of including both terms. Relative to the base model, adding either the frailty term or the self-exciting term leads to a significant improvement in model fit. However, the inclusion of the self-exciting term leads to a bigger improvement in the test statistic than the inclusion of the frailty term. Relative to the model that includes observable risk factors and frailty, the inclusion of self-exciting term as a third source of clustering yields further improvements. The inclusion of the frailty term in the model with observable risk factors and the contagion term does not significantly improve model fit, however. These tests suggest that even after controlling for frailty, the self-exciting term representing the contagion channel plays a prominent role.

6.4 Time-change tests

To further analyze the empirical significance of the frailty and contagion channels, we formally test the goodness-of-fit of our model alternatives. An important result of Das et al. (2007) states that if there are no latent frailty factors and defaults are doubly-stochastic,[15] then the default count N_t can be transformed into a standard Poisson process by a change of time given by the cumulative intensity. This result allows one to evaluate the goodness-of-fit of doubly-stochastic models with observable risk factors by testing the Poisson property of the time-scaled default times.

Our model violates the doubly-stochastic hypothesis because it addresses the frailty and contagion channels. Therefore, the tests of Das et al. (2007) do not apply. We generalize the time-change result of Das et al. (2007) in Proposition D.1 in Appendix D. Our result also covers models in which the doubly-stochastic assumption is violated because of the presence of a latent frailty factor or a self-exciting term. We show that the default count N_t can always be transformed into a standard Poisson process by a change of time given by the cumulative *posterior mean* h_t of the intensity. This result allows us to compare models addressing different sources of default clustering by testing the Poisson property of the time-scaled default times.

We follow Das et al. (2007) and begin by testing the Poisson property of defaults per bin. For a bin of size $w > 0$, we construct an increasing sequence of times t_i^w so that the cumulative posterior intensity mean h_t between two consecutive times is equal to w. Formally, $\int_{t_{i-1}^w}^{t_i^w} h_s \mathrm{d}s = w$ for each i. Let P_i^w be the number of defaults between t_{i-1}^w and t_i^w. Proposition D.1 implies that the P_i^w are independent samples from a Poisson distribution with parameter w. We test this property for different bin sizes w.

[15] A model with observable risk factors is doubly-stochastic if default arrivals are conditionally Poisson given the paths of the factors. A doubly-stochastic model addresses only the default correlation caused by firms' joint exposure to the risk factors.

Table 7 reports the first four empirical moments of the binned event counts P_i^w for each of our model alternatives and a range of bin sizes. It also reports the moments of the corresponding theoretical Poisson distributions. We observe that the base model tends to generate binned event counts that are overdispersed, with the variances of the P_i^w exceeding the theoretical counterparts. Overdispersed counts are a sign of missing explanatory factors (see Cameron & Trivedi (1998)). This is consistent with the rejection of the base model by the martingale specification tests (Section 6.2). The model including frailty tends to generate underdispersed binned event counts. Underdispersion is a sign of positive correlation between inter-default times (see Cameron & Trivedi (1998)). This observation complements the analysis of the maximum likelihood estimates in Section 5: it indicates that a model ignoring the contagion channel may in fact overstate the role of frailty. The models including the frailty channel and the contagion channel generate binned event counts whose moments closely match the theoretical ones.

Table 8 reports the p-values of several tests of the i.i.d. Poisson property of the binned event counts (see Karlis & Xekalaki (2000) and Das et al. (2007)). The Potthoff-Whittinghill-Bohning test evaluates the moments and the Kocherlakota-Kocherlakota test evaluates the generating function of the P_i^w. The other tests examine the independence, dispersion, tail properties, and distribution of the P_i^w.

The base model fails almost all of these tests at high significance levels. It generates correlated binned event counts and does not match the corresponding Poisson distributions. We firmly reject the null hypothesis that the intensity is influenced only by the common factors that prior studies have identified as significant predictors of U.S. defaults.

The model including the frailty channel but ignoring the contagion channel fails the Potthoff-Whittinghill-Bohning tests of the moments of the Poisson distribution. It also fails a chi-squared goodness-of-fit test. The unrestricted model addressing all clustering channels fails only few tests, often at low significance levels. All things considered, the data suggest that the contagion channel is a prominent source of default clustering, even after controlling for the frailty channel.

7 Out-of-sample analysis

To develop additional insight into the distinctive properties of alternative clustering sources, we finally evaluate the out-of-sample forecast accuracy of our models. We use Monte Carlo simulation to generate the conditional distribution of the number of defaults in the year ahead, given a model fitted to the data available at the beginning of the year. Our approach, detailed in Appendix E, eliminates the need to simulate the paths of the frailty variable during the forecast period. It is computationally efficient and generates forecast distributions with small variance.

Figure 8 contrasts the forecast distribution with the realized number of defaults, for each year between 1991 and 2012, and for each alternative model. We observe that the

base model lags default clusters, which complements our findings from the in-sample analysis. It fails to predict the clusters associated with the burst of the dotcom bubble and the financial crisis. The model including frailty but ignoring the contagion channel performs better at predicting these clustering events. However, it generates forecast distributions with excessively heavy tails. This may be a result of the underdispersion suggested by the moments of the binned default count in Table 7. Moreover, the mean of the forecast distribution shows very little variation through the test period. Thus, if the mean were used as a point forecast, then the forecast would change little every year, despite the significant variation of observed default rates. The model that includes the contagion channel but ignores the frailty channel generates much more concentrated forecast distributions. The forecast mean matches the realized number of defaults quite well. The unrestricted model, which addresses all clustering sources, performs better still, producing forecast distributions with lower variance during the first third of the test period.

To test forecast accuracy, we consider the 99% value at risk (VaR) of the forecast distribution, a standard measure of portfolio credit risk. For a given model, we evaluate the sequence of "hit" indicators associated with violations of the VaR in different periods. Under the null, these indicators are independent draws from a Bernoulli distribution with success probability $(1 - 0.99) = 0.01$. Table 9 reports the violation rates and the p-values of an unconditional coverage test due to Kupiec (1995), which tests whether the actual violation rate is significantly different from 0.01. We also report the p-values of an independence test due to Christoffersen (1998), a combined test of independence and Bernoulli distribution against a Markov chain alternative due to Christoffersen (1998), and the out-of-sample dynamic quantile test of Engle & Manganelli (2004). The dynamic quantile test examines if there is significant correlation between the hit indicators and the level of the VaR. With a violation rate of 0/22, the model including the frailty channel but ignoring the contagion channel outperforms all other models according to these tests. With a violation rate of 1/22, the unrestricted model comes in second. The base model is rejected by these tests due its sluggishness. The model including the contagion channel but ignoring the frailty channel is rejected because its VaR forecasts are too low.

The tests of the hit indicators may unduly favor the frailty model because it generates forecast distributions with excessively heavy tails, and VaR forecasts that are excessively large. If one were to estimate risk capital according to the VaR, one would end up holding too much capital relative to the actual risk. While one would be covered in most situations, scarce risk capital would be wasted. To address this issue, we consider the mean relative bias and the root mean square relative bias of the VaR forecasts (see Hendricks (1996)). The former is the percentage deviation of a forecast from the average forecast of all models, averaged over all forecast periods. The latter accounts for the standard deviation of the former. The outcomes of these measures, reported in Table 9, indicate that the frailty model does indeed tend to overstate the VaR and to produce the largest forecast volatility. The unrestricted model, which addresses all clustering sources, generates the VaR forecasts with the least bias and volatility.

All things considered, the results of the out-of-sample tests indicate that the frailty and self-exciting terms are equally important for obtaining accurate forecasts of correlated default risk. A default timing model in which the intensity is influenced only by common risk factors does not accurately predict correlated default risk out-of-sample. A model including frailty but ignoring the contagion channel tends to overstate correlated default risk, while a model addressing the contagion channel but ignoring frailty tends to understate it. The unrestricted model provides the most balanced and accurate out-of-sample forecasts of correlated default risk.

8 Concluding remarks

The U.S. financial markets have witnessed several significant clusters of corporate defaults over the past few decades. Using data on industrial and financial default timing in the U.S. between 1970 and 2012, we find strong evidence of the presence of several sources of default clustering. One source is firms' joint exposure to a common macro-economic risk factor that is represented by the U.S. GDP growth rate. Another source is the influence on firms of a common unobservable frailty risk factor with mean-reverting behavior. A third source is a contagion mechanism through which the default by one firm has a direct impact on the health of other firms. The impact is governed by the debt outstanding at default and decays with time.

The empirical analysis employs a new reduced-form model of correlated default timing. We allow the conditional rate of defaults in the sample to depend on dynamic factors that are observable throughout the sample period, a latent factor with square-root dynamics, and past failures. Goodness-of-fit tests are used to evaluate this model and its nested alternatives, and to establish the empirical significance of the different potential sources of default clustering.

Our findings have important implications for the management of credit risk at financial institutions. The models widely used to estimate risk capital address only the default clustering due to firms' joint exposure to observable risk factors. We find strong evidence however that a model ignoring the frailty and contagion channels understates the risk of large default losses. This indicates that banks using conventional models to estimate risk capital might end up holding too little capital to withstand the large losses induced by default clusters like the one generated by the last financial crisis. Our results suggest to estimate risk capital using an approach that also addresses the presence of a mean-reverting frailty factor influencing firms, and the impact of past failures on default rates. We find that such an approach generates the most accurate out-of-sample forecasts of correlated default risk.

An understanding of the sources of default clustering is also essential for the rating and risk analysis of securities that are exposed to correlated default risk, such as collateralized debt obligations. Our findings suggest to account for the impact of frailty and

contagion when studying such securities.

Our model can be applied to other situations in which a common unobservable factor and a contagion mechanism are suspected to play an important role in the timing of events; for example, mortgage prepayments and defaults, order arrivals in a limit order book, and jumps in asset prices.

A Reduced-form model

This appendix provides some technical details regarding our reduced-form model of default timing (Section 4). The default data is a realization of a marked point process $(T_n, U_n)_{n\geq 1}$ defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with an information filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions of right-continuity and completeness (see Protter (2004)). That is, $(T_n)_{n\geq 1}$ is an increasing sequence of default stopping times tending to ∞ and $(U_n)_{n\geq 1}$ is a sequence of $\mathcal{F}_{T_n}$-measurable random variables that represent the debt outstanding at default. We assume that there exists a positive process λ such that the variables $N_t - \int_0^t \lambda_s ds$ form a local martingale, where $N_t = \sum_{n\geq 1} 1_{\{T_n \leq t\}}$. The process λ is the *intensity* of N. We have that $\mathbb{E}[N_{t+\Delta} - N_t \,|\, \mathcal{F}_t] \approx \lambda_t \Delta t$ for small $\Delta > 0$, suggesting the interpretation of a conditional mean arrival rate for λ.

B Maximum likelihood estimation

Due to the presence of the frailty factor Z, the problem of estimating the parameter $\theta = (a, b, \kappa, \sigma, k, z)$ of the reduced-form model (1) is not standard unless $z = \sigma = 0$. We implement a variant of the filtered likelihood estimation method developed by Giesecke & Schwenkler (2014) for point process models with incomplete data. This appendix provides some details.

B.1 The likelihood function

Let D_t denote the data available at time t. These include the default data $\{T_n, U_n : n \leq N_t\}$ and the (interpolated) covariate data $\{X_s : s \leq t\}$. Let $[0, \tau]$ be the sample period. The likelihood function $\mathcal{L}_\tau(\theta)$ is the Radon-Nikodym density of the law of the data D_τ relative to its true distribution; see Section 3.1 of Giesecke & Schwenkler (2014) for a more precise statement. A maximum likelihood estimator (MLE) is a maximizer of the likelihood function. To compute the likelihood, let

$$M_\tau = \exp\left(\sum_{n\geq 1} \log(\lambda_{T_n -}) 1_{\{T_n \leq \tau\}} - \int_0^\tau (\lambda_s - 1) \mathrm{d}s\right), \tag{5}$$

and assume that $\mathbb{E}[1/M_\tau] = 1$. We define an equivalent measure $\mathbb{P}^*$ on the σ-field $\mathcal{F}_\tau$ with Radon-Nikodym density

$$\frac{\mathrm{d}\mathbb{P}^*}{\mathrm{d}\mathbb{P}} = \frac{1}{M_\tau}.$$

Proposition 3.1 of Giesecke & Schwenkler (2014) states that

$$\mathcal{L}_\tau(\theta) \propto \mathbb{E}^*\left[M_\tau \mid D_\tau\right]. \tag{6}$$

The likelihood function (6) depends on the parameter θ because M_τ is a path functional of the intensity $\lambda = e^{a\cdot(1,X)} + Y + Z$. If $z = \sigma = 0$, the data set D_τ contains all relevant information to evaluate M_τ so the likelihood function simplifies to $\mathcal{L}_\tau(\theta) \propto M_\tau$. If $z, \sigma > 0$, then the conditional expectation (6) is not trivial because the data D_τ do not include any observations of the frailty Z. The conditional expectation (6) is taken with respect to the conditional $\mathbb{P}^*$-law of the frailty given the data D_τ. Girsanov's and Lévy's theorems imply that the Brownian motion W driving the frailty Z remains a Brownian motion under $\mathbb{P}^*$. As a result, Z follows the CIR process (2) under $\mathbb{P}^*$. Furthermore, Girsanov's and Watanabe's theorems imply that N is a standard Poisson process under $\mathbb{P}^*$. Therefore, N is independent of the frailty Z under $\mathbb{P}^*$. Also, Z is independent of X by assumption. Thus, the conditional $\mathbb{P}^*$-law of Z given D_τ is governed by (k, z, σ).

B.2 Approximate likelihood function

There is no closed-form expression for the likelihood (6) if $\sigma > 0$ and $z > 0$. Giesecke & Schwenkler (2014) propose to approximate the likelihood using a rectangular quadrature method. We will use a slightly modified approximation based on a trapezoidal quadrature method that exploits the fact that our frailty Z follows a CIR process. This approximation does not interpolate the path of the frailty. Therefore, it is more accurate and less susceptible to the implementation choice.

Since Z follows a CIR process under $\mathbb{P}^*$, a key result of Broadie & Kaya (2006) implies that for times $0 \le t_1 < t_2 \le \tau$ and points $z_1, z_2 \in \mathbb{R}_+$,

$$\begin{aligned}\phi(z_1, t_1; z_2, t_2) &= \mathbb{E}^*\left[\exp\left(-\int_{t_1}^{t_2} \lambda_u \mathrm{d}u\right) \,\middle|\, D_\tau, Z_{t_1} = z_1, Z_{t_2} = z_2\right] \\ &= \Phi(z_1, t_1; z_2, t_2) \exp\left(-\int_{t_1}^{t_2} \left(e^{a\cdot(1,X_u)} + Y_u\right) \mathrm{d}u\right),\end{aligned} \tag{7}$$

where

$$\Phi(z_1, t_1; z_2, t_2) = \frac{I_q\left(\sqrt{z_1 z_2}\frac{4\zeta e^{-0.5\zeta\Delta}}{1-e^{-\zeta\Delta}}\right)}{I_q\left(\sqrt{z_1 z_2}\frac{4k e^{-0.5k\Delta}}{1-e^{-k\Delta}}\right)} \frac{\zeta e^{-0.5(\zeta-k)\Delta}(1-e^{-k\Delta})}{k(1-e^{-\zeta\Delta})} e^{(z_1+z_2)\left(\frac{k(1+e^{-k\Delta})}{1-e^{-k\Delta}} - \frac{\zeta(1+e^{-\zeta\Delta})}{1-e^{-\zeta\Delta}}\right)}$$

for $\Delta = t_2 - t_1$, $\zeta = \sqrt{k^2 + 2c^2}$, and I_q the modified Bessel function of the first kind of order $q = \frac{2kz}{c^2} - 1$. An application of the law of iterated expectations leads to a reformulation of the likelihood in terms of a product of terms as in (7).

Proposition B.1. *For $t \leq \tau$ and a function u on $\mathbb{R}_+$ such that $u(\lambda_t)$ is integrable,*

$$\mathbb{E}^*\left[u(\lambda_t)M_t \mid D_t\right] = e^t \mathbb{E}^*\left[u(\lambda_t)\phi(Z_{T_{N_\tau}}, T_{N_\tau}; Z_\tau, \tau)\prod_{n=1}^{N_\tau}\lambda_{T_n-}\phi(Z_{T_{n-1}}, T_{n-1}; Z_{T_n}, T_n) \,\middle|\, D_t\right].$$

Proof. Letting $\Pi_t = u(\lambda_t)\exp\left(\sum_{n=1}^{N_t}\log(\lambda_{T_n-})\right)$, we calculate

$$\begin{aligned}
&\mathbb{E}^*\left[u(\lambda_t)M_t \mid D_t\right]\exp(-t)\\
&= \mathbb{E}^*\left[\Pi_t \exp\left(-\int_0^{T_{N_t}}\lambda_u \mathrm{d}u\right)\mathbb{E}^*\left[\exp\left(-\int_{T_{N_t}}^{t}\lambda_u \mathrm{d}u\right) \,\middle|\, D_t, (Z_u)_{u\leq T_{N_t}}, Z_t\right] \,\middle|\, D_t\right]\\
&= \mathbb{E}^*\left[\Pi_t \exp\left(-\int_0^{T_{N_t}}\lambda_u \mathrm{d}u\right)\phi(Z_{T_{N_t}}, T_{N_t}; Z_t, t) \,\middle|\, D_t\right]. \qquad (8)
\end{aligned}$$

Line (8) uses the fact that, by Girsanov's theorem, the change of measure induced by M_τ does not affect the law of Z. Consequently, Z is a Markov process under $\mathbb{P}^*$. We iterate to complete the proof. □

Proposition B.1 allows us to approximate the likelihood (6) using a trapezoidal quadrature rule that does not interpolate the path of the frailty Z. Algorithm B.2 summarizes the numerical scheme for terms of the form $\mathbb{E}^*\left[u(\lambda_\tau)M_\tau \mid D_\tau\right]$; the likelihood is given by $u \equiv 1$. Convergence of the scheme as the discretization becomes finer follows along the lines of Theorem 4.1 of Giesecke & Schwenkler (2014). Note that the transition law of Z is non-central chi-squared since Z follows a CIR process.

Algorithm B.2. *Fix a state-space discretization $\{z^1, \ldots, z^m\}$ for the frailty Z. Let $\mathcal{A}^i$ be a neighborhood of z^i. Define the terms*

$$\begin{aligned}
F^n(k,l) &= e^{T_n - T_{n-1}}\left(e^{a\cdot(1, X_{T_n-})} + Y_{T_n-} + cz^k\right)\phi(z^l, T_{n-1}; z^k, T_n),\\
p^n(k,l) &= \mathbb{P}^*\left[Z_{T_n} \in \mathcal{A}^k \,\middle|\, D_\tau, Z_{T_{n-1}} = z^l\right],
\end{aligned}$$

for $1 \leq k, l \leq m$ and $1 \leq n \leq N_\tau$. In addition, set $T_{N_\tau+1} = \tau$ and

$$\begin{aligned}
F^{N_\tau+1}(k,l) &= e^{\tau - T_{N_\tau}}\, u\left(e^{a\cdot(1, X_\tau)} + Y_\tau + cz^k\right)\phi(z^l, T_{N_\tau}; z^k, \tau)\\
p^{N_\tau+1}(k,l) &= \mathbb{P}^*\left[Z_\tau \in \mathcal{A}^k \,\middle|\, D_\tau, Z_{T_{N_\tau}} = z^l\right],
\end{aligned}$$

For a given $\theta \in \Theta$, do:

(1) Initialization: Let $\xi = (\xi(1), \ldots, \xi(m))$ be an m-dimensional vector such that $\xi(i) = 1$ if $y \in \mathcal{A}^i$ and otherwise $\xi(i) = 0$.

(2) Iteration: For $n = 1, \ldots, N_\tau + 1$ do

(a) Define the $m \times m$-matrix $\hat{Q}_j$ with elements $F^n(k,l)p^n(k,l)$

(b) Update the vector ξ to $\xi \leftarrow \hat{Q}_j \cdot \xi$

(3) Termination: Compute an approximation of $\mathbb{E}^*\left[u(\lambda_\tau)M_\tau \mid D_\tau\right]$ as $\sum_{i=1}^m \xi(i)$.

B.3 Asymptotic properties

Assume that the true parameter θ_0 lies in the interior of the parameter space Θ. If $z = \sigma = 0$, we can compute an MLE $\hat{\theta}_\tau$ analytically and Theorem 3.2 of Giesecke & Schwenkler (2014) implies that the MLE is consistent.[16] Under regularity and identifiability conditions, Theorem 3.3 of Giesecke & Schwenkler (2014) states that $\hat{\theta}_\tau$ will be asymptotically normal so that $\sqrt{\tau}(\hat{\theta}_\tau - \theta_0)$ converges in distribution to a multivariate normal distribution with mean 0 and variance-covariance matrix

$$\Sigma = \left(- \lim_{\tau \to \infty} \frac{1}{\tau} \nabla^2 \log \mathcal{L}_\tau(\theta_0) \right)^{-1}.$$

We use the following finite-horizon approximation of Σ:

$$\hat{\Sigma}_\tau = \left(-\frac{1}{\tau} \nabla^2 \log \mathcal{L}_\tau(\theta_0) \right)^{-1}.$$

If $\sigma > 0$ and $z > 0$, the likelihood is approximated using Algorithm B.2. Define $\mathcal{L}^A_\tau(\theta)$ as the approximate likelihood computed by Algorithm B.2. An approximate MLE $\hat{\theta}^A_\tau$ maximizes the approximate likelihood function:

$$\hat{\theta}^A_\tau \in \arg\max_{\theta \in \Theta} \mathcal{L}^A_\tau(\theta).$$

Proposition 4.3 of Giesecke & Schwenkler (2014) implies that, under mild technical conditions, the approximate MLE $\hat{\theta}^A_\tau$ will be consistent and asymptotically normal as the sample period grows and the discretization becomes finer. A finite-horizon approximation of the asymptotic variance-covariance matrix is given by

$$\hat{\Sigma}^A_\tau = \left(-\frac{1}{\tau} \nabla^2 \log \mathcal{L}^A_\tau(\hat{\theta}^A_\tau) \right)^{-1}.$$

B.4 Implementation

The numerical algorithm and approximate likelihood maximization are implemented in R and run on an eight-core 32GB Sun blade system. We use the Nelder-Mead optimization method to locate the maximizers of the likelihood function. To address the issue of local optima, we run a number of optimizations with random initial values. In order to deal with very small or very large values of the intensity and the parameters, we scale the time unit to one week.

[16] Condition (A3) of Giesecke & Schwenkler (2014) is irrelevant in our setting because the parameters driving the explanatory factor X can be estimated separately from the default parameter θ.

C Posterior mean of intensity

Since the data do not include observations of the frailty Z, the intensity λ_t cannot be measured unless $z = \sigma = 0$. We consider the posterior mean of λ_t, denoted by h_t. This appendix explains the computation of h_t.

Let $\mathbb{G} = (\mathcal{G}_t)_{t\geq 0}$ be the right-continuous and complete filtration generated by the observable data, i.e., $\mathcal{G}_t = \sigma(D_t) \subset \mathcal{F}_t$. The posterior mean h_t of λ_t is the optional projection of λ onto $\mathbb{G}$.[17] It satisfies

$$h_t = \mathbb{E}\left[\lambda_t \mid \mathcal{G}_t\right] \tag{9}$$

almost surely, and represents the default rate given the observable data available at t (more precisely, it is the intensity relative to $\mathbb{G}$). We exploit the change of measure defined by (5) to efficiently compute h_t using two calls of Algorithm B.2, one with $u(\lambda) \equiv 1$ and one with $u(\lambda) = \lambda$. This is based on the following result.

Proposition C.1. *Suppose* $\mathbb{E}[1/M_\tau] = 1$ *for* M_τ *defined in* (5)*. For* $t \leq \tau$*, the posterior mean* h_t *satisfies almost surely*

$$h_t = \frac{\mathbb{E}^*\left[\lambda_t M_t \mid D_t\right]}{\mathbb{E}^*\left[M_t \mid D_t\right]}. \tag{10}$$

Proof. Theorem T3 of Section VI of Brémaud (1980) states that, given $\mathbb{E}[1/M_\tau] = 1$, the exponential martingale M induces an equivalent probability measure $\mathbb{P}^*$ on the σ-field $\mathcal{F}_\tau$ with Radon-Nikodym density $\frac{\mathrm{d}\mathbb{P}^*}{\mathrm{d}\mathbb{P}} = 1/M_\tau$. Moreover, the process $(1/M_t)_{t\leq\tau}$ is a martingale with respect to $\mathbb{F}$ by Theorem II.T8 of Brémaud (1980). Since $M_t > 0$ for all $t \leq \tau$, it follows that $\mathbb{P}$ is also absolutely continuous with respect to $\mathbb{P}^*$ on $\mathcal{F}_t$ with Radon-Nikdoym density

$$\left.\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{P}^*}\right|_{\mathcal{F}_t} = M_t$$

for $t \leq \tau$. Optional projection indicates that $\mathbb{P}$ is also absolutely continuous with respect to $\mathbb{P}^*$ on the σ-algebra $\mathcal{G}_t \subseteq \mathcal{F}_t$ for all $t \leq \tau$ with Radon-Nikodym density $\mathbb{E}^*\left[M_t \,|\, \mathcal{G}_t\right]$. Now, λ is positive and $\mathbb{F}$-adapted. This implies that

$$h_t = \mathbb{E}\left[\lambda_t \,|\, \mathcal{G}_t\right] = \frac{\mathbb{E}^*\left[\lambda_t M_t \mid \mathcal{G}_t\right]}{\mathbb{E}^*\left[M_t \mid \mathcal{G}_t\right]}.$$

In order to obtain (10), note that $\mathcal{G}_t = \sigma(D_t)$. □

D Goodness-of-fit tests

This appendix extends a key result of Das et al. (2007) for doubly-stochastic models with observable risk factors. The extension allows us to construct goodness-of-fit tests for

[17] See Protter (2004) for details on the optional projection.

our default timing model (1), which violates the doubly-stochastic hypothesis due to the presence of the frailty and self-exciting terms. A key role is played by the posterior mean h_t of the intensity, which is discussed in Appendix C.

Proposition D.1. *Let H_t be the right-continuous inverse to $C_t = \int_0^t h_u \mathrm{d}u$, i.e.,*

$$H_t = \max\left\{ s \geq 0 : \int_0^s h_u \mathrm{d}u \leq t \right\}. \tag{11}$$

The time-changed process J defined by $J_t = N_{H_t}$ is a standard Poisson process on $[0, \infty)$ relative to $\mathbb{P}$ and the minimal right-continuous completion $\mathbb{H}$ of $(\mathcal{G}_{H_t})_{t \geq 0}$.

Proof. A result of Meyer (1971) implies that a counting process with compensator that is continuous and increasing to ∞ almost surely can be transformed into a standard Poisson process by a change of time given by the compensator. Relative to the filtration $\mathbb{G}$ generated by the observable data (see Appendix C), the counting process N has compensator C given by $C_t = \int_0^t h_u \mathrm{d}u$, where h is the posterior mean of the intensity given by Proposition C.1. Since each H_t is a stopping time with respect to $\mathbb{G}$, we can define the stopping time σ-algebra $\mathcal{H}_t = \mathcal{G}_{H_t}$, which is the smallest σ-algebra containing all right-continuous left limit processes sampled at H_t. Meyer's theorem implies that the time-scaled process J is a standard Poisson process in the time-changed filtration $\mathbb{H}$, which is the minimal right-continuous completion of $(\mathcal{H}_t)_{t \geq 0}$. □

Proposition D.1 states that the default count N_t can be transformed into a standard Poisson process by a change of time given by the cumulative posterior mean of the intensity. The posterior mean is given by Proposition C.1, and can be computed using Algorithm B.2. This result allows us to test model fit by testing the Poisson property of the time-scaled default times.

E Out-of-sample simulation

This appendix describes the simulation procedure we use to forecast the conditional distribution of the total number of defaults out-of-sample.

Inspired by Das et al. (2007), Duffie et al. (2007) and Duffie et al. (2009), we employ a daily Gaussian vector auto-regression of order 41 for the vector containing the rolling S&P 500 yearly return and volatility, the yield of the 3-month Treasury Bill, and the spread between the yields of the 10-year and the 1-year Treasury bonds. We employ a monthly ARMA(1,1) model for the growth rate of the industrial production growth, a quarterly MA(1) model for the GDP growth rate, and a weekly AR(1) model for the spread between the yields of BAA and AAA rated corporate bonds. We choose the optimal autoregression and moving-average orders according to the Akaike information criterion.

Using data D_t available at time t, we estimate the models of the explanatory variables and the models of default timing in Table 3. We then compute the model-implied

conditional distribution of the number of defaults during $(t, t+1]$. The distribution is calculated by Monte Carlo simulation of default times using 100,000 trials. The simulation is based on the (daily) discretization of the posterior mean h_t during $(t, t+1]$ as described in Giesecke & Teng (2012), and the simulation of the explanatory variables during that interval. We interpolate the simulated paths of the less frequently observed explanatory variables to obtain a series of daily values for all variables.

The feasibility of event simulation using h_t along with our ability to compute h_t using Proposition C.1 eliminates the need to generate paths of the frailty during $(t, t+1]$, which would be required if the simulation were based on the intensity λ. This increases the efficiency of event prediction by an order of magnitude relative to the prediction based on λ that is standard in the frailty literature.

References

Aït-Sahalia, Yacine, Julio Cacho-Diaz & Roger Laeven (2013), Modeling financial contagion using mutually exciting jump processes. Working Paper.

Aragon, George O. & Philip E. Strahan (2012), 'Hedge funds as liquidity providers: Evidence from the lehman bankruptcy', *Journal of Financial Economics* **103**(3), 570 – 587.

Berndt, Antje, Peter Ritchken & Zhiqiang Sun (2010), 'On correlation and default clustering in credit markets', *Review of Financial Studies* **23**(7), 2680–2729.

Boone, Audra L. & Vladimir I. Ivanov (2012), 'Bankruptcy spillover effects on strategic alliance partners', *Journal of Financial Economics* **103**(3), 551 – 569.

Bowsher, Clive (2007), 'Modelling security market events in continuous time: Intensity based, multivariate point process models', *Journal of Econometrics* **141**, 876–912.

Brémaud, Pierre (1980), *Point Processes and Queues – Martingale Dynamics*, Springer-Verlag, New York.

Broadie, Mark & Ozgur Kaya (2006), 'Exact simulation of stochastic volatility and other affine jump diffusion processes', *Operations Research* **54**(2), 217–231.

Brunnermeier, Markus K. (2009), 'Deciphering the liquidity and credit crunch 2007-2008', *Journal of Economic Perspectives* **23**(1), 77–100.

Cameron, A. Colin & Frank A. G. Windmeijer (1996), 'R-squared measures for count data regression models with applications to health-care utilization', *Journal of Business & Economic Statistics* **14**(2), 209–220.

Cameron, A. Colin & Pravin K. Trivedi (1998), *Regression Analysis of Count Data*, Econometric Society Monographs, Cambridge University Press.

Campbell, John Y., Jens Hilscher & Jan Szilagyi (2008), 'In search of distress risk', *Journal of Finance* **63**(6), 2899–2939.

Chakrabarty, Bidisha & Gaiyan Zhang (2012), 'Credit contagion channels: Market microstructure evidence from lehman brothers' bankruptcy', *Financial Management* **41**(2), 320–343.

Charles, Amélie, Olivier Darné & Jae H. Kim (2011), 'Small sample properties of alternative tests for martingale difference hypothesis', *Economics Letters* **110**(2), 151 – 154.

Chava, Sudheer & Robert Jarrow (2004), 'Bankruptcy prediction with industry effects', *Review of Finance* **8**, 537–569.

Christoffersen, Peter (1998), 'Evaluating interval forecasts', *International Economic Review* **39**, 842–862.

Collin-Dufresne, Pierre, Robert Goldstein & Jean Helwege (2009), How large can jump-to-default risk premia be? Modeling contagion via the updating of beliefs. Working Paper, Columbia University.

Collin-Dufresne, Pierre, Robert Goldstein & Julien Hugonnier (2004), 'A general formula for the valuation of defaultable securities', *Econometrica* **72**, 1377–1407.

Collin-Dufresne, Pierre, Robert S. Goldstein & J. Spencer Martin (2001), 'The determinants of credit spread changes', *Journal of Finance* **56**(6), 2177–2207.

Cox, D. R. (1972), 'Regression models and life-tables', *Journal of the Royal Statistical Society. Series B (Methodological)* **34**(2), 187–220.

Das, Sanjiv, Darrell Duffie, Nikunj Kapadia & Leandro Saita (2007), 'Common failings: How corporate defaults are correlated', *Journal of Finance* **62**, 93–117.

Delloye, Martin, Jean-David Fermanian & Mohammed Sbai (2006), 'Dynamic frailties and credit portfolio modeling', *Risk* **19**(1), 101–109.

Domínguez, Manuel A. & Ignacio N. Lobato (2003), 'Testing the martingale difference hypothesis', *Econometric Reviews* **22**(4), 351–377.

Duffie, Darrell, Andreas Eckner, Guillaume Horel & Leandro Saita (2009), 'Frailty correlated default', *Journal of Finance* **64**, 2089–2123.

Duffie, Darrell, Leandro Saita & Ke Wang (2007), 'Multi-period corporate default prediction with stochastic covariates', *Journal of Financial Economics* **83**(3), 635–665.

Duffie, Darrell & Nicolae Garleanu (2001), 'Risk and valuation of collateralized debt obligations', *Financial Analysts Journal* **57**(1), 41–59.

Dumontaux, Nicolas & Adrian Pop (2013), 'Understanding the market reaction to shockwaves: Evidence from the failure of lehman brothers', *Journal of Financial Stability* **9**(3), 269 – 286.

Engle, Robert & Simone Manganelli (2004), 'Caviar: Conditional autoregressive value at risk by regression quantiles', *Journal of Business and Economic Statistics* **22**(4), 367–381.

Errais, Eymen, Kay Giesecke & Lisa Goldberg (2010), 'Affine point processes and portfolio credit risk', *SIAM Journal on Financial Mathematics* **1**, 642–665.

Feldhütter, Peter & Mads Stenbo Nielsen (2012), 'Systematic and idiosyncratic default risk in synthetic credit markets', *Journal of Financial Econometrics* **10**(2), 292–324.

Fernando, Chitru S., Anthony D. May & William L. Megginson (2012), 'The value of investment banking relationships: Evidence from the collapse of lehman brothers', *The Journal of Finance* **67**(1), 235–270.

Giesecke, Kay (2004), 'Correlated default with incomplete information', *Journal of Banking and Finance* **28**, 1521–1545.

Giesecke, Kay, Francis A. Longstaff, Stephen Schaefer & Ilya Strebulaev (2011), 'Corporate bond default risk: A 150-year perspective', *Journal of Financial Economics* **102**(2), 233 – 250.

Giesecke, Kay, Francis A. Longstaff, Stephen Schaefer & Ilya Strebulaev (2014), 'Macroeconomic effect of corporate default crises: A long-term perspective', *Journal of Financial Economics* **111**(2), 297–310.

Giesecke, Kay & Gerald Teng (2012), Numerical solution of jump-diffusion SDEs. Working Paper, Stanford University.

Giesecke, Kay & Gustavo Schwenkler (2014), Filtered likelihood for point processes. Working paper.

Gilchrist, Simon & Egon Zakrajšek (2012), 'Credit spreads and business cycle fluctuations', *American Economic Review* **102**(4), 1692–1720.

Hamilton, David (2005), Moodys senior ratings algorithm and estimated senior ratings. Moodys Investors Service.

Hawkes, Alan G. (1971), 'Spectra of some self-exciting and mutually exciting point processes', *Biometrika* **58**(1), 83–90.

Hendricks, Darryll (1996), 'Evaluation of value-at-risk models using historical data', *Economic Policy Review* **2**(1), 39–69.

Jarrow, Robert A. & Fan Yu (2001), 'Counterparty risk and the pricing of defaultable securities', *Journal of Finance* **56**(5), 555–576.

Jorion, Philippe & Gaiyan Zhang (2007), 'Good and bad credit contagion: Evidence from credit default swaps', *Journal of Financial Economics* **84**(3), 860–883.

Karlis, Dimitris & Evdokia Xekalaki (2000), 'A simulation comparison of several procedures for testing the Poisson assumption', *The Statistician* **49**, 355–382.

Koopman, Siem Jan, Andre Lucas & Andre Monteiro (2008), 'The multi-stage latent factor intensity model for credit rating transitions', *Journal of Econometrics* **142**(1), 399–424.

Koopman, Siem Jan, Andre Lucas & Bernd Schwaab (2011), 'Modeling frailty-correlated defaults using many macroeconomic covariates', *Journal of Econometrics* **162**(2), 312–325.

Kupiec, Paul H. (1995), 'Techniques for verifying the accuracy of risk measurement models', *Journal of Derivatives* **3**(2), 73–84.

Lando, David & Mads Stenbo Nielsen (2010), 'Correlation in corporate defaults: Contagion or conditional independence?', *Journal of Financial Intermediation* **19**(3), 355–372.

Lang, Larry & Rene Stulz (1992), 'Contagion and competitive intra-industry effects of bankruptcy announcements', *Journal of Financial Economics* **32**, 45–60.

Longstaff, Francis & Arvind Rajan (2008), 'An empirical analysis of collateralized debt obligations', *Journal of Finance* **63**(2), 529–563.

Meyer, Paul-André (1971), Démonstration simplifée d'un théorème de Knight, *in* 'Séminaire de Probabilités V, Lecture Notes in Mathematics 191', Springer-Verlag Berlin, pp. 191–195.

Mortensen, Allan (2006), 'Semi-analytical valuation of basket credit derivatives in intensity-based models', *Journal of Derivatives* **13**, 8–26.

Protter, Philip (2004), *Stochastic Integration and Differential Equations*, Springer-Verlag, New York.

Shumway, Tyler (2001), 'Forecasting bankruptcy more accurately: A simple hazard model', *Journal of Business* **74**, 101–124.

Stulz, Rene M. (2010), 'Credit default swaps and the credit crisis', *Journal of Economic Perspectives* **24**(1), 73–92.

Winkelmann, Rainer (2008), *Econometric Analysis of Count Data*, Springer.

Variable	Mean	Standard deviation	Skewness	Kurtosis	Min.	First quartile	Median	Third quartile	Max.	Correlation
Number of defaults per month	3.88	4.42	1.81	4.29	0	1	3	6	29	1.00
Total debt outstanding at default per month (millions of USD)	671.53	1265.53	4.09	23.23	0.00	1.31	179.76	776.59	11314.78	0.72
Growth rate of U.S. industrial production	2.32	4.14	−0.87	1.55	−13.99	0.55	2.66	4.79	10.73	−0.42
Growth rate of U.S. GDP (annualized)	6.71	4.27	0.76	2.96	−8.40	4.40	6.05	8.53	25.50	−0.46
Trailing 1-year return of S&P 500	8.11	17.06	−0.42	0.05	−42.51	−2.33	9.92	20.38	52.64	−0.28
Trailing 1-year volatility of S&P 500	17.89	7.82	0.46	−0.49	1.80	11.34	17.28	23.00	38.31	0.12
3M Treasury Bill yield	5.28	3.23	0.56	0.68	0.01	3.27	5.09	7.12	16.30	−0.40
10Y-to-1Y Treasury bond yield spread	1.05	1.26	−0.34	−0.32	−3.07	0.24	1.08	1.92	3.40	0.30
BAA-to-AAA corporate bond yield spread	1.12	0.46	1.69	3.65	0.55	0.80	0.98	1.31	3.38	0.04

Table 1: *Summary statistics of monthly data*. This table provides summary statistics for the time series of monthly defaults, total debt outstanding at default per month, and explanatory variables (sampled monthly). Each series has 516 values. The column "Correlation" indicates the linear correlation coefficient between monthly defaults and another variable.

	Regression 1		Regression 2	
Variable	Coefficient	t-statistic	Coefficient	t-statistic
Constant	2.651	*** 23.416	0.715	*** 3.916
Industrial production growth	−0.066	*** −8.872	−0.029	*** −3.320
GDP growth	−0.105	*** −13.234	−0.065	*** −7.200
S&P 500 yearly return	−0.008	*** −4.733	−0.004	−1.927
S&P 500 yearly volatility	0.016	*** 5.055	0.006	1.798
3-month T-Bill yield	−0.043	*** −3.416	0.032	* 2.260
Spread, 10Y vs. 1Y Treasury bond yields	−0.081	** −2.675	−0.078	** −2.604
Spread, BAA vs. AAA corporate bond yields	−0.689	*** −11.385	−0.398	*** −6.437
Lag one defaults per month			0.045	*** 6.403
Lag one total debt outstanding at default per month			0.169	*** 8.954
Deviance R^2	0.3958		0.5639	
Likelihood ratio statistic, Regression 1 vs. Regression 2	*** 397.34			

Table 2: *Coefficient estimates and t-statistics of Poisson regressions of monthly defaults.* Regression 1: *Expected number of defaults in* $[t, t+1/12) =$ *exp (Constant +* β_1 *Industrial production growth at time t +* β_2 *GDP growth at time t +* β_3 *S&P500 return at time t +* β_4 *S&P500 volatility at time t +* β_5 *3M Treasury yield at time t +* β_6 *Treasury yield spread at time t +* β_7 *Corporate yield spread at time t.* Regression 2: the variables *Number of defaults in* $[t-1, t-1+1/12)$ and *Total debt outstanding at default in* $[t-1, t-1+1/12)$ are included as additional regressors. * indicates significance at the 95% level, ** significance at the 99% level, and *** significance at the 99.9% level. The regressions are based on 515 data points. The deviance R^2 is computed in accordance with Cameron & Windmeijer (1996). The likelihood ratio test statistic has an asymptotic chi-squared distribution with one degree of freedom.

Model	**Parameter Restrictions**	**Intensity**
Complete	None	$\lambda_t = \exp(a_0 + \sum_{i=1}^{d} a_i X_{i,t}) + Y_t + Z_t$
Base + Frailty	$b = 0$	$\lambda_t = \exp(a_0 + \sum_{i=1}^{d} a_i X_{i,t}) + Z_t$
Base + Contagion	$z = \sigma = 0$	$\lambda_t = \exp(a_0 + \sum_{i=1}^{d} a_i X_{i,t}) + Y_t$
Base	$b = z = \sigma = 0$	$\lambda_t = \exp(a_0 + \sum_{i=1}^{d} a_i X_{i,t})$

Table 3: *Reduced-form model alternatives.* The table reports the intensity specification for each reduced-form model we estimate. Each model addresses a different set of clustering sources.

	Complete	Base	Base + Contagion	Base + Frailty
Constant a_0	** −1.6926	*** 1.0141	** −1.6639	0.5649
	(0.6765)	(0.1349)	(0.6453)	(0.8971)
Industrial Production coefficient a_1	−0.0013	−0.0509	−0.0027	−0.0616
	(0.0334)	(0.0086)	(0.0318)	(0.0471)
GDP Growth coefficient a_3	*** −0.1235	*** −0.1117	*** −0.1197	0.0031
	(0.0360)	(0.0098)	(0.0343)	(0.0391)
S&P 500 Return coefficient a_4	0.0017	* −0.0045	0.0015	−0.0193
	(0.0078)	(0.0018)	(0.0074)	(0.0173)
S&P 500 Volatility coefficient a_5	0.0038	** 0.0112	0.0041	0.0123
	(0.0208)	(0.0038)	(0.0222)	(0.0246)
3M T-Bill Yield coefficient a_6	0.0284	** −0.0466	0.0277	*** −0.6734
	(0.0612)	(0.0149)	(0.0857)	(0.1546)
10Y-1Y Treasury Yield Spread coefficient a_8	−0.0007	* −0.0821	−0.0019	0.4748
	(0.1751)	(0.0352)	(0.2770)	(0.2836)
BAA-AAA Corporate Yield Spread coefficient a_9	−0.01804	*** −0.7305	−0.0192	−0.0971
	(0.3907)	(0.0741)	(0.4253)	(0.3035)
Frailty volatility σ	0.2221			*** 0.1288
	(0.1602)			(0.0146)
Frailty mean reversion rate k	* 6.0203			*** 0.0156
	(2.9545)			(0.0002)
Frailty mean reversion level z	* 0.0041			*** 0.5335
	(0.0020)			(0.0577)
Contagion sensitivity b	*** 0.0103		*** 0.0101	
	(0.0017)		(0.0017)	
Contagion decay rate κ	*** 0.0571		*** 0.0563	
	(0.0096)		(0.0098)	
Log-likelihood	606.29	446.12	606.21	599.24

Table 4: Parameter estimates for each of the four model alternatives described in Table 3. The asymptotic standard errors are given parenthetically. They are computed using the Hessian matrix of the log-likelihood at the parameter estimates reported. * indicates significance at the 95% level, ** significance at the 99% level, and *** significance at the 99.9% level.

	Complete	Base	Base + contagion	Base + frailty
Automatic Portmanteau (AQ) test	0.235	*** 0.000	0.225	0.438
Automatic Variance Ratio (AVR) test	0.181	*** 0.000	0.169	0.600
Cramer-von Mises (CvM) test	0.704	*** 0.000	0.679	0.073
Kolmogorov-Smirnov (KS) test	0.831	*** 0.000	0.783	0.064

Table 5: *Martingale difference tests.* We perform a series of tests of the martingale difference hypothesis for the increments of the fitted M_t in (4) between defaults. A martingale difference sequence has no dependence in mean. The tests evaluate both linear (AQ test) as well as non-linear dependence in mean (AVR, CvM, and KS tests). The asymptotic distribution of the AQ test is chi-squared with one degree of freedom. The p-values of the remaining tests are computed via bootstrapping with $1,000$ bootstrap samples. For the AVR test, we perform a wild bootstrap based on Mammen's two-point distribution. The CvM and KS test statistics are computed as indicated by Domínguez & Lobato (2003). *** indicates significance at the 99.9% level.

Benchmark model Alternative model	Base Base + contagion	Base Base + frailty	Base Complete	Base + contagion Complete	Base + frailty Complete
Test statistic	320.18	299.60	320.34	0.16	14.10
Degrees of freedom	2	3	5	3	2
p-value	*** 0.000	*** 0.000	*** 0.000	0.984	*** 0.001

Table 6: *Likelihood ratio tests*. This table reports likelihood ratio test statistics, p-values, and degrees of freedom for the corresponding asymptotic distributions. A likelihood ratio test evaluates the fit of an alternative model relative to a nested benchmark model. The test statistic is given by twice the difference between the maximum log-likelihood of the alternative model and the benchmark model. The log-likelihood values are reported in Table 4. The test statistic has, asymptotically, a chi-squared distribution with degrees of freedom equal to the number of additional parameters included in the alternative. *** indicates significance at the 99.9% level.

$w = 2$					
	Observations	Mean	Variance	Skewness	Kurtosis
Theoretical		2.000	2.000	0.707	3.500
Base model	727	1.997	2.617	1.134	6.161
Base + contagion model	729	1.992	1.585	0.565	3.268
Base + frailty model	775	1.874	1.302	0.613	3.606
Complete model	730	1.989	1.597	0.611	3.206
$w = 4$					
	Observations	Mean	Variance	Skewness	Kurtosis
Theoretical		4.000	4.000	0.500	3.250
Base model	364	3.989	6.997	0.846	4.581
Base + contagion model	365	3.978	3.247	0.391	2.989
Base + frailty model	388	3.742	2.466	0.393	2.999
Complete model	365	3.978	3.527	0.592	3.301
$w = 6$					
	Observations	Mean	Variance	Skewness	Kurtosis
Theoretical		6.000	6.000	0.408	3.167
Base model	243	5.975	13.983	0.800	4.274
Base + contagion model	243	5.975	5.371	0.343	2.854
Base + frailty model	259	5.606	3.720	0.309	3.203
Complete model	244	5.951	5.586	0.232	2.690
$w = 8$					
	Observations	Mean	Variance	Skewness	Kurtosis
Theoretical		8.000	8.000	0.354	3.125
Base model	182	7.978	21.900	0.570	3.618
Base + contagion model	183	7.934	7.655	0.206	3.055
Base + frailty model	194	7.485	4.987	0.069	2.572
Complete model	183	7.934	7.633	0.236	2.488
$w = 10$					
	Observations	Mean	Variance	Skewness	Kurtosis
Theoretical		10.000	10.000	0.316	3.100
Base model	146	9.945	33.087	0.502	3.353
Base + contagion model	146	9.945	9.431	0.095	3.093
Base + frailty model	155	9.368	6.104	0.168	2.876
Complete model	146	9.945	9.680	0.130	2.549

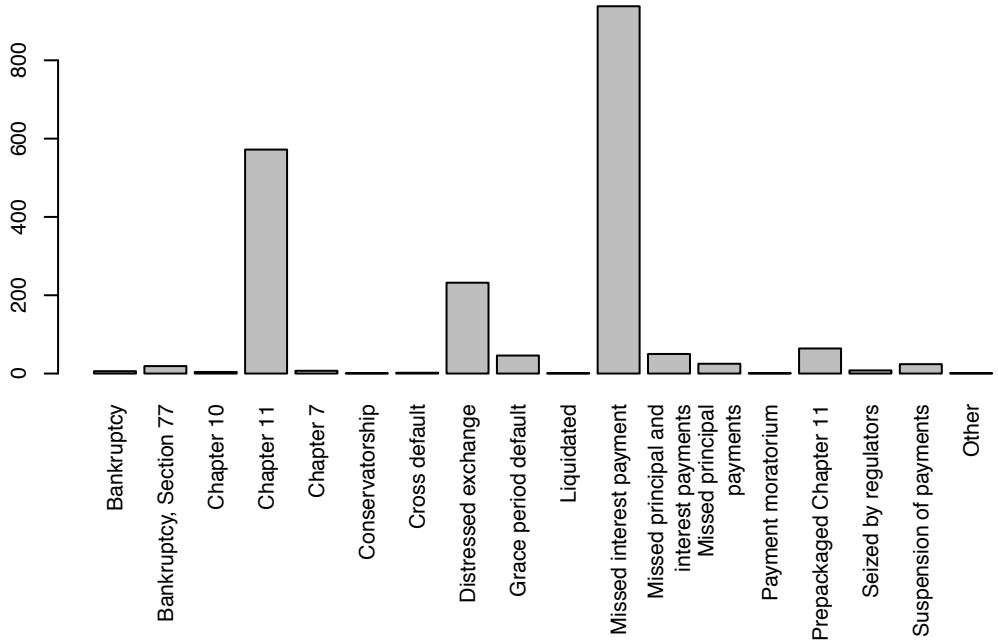
Table 7: *Moments of binned event counts.* This table reports the first four empirical moments of the binned event counts P_i^w for bin sizes $w \in \{2, 4, 6, 8, 10\}$, for each alternative model. We also report the moments of the theoretical Poisson distributions with rate w.

Potthoff-Whittinghill-Bohning Test					Kocherlakota-Kocherlakota Test				
w	Base	Base + frailty	Base + contagion	Complete	w	Base	Base + frailty	Base + contagion	Complete
2	*** 0.000	*** 0.000	*** 0.000	*** 0.000	2	0.575	0.508	0.688	0.693
4	*** 0.000	*** 0.000	* 0.012	0.126	4	0.158	0.466	0.719	0.807
6	*** 0.000	*** 0.000	0.253	0.499	6	* 0.015	0.482	0.833	0.916
8	*** 0.000	** 0.001	0.737	0.717	8	** 0.001	0.515	0.965	0.949
10	*** 0.000	** 0.002	0.660	0.820	10	*** 0.000	0.489	0.952	0.978
Independence Test					Fisher Dispersion Test				
w	Base	Base + frailty	Base + contagion	Complete	w	Base	Base + frailty	Base + contagion	Complete
2	*** 0.000	0.911	0.961	0.517	2	*** 0.000	1.000	1.000	1.000
4	*** 0.000	0.655	0.148	0.163	4	*** 0.000	1.000	0.997	0.949
6	*** 0.000	* 0.030	0.118	0.064	6	*** 0.000	1.000	0.884	0.771
8	*** 0.000	0.074	0.296	** 0.004	8	*** 0.000	1.000	0.647	0.657
10	*** 0.000	** 0.002	* 0.018	** 0.001	10	*** 0.000	1.000	0.674	0.592
Upper Tail Test					Chi-Squared Test				
w	Base	Base + frailty	Base + contagion	Complete	w	Base	Base + frailty	Base + contagion	Complete
2	** 0.004	1.000	0.994	0.985	2	*** 0.000	*** 0.000	*** 0.000	*** 0.000
4	*** 0.000	1.000	0.921	0.796	4	*** 0.000	*** 0.000	** 0.008	0.115
6	*** 0.000	1.000	0.742	0.630	6	*** 0.000	*** 0.000	0.248	0.517
8	*** 0.000	1.000	0.741	0.672	8	*** 0.000	*** 0.000	0.763	0.746
10	*** 0.000	1.000	0.769	0.632	10	*** 0.000	*** 0.000	0.683	0.849

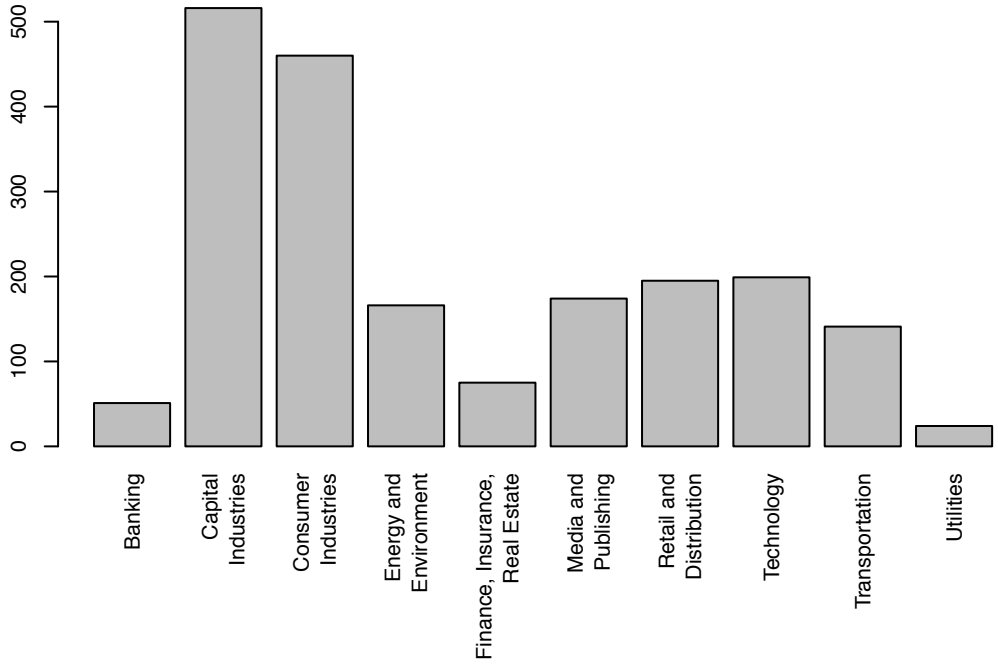
Table 8: *Poisson distribution tests.* p-values of Poisson distribution tests for the counts of the binned event counts P_i^w. The Potthoff-Whittinghill-Bohning test compares the empirical moments of the P_i^w to the theoretical moments. Its asymptotic distribution is standard normal (Karlis & Xekalaki (2000)). The Kocherlakota-Kocherlakota test evaluates the empirical moment generating function of the P_i^w and compares to its theoretical counterpart, see Karlis & Xekalaki (2000). The upper tail test examines the fatness of the right tail of the empirical distribution of the P_i^w; see Das et al. (2007). We compute empirical p-values for the Kocherlakota-Kocherlakota and the upper tail tests via bootstrapping with $100,000$ samples. The independence test is a BDS test with a distance parameter of one standard deviation. The Fisher dispersion test examines the variance of the realized P_i^ws and is described in Das et al. (2007). Finally, we also perform a standard chi-squared goodness of fit test for the Poisson distribution. The number of samples of P_i^w for each model are as follows (ordered by increasing bin size). Base: 727, 364, 243, 182, 146. Base + frailty: 775, 388, 259, 194, 155. Base + contagion: 729, 365, 243, 183, 146. Complete: 730, 365, 244, 183, 146. * indicates rejection with 95% confidence, ** rejection with 99% confidence, and *** rejection with 99.9% confidence.

Model	Base	Base + frailty	Base + contagion	Complete
Violation Rate	5/22	0/22	2/22	1/22
Unconditional Coverage Test	*** 0.000	0.506	* 0.020	0.221
Independence Test	0.951	1.000	0.107	0.752
Markov Test	*** 0.000	0.802	* 0.018	0.450
Dynamic Quantile Test	* 0.020	1.000	0.318	0.972
Mean Relative Bias	−11.98%	26.44%	−8.13%	−6.32%
Root Mean Squared Relative Bias	17.31%	33.30%	14.37%	13.77%

Table 9: *Out-of-sample forecast accuracy tests.* The table reports p-values of various tests of forecast accuracy and other measures of accuracy. The violation rate indicates the number of times that the realized number of defaults in a given year exceeds the forecast value-at-risk. The null hypothesis of the unconditional coverage test of Kupiec (1995) is that the violation rate does not exceed 1%. The null hypothesis of the independence test of Christoffersen (1998) is that a binary first-order Markov chain for the hit indicators has transition matrix given by the identity matrix. The asymptotic distribution of these tests is chi-squared with one degree of freedom. The Markov test combines the coverage and independence tests. It takes the null of the coverage test and the alternative of the independence test, see Christoffersen (1998). The asymptotic distribution of this test is chi-squared with two degrees of freedom. The null hypothesis of the dynamic quantile test due to Engle & Manganelli (2004) is that there is no correlation between the hit indicators and the number of defaults per year. The asymptotic distribution is chi-squared with three degrees of freedom. Finally, the mean relative bias is the average relative difference of the value-at-risk predicted by each model, compared to the average prediction of all models. The root mean squared relative bias is the standard deviation of the latter. * indicates rejection with 95% confidence, *** indicates rejection with 99% confidence, and *** rejection with 99.9% confidence.



(a) *Number of defaults by Moody's event category.*



(b) *Number of defaults by industry sector according to Moody's 11 Code.*

Figure 2: *Defaults by event category and industry sector.*

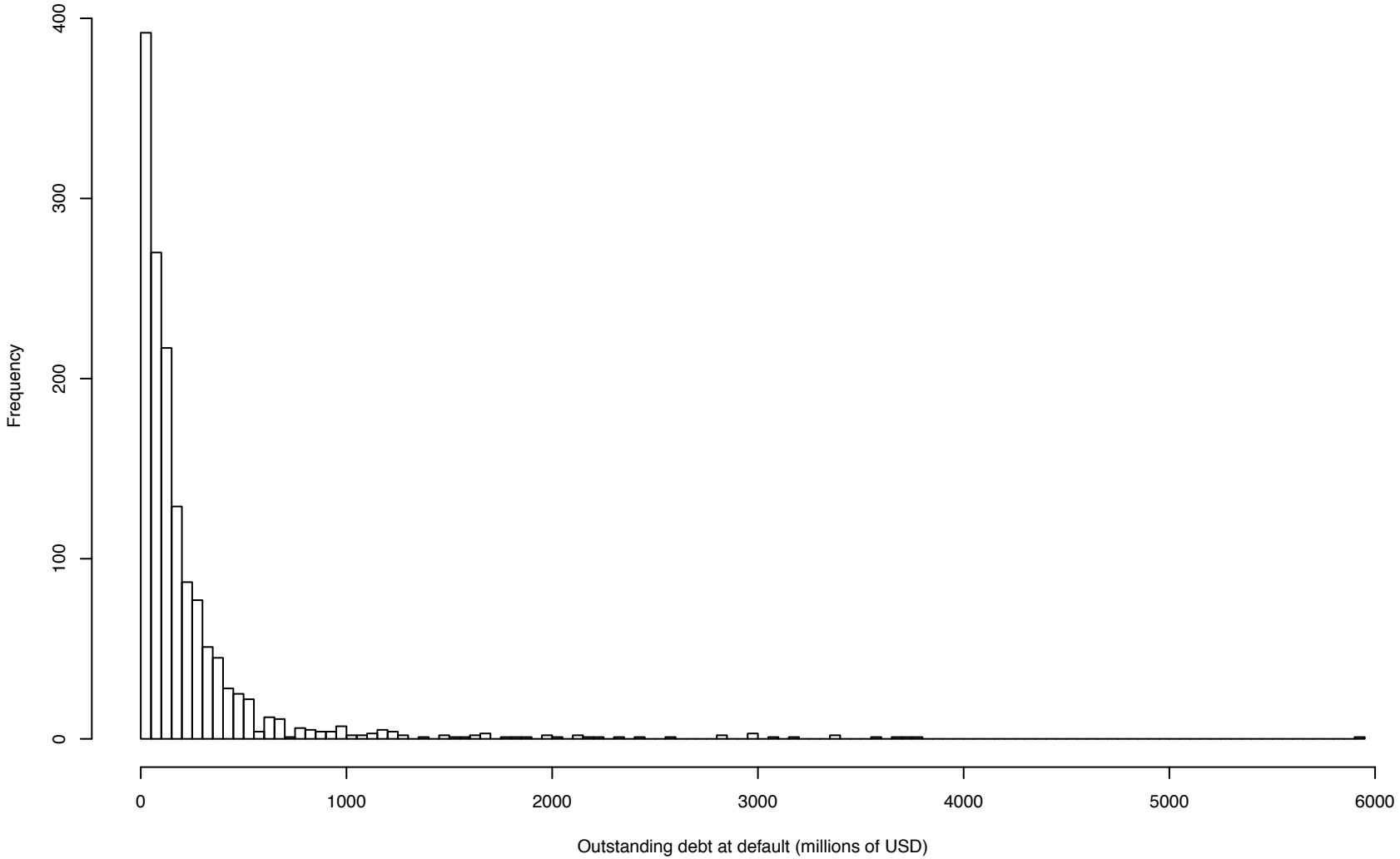


Figure 3: *Distribution of the total amount of debt outstanding at default.* This chart presents a histogram of the total amount of debt outstanding at default for our sample of events.

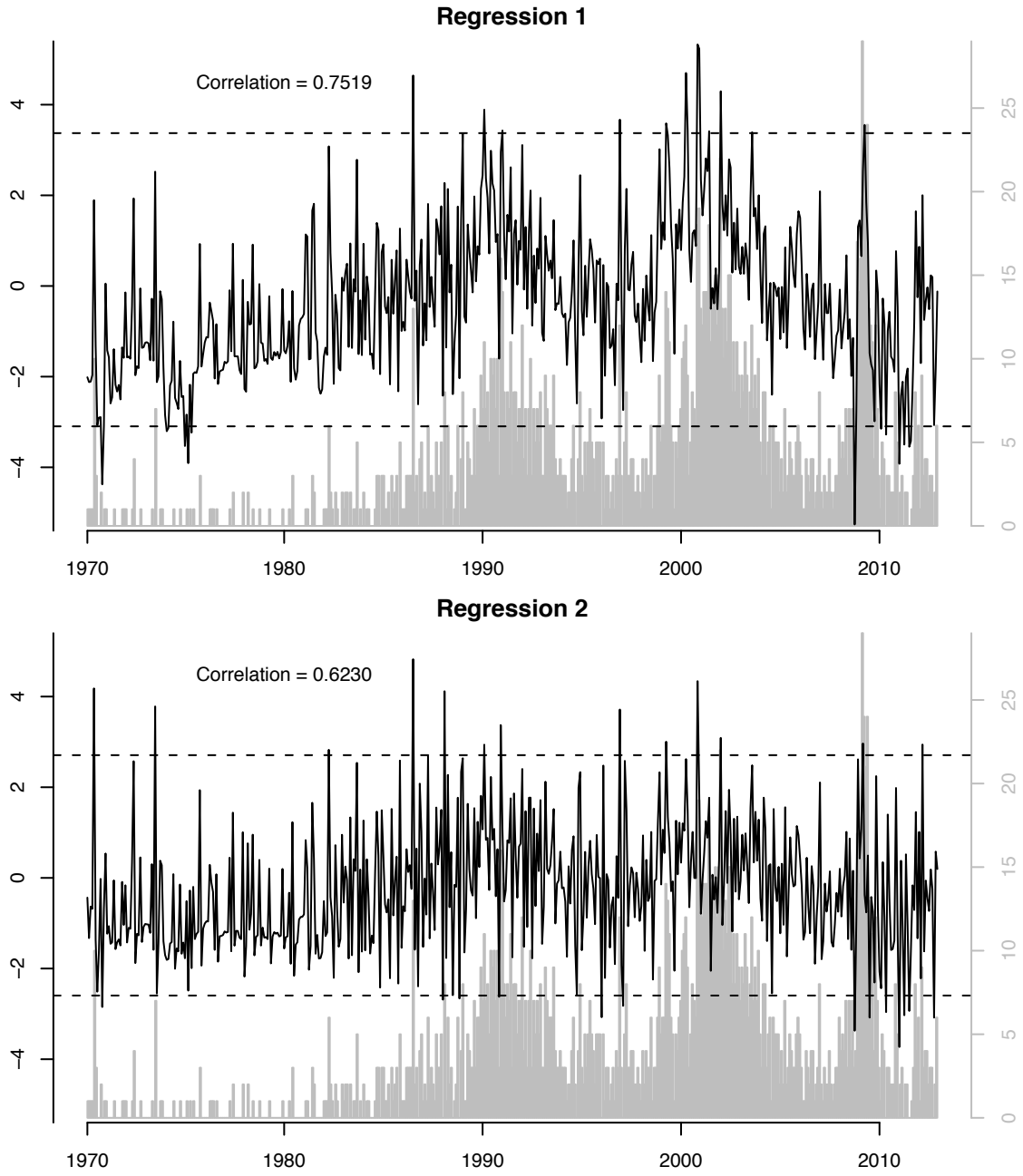


Figure 4: *Regression residuals and monthly defaults.* This chart shows the deviance residuals from the Poisson regressions of monthly defaults (solid line, left scale), along with 95% empirical confidence bands of the residuals (dashed line, left scale), monthly defaults (grey bars, right scale), and the linear correlation coefficient of residuals and monthly defaults. Top panel: Regression 1. Bottom panel: Regression 2. See Table 2 for the definition of the regressions. The deviance residuals of a Poisson regressions are analogous to the residuals of a linear regression.

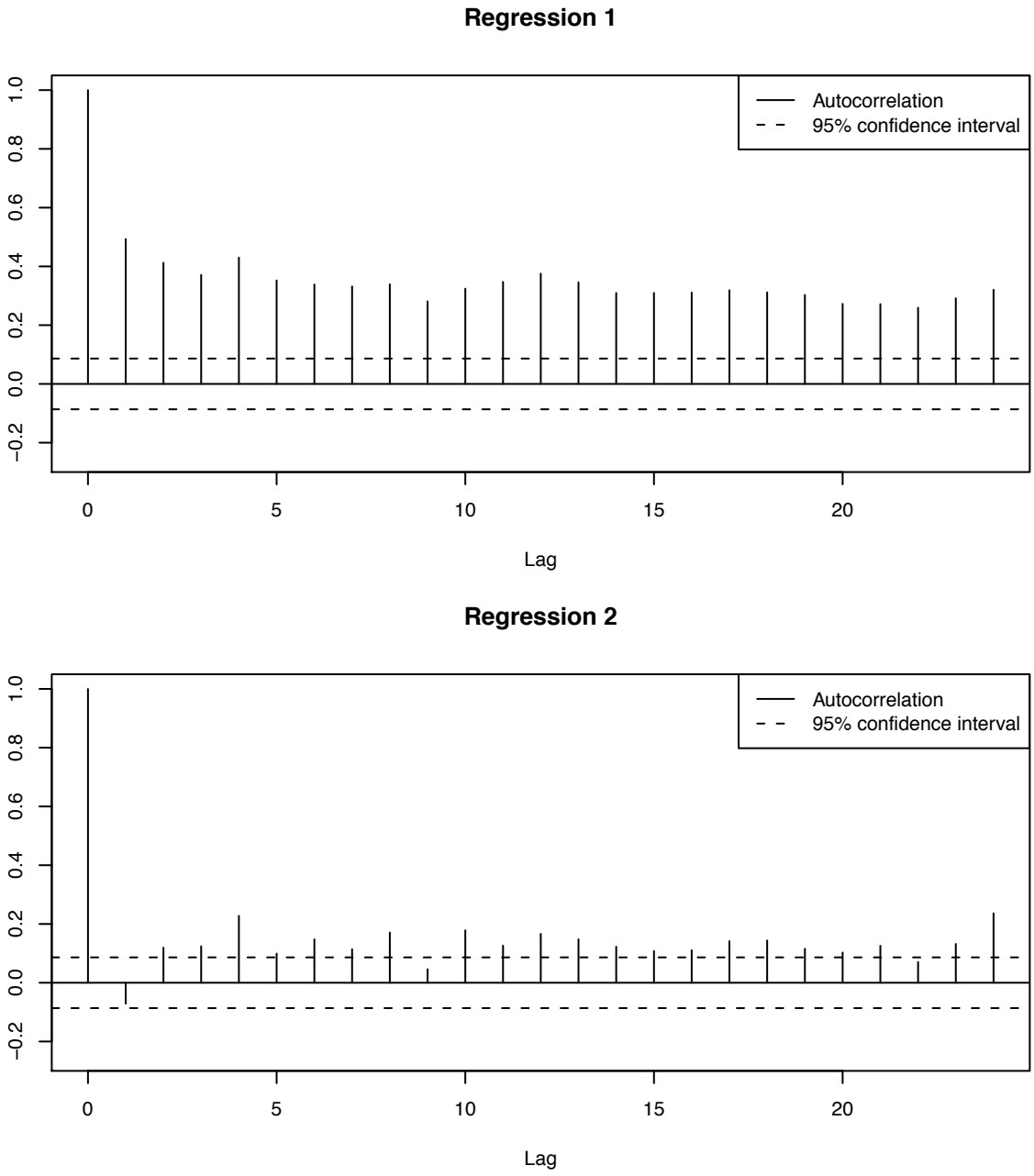


Figure 5: *Autocorrelation of regression residuals.* This chart plots the autocorrelation function up to lag 24 of the regression residuals shown in Figure 4. For lag 0, the autocorrelation is 1. For lag $l > 0$, the autocorrelation is the quotient of the covariance between the current residual and the lag l residual to the error variance S^2 (see Table 2). The dashed lines show the 95% confidence bands for the autocorrelation function assuming independent and identically distributed residuals. Top panel: Regression 1. Bottom panel: Regression 2. See Table 2 for the definition of the regressions.

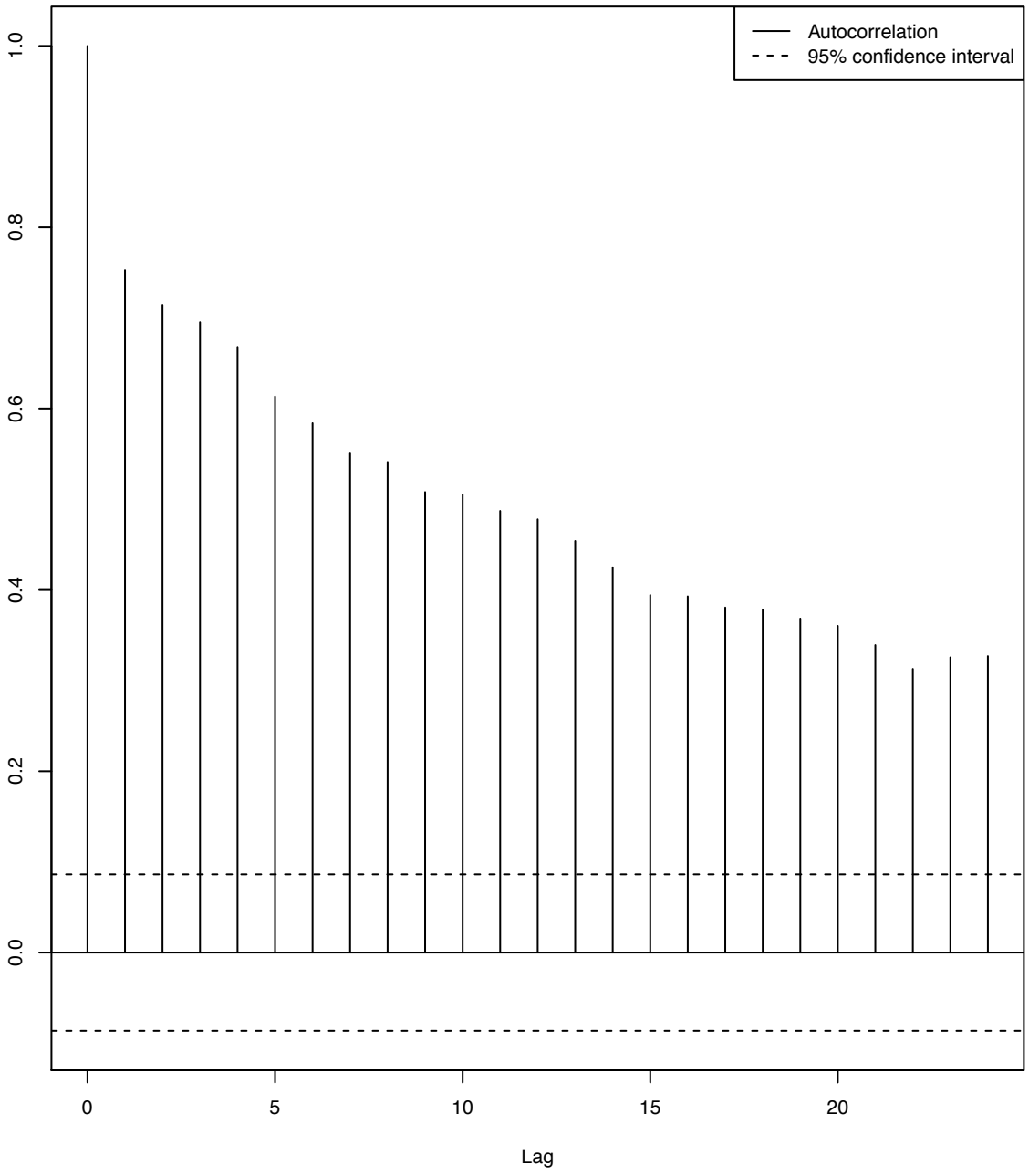


Figure 6: *Autocorrelation of monthly defaults*. This chart shows the autocorrelation function up to lag 24 of the number of defaults per month. The dashed lines indicate 95% confidence bands for the autocorrelation function, obtained under the assumption of independent and identically distributed values.

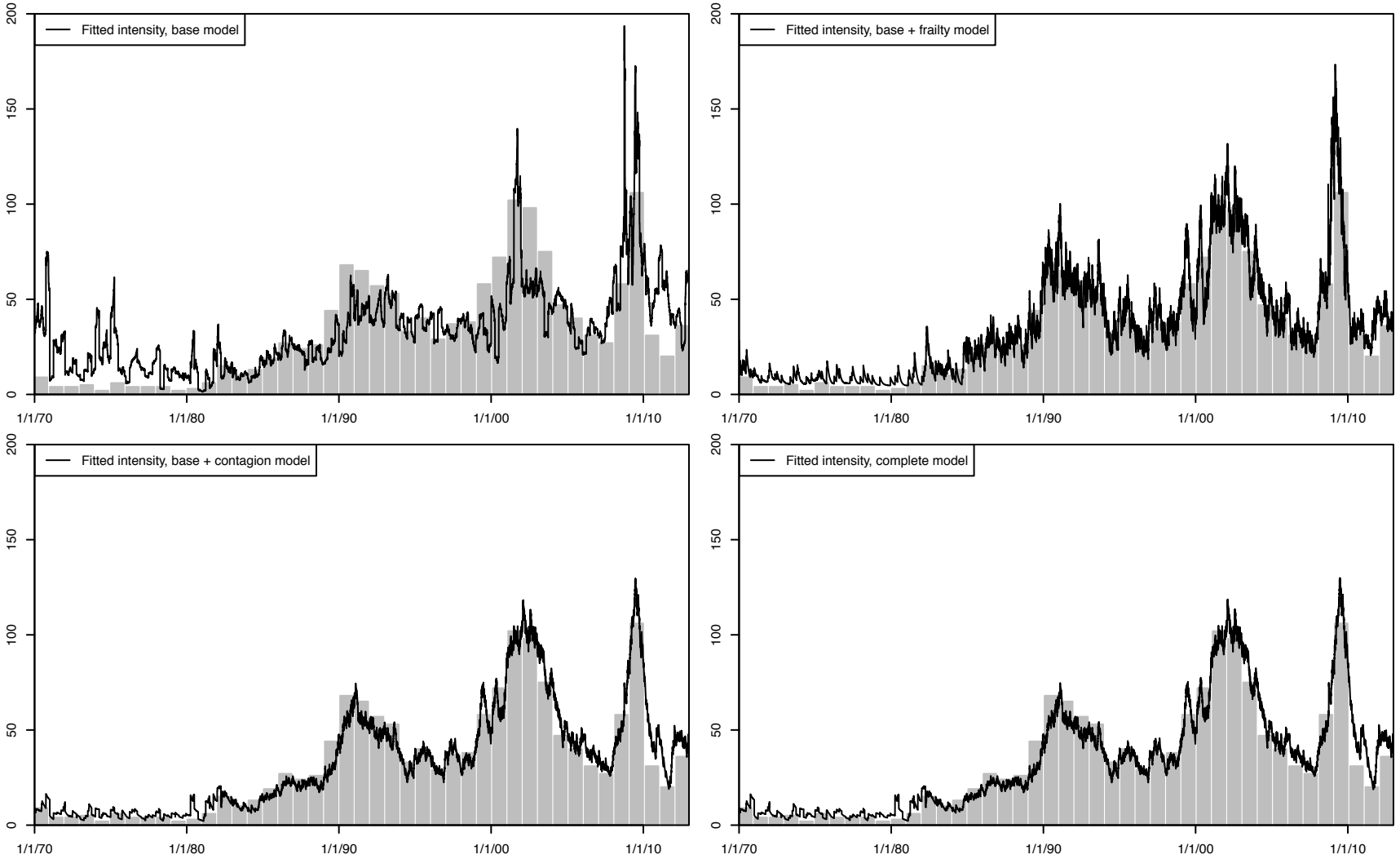


Figure 7: *Fitted intensities.* A figure shows the fitted intensity (measured in defaults per year) implied by the maximum likelihood estimates reported in Table 4. A grey bar represents the realized number of defaults per year.

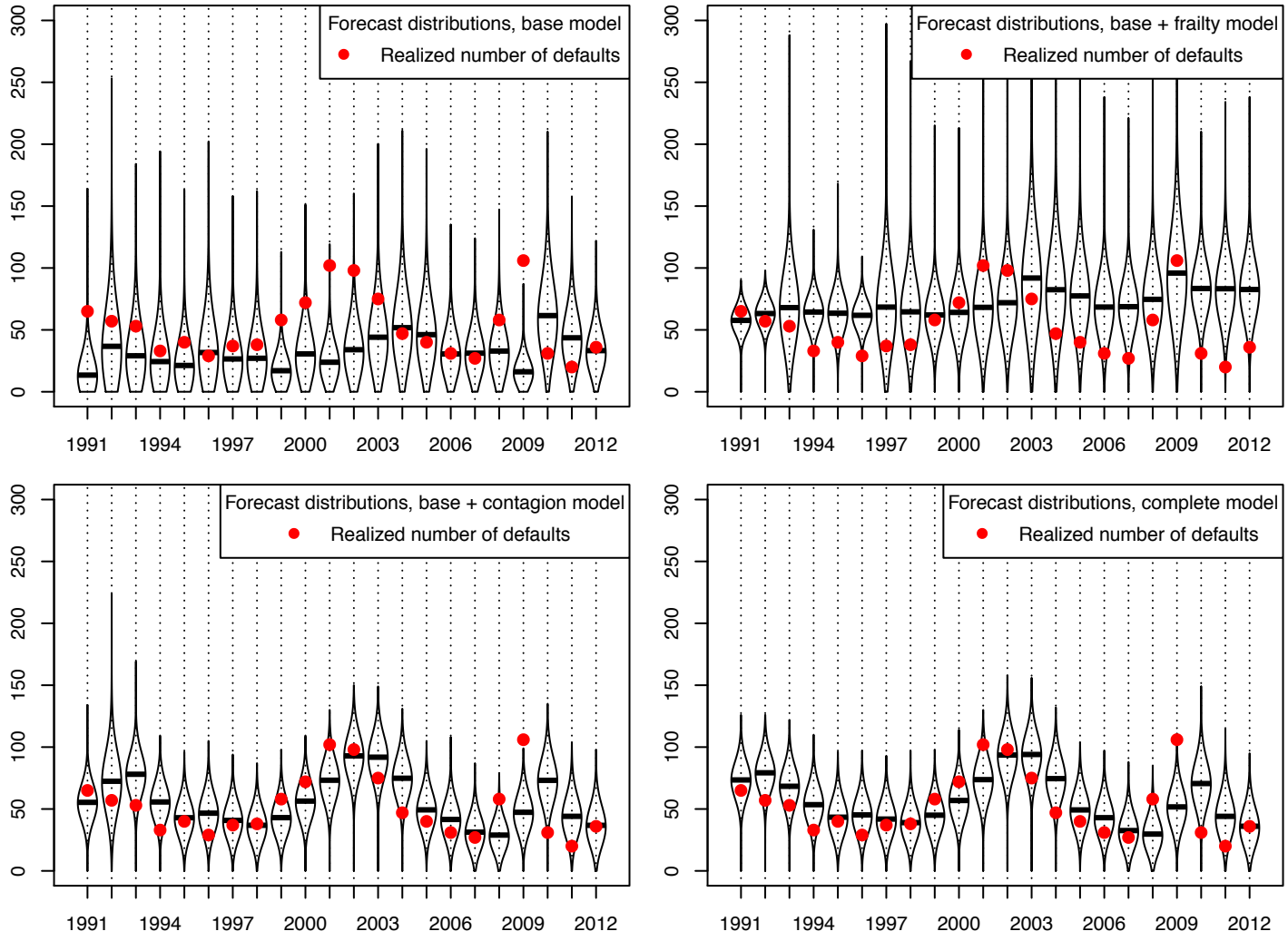


Figure 8: *Out-of-sample forecast distribution of the number of defaults in the year ahead.* The plot shows the Gaussian kernel-smoothed forecast distributions, with a horizontal line indicating the mean. The dot indicates the realized number of defaults. The forecast distributions are computed via Monte Carlo simulation as described in Appendix E.