The Determinants of Operational Losses

Anna Chernobai Philippe Jorion Fan Yu[1]

First Draft: November 12, 2007

Current Version: March 18, 2008

[1] Chernobai is from Syracuse University, Email: annac@syr.edu. Jorion is from the University of California, Irvine, Email: pjorion@uci.edu. Yu is from Michigan State University, Email: yuf@bus.msu.edu. The data used for this study were provided by Algorithmics Inc., a member of the Fitch Group. We are grateful to Lei Zhang and the FactSet staff for technical assistance, and to Richard Cantor for providing access to Moody's corporate default risk service. We thank Thomas Barkley, Phelim Boyle, Long Chen, Steve Dimmock, Ted Fee, Zsuzsanna Fluck, Mike Gallmeyer, Lorenzo Garlappi, Jens Hilscher, Naveen Khanna, Kofi Appiah Okyere, Milena Petrova, Jiaping Qiu, Scott Richardson, Roberto Rigobon, Mark Schroder, Richard Sloan, Franck Taieb, Dragon Tang, Raja Velu, and seminar participants at BGI, Claremont McKenna College, Georgia State University, McMaster University, Michigan State University, M&T Bank, Syracuse University, the 2008 OpRisk USA Conference, and the Second Risk Management Conference at Mont Tremblant for helpful comments. We are especially indebted to Darrell Duffie for extensive discussions on this topic.

The Determinants of Operational Losses

Abstract

We examine the microeconomic and macroeconomic determinants of operational losses in financial institutions. Using 24 years of U.S. public operational loss data from 1980 to 2003, we demonstrate that the firm-specific environment is a key determinant of operational risk; firm-specific characteristics such as size, leverage, volatility, book-to-market, profitability, and the number of employees are all highly significant in our models. In contrast, while there is some evidence that operational losses are more frequent and more severe during economic downturns, overall the macroeconomic environment appears less important. We further test the doubly-stochastic Poisson assumption with respect to the arrivals of operational losses, given the estimated arrival intensities. Despite the traditional view that operational risk is unsystematic, we find evidence of clustering of operational risk events at the industry level in excess of what is predicted by the stochastic frequency estimates.

JEL Classification Codes: G32, G10, G21, G28, C12.

Keywords: Operational risk, Basel II capital accord, doubly stochastic, conditional Poisson.

1 Introduction

We present a comprehensive analysis of the determinants of operational risk, defined by the Basel Committee on Banking Supervision (BCBS) as the risk of loss resulting from inadequate or failed internal processes, people and systems, or from external events. As a distinct risk category, operational risk has received prominent coverage in the recent NBER volume entitled "The risk of financial institutions" (Carey and Stulz (2006) and de Fontnouvelle, Rosengren, and Jordan (2006)). Anecdotally, many corporate failures can be traced to operational losses. Classic examples include Barings Bank and Enron. Operational risk has also been responsible for high-magnitude losses at Daiwa Bank ($1.1 billion) and Allfirst Financial ($700 million), among others. More recently, the $7.2 billion loss due to unauthorized trading at Société Générale has spawned allegations of moral hazard and a lack of internal control in the banking industry (Arnold, Larsen, Hollinger, O'Doherty, and Milne (2008)). The recent multi-billion dollar losses at Credit Suisse and AIG, attributed to "trader error" or "material weakness" related to the valuation of subprime mortgage securities, also suggest that operational losses may be correlated over time and across firms. The prospect of more frequent and severe operational losses during a recession is discomforting because this is precisely at a time when bank capital is eroded by credit losses. This increases the probability of multiple bank failures, which could create systemic risk.

Consequently, regulators are now prodding the banking industry toward better measurement and management of operational risk. For example, the recently finalized Basel II capital accord (BCBS (2006a)) requires an operational risk capital charge. Large internationally active banks currently reserve $2 to 7 billion of capital against operational losses (de Fontnouvelle, DeJesus-Rueff, Jordan, and Rosengren (2006)).[1] Because the advanced measurement approach (AMA) allows banks to use a value at risk (VaR) measure of the loss distribution to compute the regulatory capital amount, the industry has feverishly started to collect data on operational losses. The available historic operational loss data from the industry, when combined with firm characteristics and the business environment, can shed light on the main features of operational risk, including its key determinants.[2]

[1] For example, CitiGroup and JPMorgan Chase reported $8.1 billion and $5.7 billion of operational risk capital in their 2006 annual reports, respectively.

[2] Banks often rely on databases of external losses to supplement their internal loss history and carry out scenario analyses. If they do not "condition" on the business environment and the characteristics of firms suffering these

The distribution of operational losses can be decomposed into a frequency distribution, which describes the arrival of losses, and a severity distribution, which describes the size of losses when they occur. It is commonly assumed that operational losses are independent from each other. As we know from market risk models, the tail of the loss distribution is very much influenced by the dependence structure of loss events. Independent events tend to diversify each other, leading to lower VaR measures than otherwise. To date, very little work has been done on the comovements in operational losses. Indeed, in its recent survey of industry practices, the BCBS notes that the work on the dependence structure of operational risk is "very much in its infancy" (BCBS (2006b, p. 25)). Most banks assume independent events, while some banks allow correlations across business lines or loss event types. BCBS also notes that "dependence structures could occur as a result of business cycles (e.g., economic difficulties that cause an increase in rogue trading and fraud), firm-specific factors (e.g., a new senior manager changes the control environment across a number of business lines), or cross-dependence of large events (e.g., flooding results in widespread looting and increases the number of fraudulent transactions)" (BCBS (2006b, p. 24)). This is consistent with what we document below as the major contributory factors cited for operational loss events: a lack of internal control, managerial action/inaction, and changing market conditions.

We begin our investigation by treating the arrival of operational loss events as a point process, whose intensity is driven by firm-specific and macroeconomic covariates; conditional on the history of these covariates, the arrivals of losses, either within the same firm or across different firms, are independent events. Such processes are called *Cox*, *doubly stochastic*, or *conditional Poisson processes*. This intensity-based framework, in the spirit of traditional factor models for equity returns, is commonly found in the credit risk literature (Jarrow and Turnbull (1995), Lando (1998), and Duffie and Singleton (1999)).[3]

Using a database on publicly reported operational losses on U.S. financial institutions from 1980 to 2003, covering 1,159 losses at 157 firms, we first estimate the arrival intensity of the losses as a function of microeconomic and macroeconomic covariates using maximum likelihood methods. Following BCBS guidelines, our database categorizes the loss events into seven different types,

losses, their estimated loss distribution and the results of their scenario analyses may be misleading.

[3]While it allows for substantial generality as well as ease of estimation, violations of the conditional Poisson assumption can occur when there are missing covariates (Duffie, Eckner, Horel, and Saita (2006)), or if there are contagion effects caused by inter-firm linkages or information spillover (Jarrow and Yu (2001), Collin-Dufresne, Goldstein, and Helwege (2003), Jorion and Zhang (2007), and Yu (2007)).

which we roughly aggregate into internal or external events. With the exception of truly exogenous losses caused by physical disasters such as the terrorist attack of September 11th, 2001, we find that the arrival rate of losses depends strongly on firm-specific characteristics, and particularly so for internal events. Specifically, firms with higher leverage and equity volatility and lower market-to-book ratio have more frequent losses. The dependence on leverage, volatility, and market-to-book, typical firm characteristics that predict the probability of default, is consistent with the hypothesis that financially constrained firms are not devoting enough resources to internal control and oversight. For internal events, especially internal frauds, we also find that the loss frequency depends positively on profitability, suggesting that moral hazard may be a widespread problem in the banking industry. Another interesting firm-specific covariate is the number of employees, which relates to the incidence of internal losses in a concave manner. This is consistent with the human element of internal losses. Interestingly, many of the firm-specific covariates are weaker, but still significant, predictors of external loss events. This suggests that better internal control and oversight can help reduce the incidence of losses that have external causes.

Over our sample period, we notice a similarity between the time-series of the number of U.S. financial defaults and the number of operational loss events. However, the latter appears much less volatile. Consistent with this observation, we find only a weak macroeconomic dependence for the occurrence of operational losses. Out of the many macroeconomic covariates we have included in this analysis, only the GDP growth rate appears to be significantly related to the arrival intensity, indicating that losses are more frequent during recessions.[4] Our analysis of loss severity, given that a loss has occurred, similarly shows that losses are more severe during economic downturns.

Having estimated the arrival intensity of the losses, we conduct a diagnostic check of the conditional Poisson assumption following the procedure outlined in Das, Duffie, Kapadia, and Saita (2007).[5] We find that the conditional Poisson assumption is rejected among certain types of events, such as internal or fraud-related losses. While a detailed investigation into the nature of this violation is beyond the scope of this paper, we find some evidence that the excess clustering of events

[4] In contrast, our results on profitability are most likely driven by the cross-sectional differences in firm-level loss arrivals.

[5] Using the conditional Poisson default arrival process estimated in Duffie, Saita, and Wang (2007), Das, Duffie, Kapadia, and Saita (2007) show that the U.S. default data display too much clustering. Rather than searching for a missing covariate, which they term "frailty," Duffie, Eckner, Horel, and Saita (2006) simply extract it as a latent factor from observed defaults.

can be partially accounted for by missing covariates such as the change in the industrial production index, a post-SOX period dummy, and the industry-wide count of loss events that have been settled in the recent past. These findings suggest that the banking industry has become more vigilant in response to past loss experience as well as regulatory pressure.

By its very nature as a new research area, the empirical literature on operational risk is sparse. The latest empirical studies focus mainly on documenting the size and significance of operational losses. For instance, Cummins, Lewis, and Wei (2006) and Perry and de Fontnouvelle (2005) find a significantly negative equity market reaction to operational loss announcements. de Fontnouvelle, DeJesus-Rueff, Jordan, and Rosengren (2006) show that capital requirements for operational losses can regularly exceed those for market risks at large U.S. banks.[6] There also seems to be a dearth of literature on the topic of common factors in operational losses. A notable exception, Allen and Bali (2007) examine cyclicality in operational risk measures derived from the stock returns of financial institutions, after purging the effect of other sources of risks. This is a top-down definition, however, which does not utilize the distribution of observed operational losses.

Our study is also related to the literature on corporate fraud. For example, the early work by Dechow, Sloan, and Sweeney (1996) relates earnings manipulations investigated by the SEC to board composition and CEO characteristics. More recently, Efendi, Srivastava, and Swanson (2007) find that firms are more likely to misstate financial statements when the CEO has sizable holdings of stock options. Povel, Singh, and Winton (2007) develop a theoretical model where fraud is more likely to occur in good times, because bad firms can more easily masquerade as good firms and attract funding from investors. Relative to this literature, our paper focuses on a much broader sample of operational losses, of which accounting frauds and earnings restatements are only a small subset.[7]

Currently, there is no empirical research that systematically relates the incidence and severity of operational losses to observable firm characteristics and macroeconomic factors. We also do not know whether the dependence structure of operational losses is sufficiently well described by the

[6] Rosenberg and Schuermann (2006) use the operational loss distribution estimated in de Fontnouvelle, DeJesus-Rueff, Jordan, and Rosengren (2006) to construct an aggregate loss distribution across market, credit, and operational risks for a typical large bank.

[7] Among the seven event types defined by the BCBS, accounting frauds and restatements are mostly found within the category of internal fraud (*ET1*) and execution, delivery, and process management (*ET7*). These two event types constitute no more than 23 percent of the loss events in our sample (Table 2).

conditional Poisson framework.[8] Our paper addresses both of these issues. Specifically, we identify several important firm-specific and macroeconomic determinants of operational loss frequency and severity. Moreover, we use diagnostic tools to assess the ability of the conditional Poisson model to describe the arrival of loss events over time and across firms. Our empirical findings are likely to be of interest to bank risk managers and regulators, and useful to the development of new theoretical models for measuring and managing operational risk.

The rest of this paper is organized as follows. Section 2 provides the background for the measurement of operational losses. Section 3 summarizes the operational loss database used in our empirical analysis. Section 4 outlines our empirical methodology and documents the list of firm-specific and macroeconomic variables that might have an influence on the operational loss frequency and magnitude. Section 5 presents our empirical results. Section 6 directly tests the validity of our modeling assumptions. We conclude with Section 7.

2 The Measurement of Operational Risk

Under Pillar I of Basel II, banks are required to adopt a methodology to assess the operational risk capital charge that would serve as a shield against potential future losses given a one-year horizon. In the order of increasing sophistication and risk sensitivity, the spectrum of available approaches proposed by Basel II consist of (i) the Basic Indicator Approach (BIA), (ii) the Standardized Approach (SA), and (iii) the Advanced Measurement Approaches (AMA). Banks are allowed to choose from and must move up along the spectrum as they develop more complex operational risk measurement systems. Under the BIA and SA, the "top-down" approaches, the capital charge is proportional to a fixed percentage of a bank's gross income, pre-determined by the Basel Committee. The AMA are "bottom-up" risk-sensitive approaches in that they are built upon a bank's risk management practices and make use of internal and external historic loss data in order to determine the capital charge. The Loss Distribution Approach (LDA), allowed as part of the AMA, is based on an actuarial-type loss model. Under the LDA, the frequency and the severity of losses are examined separately and then combined into a distribution of total losses over a fixed horizon.

[8] In fact, it is often taken for granted that operational risk is largely idiosyncratic (Daniélsson, Embrechts, Goodhart, Keating, Muennich, Renault, and Shin (2001)).

Define the loss frequency over an interval $[0,\ t]$ as N_t and its probability mass function as

$$\Pr(N_t = n) = p(n),\ n = 0, 1, 2, \ldots \tag{1}$$

At any given point in time t, assuming that the events are independent of each other, this is usually taken as $N_t = N(\lambda t)$, where $N(t)$ denotes a standard Poisson process and $\lambda > 0$ is a constant that is also the expected count per unit time. Under this assumption, we have

$$p(n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}, \quad n = 0, 1, 2, \ldots \tag{2}$$

The loss severity density function for the ith loss X_i, $i = 1, 2, \ldots, N_t$ is

$$f_{X_i}(x) = f_X(x|\, i), \quad i = 1, 2, \ldots, N_t. \tag{3}$$

The total loss over the horizon is then measured as the sum of the losses:

$$L := S_t = \sum_{i=1}^{N_t} X_i. \tag{4}$$

Assuming that the frequency N_t and severity X are independent and that the losses X are *i.i.d.* draws from a continuous distribution $F_X(x)$ with common density $f_X(x)$, the two distributions can be combined into a probability distribution of aggregate loss L through a process known as convolution:

$$G_L(s) := \Pr(L \le s) = \begin{cases} \sum\limits_{n=1}^{\infty} p(n) F_X^{n*}(s), & s > 0, \\ p(0), & s = 0, \end{cases} \tag{5}$$

where $F_X^{n*}(x)$ denotes the n-fold convolution of $F_X(x)$ with itself.

Given this aggregate loss distribution, the value at risk (VaR) over a horizon of one year at the 99.9 percent confidence level can be used to estimate the operational risk capital charge.[9] Operational VaR is defined as the inverse of the aggregate loss distribution function:

$$VaR = G_L^{-1}(0.999). \tag{6}$$

Let this be VaR(L_k) for group k, defined as a business line or event type, or a business line/event type combination. If events and severities are independent within each group and across groups, then it is straightforward to use simulation methods to construct the entire distribution of losses

[9] Basel II recommends using a one-year holding period and a high confidence level for evaluation of the capital charge such as 99.9 percent. See BCBS (2006a) for details.

from the distributions $p(n)_k$ and $f_X(x)_k$. Dependencies, however, can arise in the loss events and severities within each group as well as across groups. If so, the loss distribution could have longer tails than in the independent case. As shown in BCBS (2006b, p. 24), currently a large number of banks simply treat operational losses as independent events, either unconditionally or within the same event type or business line. Only a small number of banks are considering incorporating more complex dependence structures.

In this paper, we consider scenarios in which both the arrival intensity and the severity of operational losses are stochastic and driven by various time-varying firm-specific and macroeconomic covariates. In such cases, the arrival of operational losses follows a *doubly-stochastic*, or *conditional Poisson process* which is also referred to as a *Cox process*. Such counting process, $N'_t = N(\Lambda(t))$, where

$$\Lambda(t) = \int_0^t \lambda(u)\, du, \tag{7}$$

is characterized by the stochastic intensity $\lambda(t) > 0$. When combined with a time-varying loss severity with density $f_{X_t}(x) = f_X(x|t)$, this process yields the aggregate loss process L' of the form

$$L' := S'_t = \sum_{i=1}^{N'_t} X_{t(i)}, \tag{8}$$

where $t(i)$ denotes the arrival time of the ith loss. Therefore, the arrival intensity and the density of the loss severity are the two fundamental building blocks to an operational risk model.

The above definition of the loss arrival and severity processes applies for an individual firm. More importantly, we assume that conditional on the firm-specific and macroeconomic covariates that determine the loss intensities, the arrivals of losses are independent across firms. This assumption ensures that the joint distribution of loss arrivals is completely specified by the stochastic intensities. It allows for the estimation of the intensities using convenient econometric methods based on maximum likelihood. In principle, the functional form of the loss intensity and the density of the loss severity and even the set of covariates can change across event types and business lines.

3 Data Description

In this paper, we analyze the operational loss data from the Algo FIRST database provided by Algorithmics Inc., a member of the Fitch Group. The vendor gathers information on operational losses extracted from public sources. The building of the database began in 1998. The database provides information about the operational losses in the financial and non-financial industries across the world beginning from 1920. It offers a detailed description of each event, including loss amounts, dates of loss occurrence and settlement, company name, geographical location of the event, and event trigger. In addition, the format of the data conforms to the Basel Committee definition of event types and business lines.

Because the data are collected from public sources, they are not fully representative of the entire population of operational losses. Instead, we can interpret them as unexpected events that could not be hidden from the public eye. Because larger losses are more difficult to hide, the sample may be biased toward higher-magnitude events.[10] These events, however, are precisely those that should generate concern because they have the potential to cause major failures.

To more formally describe the potential selection bias, we denote the intensity of loss arrivals as λ and the conditional probability that an event is reported in our database, given that it has occurred, as ϕ. In the literature on point processes, this is called "thinning," which gives rise to a conditional Poisson process with intensity $\lambda\phi$. It is conceivable that ϕ is the outcome of a complicated process that depends on factors both internal and external to the firm. In this case, one could argue that our empirical analysis may be uncovering the factors determining the *reporting* of an event rather than those responsible for its *occurrence*. To address this issue, we have looked through an item in the FIRST database that deals with the source of each recorded event. In an overwhelming majority of the cases, the sources of the loss announcements are not the firms themselves, but third parties such as SEC press reports, the NASD, court decisions, and affected customers. This suggests that the firms actually have little choice in deciding whether an event is reported to the public, mitigating concerns over selective disclosure.[11]

We follow the Basel II classification of risk events according to event type.[12] The seven event types are as follows:

[10] Nevertheless, we note that there is no minimum recording threshold for losses collected by this vendor. In fact, many of the losses are quite small in magnitude.

[11] Of course, one can never rule out the existence of sample selection bias completely. For instance, the frequency of losses is likely to be influenced by the budget of the regulators. However, it would be difficult for such factors to fully account for our estimated loss arrival intensity, which has a strong dependence on firm-specific covariates that differs across event types.

[12] The Basel II accord also classifies the loss events according to eight business lines, such as corporate finance, trading and sales, retailed banking, commercial banking, payment and settlement, agency services and custody, asset management, and retail brokerage. However, in this study we only focus on the differences among the event types.

1. Internal Fraud (*ET1*): Events intended to defraud, misappropriate property, or circumvent regulations or company policy, involving at least one internal party, categorized into unauthorized activity and internal theft and fraud.

2. External Fraud (*ET2*): Events intended to defraud, misappropriate property, or circumvent the law, by a third party, categorized into theft, fraud, and breach of system security.

3. Employment Practices and Workplace Safety (*ET3*): Acts inconsistent with employment, health, or safety laws or agreements, categorized into employee relations, safety of the environment, and diversity and discrimination.

4. Clients, Products, and Business Practices (*ET4*): Events due to failures to comply with a professional obligation to clients, or arising from the nature or design of a product, and include disclosure and fiduciary, improper business and market practices, product flaws, and advisory activities.

5. Damage to Physical Assets (*ET5*): Events leading to loss or damage to physical assets from natural disasters or other events such as terrorism.

6. Business Disruption and System Failures (*ET6*): Events causing disruption of business or system failures.

7. Execution, Delivery, and Process Management (*ET7*): Events due to failed transaction processing or process management that occur from relations with trade counterparties and vendors, classified into categories such as transaction execution and maintenance, customer intake and documentation and account management.

8. *Other*: Events that cannot be classified into any of the above seven event type categories.

3.1 Sample Selection

For the empirical analysis relating the distribution of operational losses to macroeconomic and firm-specific covariates, we only consider data belonging to the U.S. banking industry. This ensures some homogeneity in the sample, which reflects common processes and policies. We apply several further filters to the raw data and remove observations for which the date of loss occurrence is unknown.

After preliminary filtering, the initial data sample covers events beginning between 1980 and 2005. Each risk event has two associated dates: event starting date ("start date") that refers to the date on which a particular event originated and event ending date ("end date") on which the event ended (for example, the date on which it was revealed to the firm's authorities). Figure 1 provides an illustration of the distribution of the duration (in months) of recorded events. The event duration distribution is rightly skewed. Table 1 reports the detailed descriptive statistics for the event duration for every event type as well as the aggregate data. The table indicates that approximately half of the reported operational risk events last for no more than 20 months. Overall, three quarters of all events last under 4 years. The longest-lasting events rarely exceed 15 years.

Since many losses may take months or even years to materialize, it is likely that many events are currently taking place but have not yet been discovered. This means that the last several years of the database may be under-populated. Specifically, we extract the data from the database in early 2007. According to the distribution of event duration, by 2007, the database includes less than half of the events that originated in 2005. To minimize the effect of right censoring without losing too much information, we further limit the sample to events originating between the beginning of 1980 and the end of 2003. This leaves us with 1,419 loss events in the sample. This time window guarantees coverage of approximately 75% of events for 2003, and over 80% and 90% of events for 2002 and 2001, respectively, leaving out only a negligible fraction of events originating prior to 2001.

Finally, we exclude all firms that are not publicly traded and keep only those firms whose information is available in COMPUSTAT and CRSP databases. The finalized sample represents 1,159 loss events at 157 financial firms.

3.2 Time-Series Behavior of Operational Losses

We notice that the frequency and the severity of operational risk exhibit a highly non-uniform nature across event types and time. First, Panel A of Figure 2 illustrates the annual frequency of loss events during the 1980-2003 time period and Panel A of Table 2 breaks down the annual frequencies by event type. The most frequent losses appear within the clients-related category, *ET4* (47.4%), followed by internal fraud, *ET1* (15.5%). The frequency of operational risk events has been on the increase since 1980, but experienced a sharp decline after 2001.[13] This pattern is remarkably similar to that of the number of banking industry defaults in the same period (Panel C of Figure 2), suggesting that there is a high degree of dependence between credit risk and operational risk.[14] Although the overall patterns are similar, we further note that the number of operational losses appears less volatile than the number of defaults, suggesting that the business cycle dependence of operational risk may be slightly weaker. Furthermore, an inspection of Table 2 shows that all types of operational loss events do not follow the same time-series pattern. This is consistent with different types of events being triggered by different sources of risk.

Second, Panel B of Figure 2 shows that the severity of losses follows a more erratic pattern than the frequency of losses, experiencing peaks and troughs throughout the sample period. Panel B of Table 2 documents the annually aggregated loss amounts by event type. The most frequent losses (*ET4*) also account for the largest total amount of losses (52.9%). In contrast, fewer than 2% of all events—losses belonging to the *ET5* category (Damage to Physical Assets)—account for over one tenth of the total loss amount (10.6%). The bottom of Panel B reports the average loss amount per event. It shows that events in *ET5* on average cost $295 million. Panel A shows that 16 of the 22 events for this category occurred during 2001.

In examining the frequency data, our primary focus is to identify the underlying factors that cause operational risk events to take place. We have previously discussed two types of dates provided for our sample. For frequency data, the relevant event date would be the "start date" of occurrence. For some events, the event start dates are unknown and therefore recorded as January 1 of the corresponding year; to correct for this artificial bunching of dates, we include a January dummy variable in all of our econometric frequency models. Such a phenomenon does not come as a surprise: in the case of internal fraud, for example, it is often impossible to determine the precise date at which the first incident occurred, hence the fraudulent activity is recorded as having

[13] It is unlikely that this decline is because of the extended duration issue described earlier (for example, even increasing the 2003 loss count by 25 percent would still lead to a declining pattern in Panel A of Figure 2). One reason for such a decline could be the release of the Basel II capital accord, in particular the early amendments to the accord calling for regulatory capital for operational risk (BCBS (1998, 1999)). This development likely caused banks to increase their effort to reduce operational losses. The period of declining operational losses also coincides with the passage of the Sarbanes-Oxley Act, which seeks to improve the quality of financial reporting and internal control, particularly with respect to corporate fraud. On the other hand, banks were already subject to the FDICIA (Federal Deposit Insurance Corporation Improvement Act) rules during this period, many of which are similar to the Sarbanes-Oxley requirements.

[14] Our data on financial industry defaults come from Moody's default risk service, and include 173 defaults from 1980 to 2003.

started on January 1 of the year in which the activity began. On the other hand, in analyzing the magnitudes of loss events, we consider the "end date" of the event. Because the total loss amount can only be known (or estimated) on or after the event completion date, the last month of the duration interval is used as the timing of loss amount in our severity models.[15] This allows us to analyze the determinants of the total loss amount, for which we include the event duration as an explanatory covariate.

Table 3 presents the descriptive statistics for the monthly aggregated operational loss data classified by event type. For the severity distribution, the mean values significantly in excess of the median values suggest heavily right-skewed loss distributions, confirming the evidence of heavy tails in operational loss data documented in other studies such as Chavez-Demoulin, Embrechts, and Nešlehová (2006). The presence of high-scale events is evident from the extreme values for the maximum losses: for *ET4* (Clients, Products, and Business Practices), the largest monthly loss exceeds $5 billion. During months with non-negative loss counts, *ET5* (Damage to Physical Assets) and other unclassified events account for the highest mean losses—$1.6 billion and $0.9 billion per month, respectively.

4 Methodology

4.1 Econometric Framework

We work with data organized as a cross-sectional time-series panel. The panel represents individual firms and is unbalanced due to unequal lengths of time the firms are represented within the sample. For all categories of losses, we first estimate regressions of the number of losses per month and the average loss amount given loss occurrence, where the independent variables in the frequency model and the severity model can include various contemporaneous firm-specific financial indicators and macroeconomic covariates measured at annual or quarterly intervals. As the loss frequencies vary by each event type, the sample lengths in the severity models are not uniform. Because the sample size is small for some categories, parsimony is essential; we only keep variables that are significant in the regression analysis. The firms' balance sheet data are obtained from COMPUSTAT, and the

[15] Strictly speaking, the loss amount only becomes known after the settlement date, which occurs after the event end date. However, our focus is not on the market reaction to operational losses, but characteristics (frequency and severity) and determinants of the losses as they occur. In other words, our perspective is that of a financial firm trying to assess its operational risk using a database of historical loss events.

market values of equity are obtained from CRSP.

We use the maximum likelihood estimator (MLE) for the frequency models based on the conditional Poisson count of events—an assumption that we will test later.[16] Because the frequency model provides as an output the predicted (or expected) number of loss events each month, the estimation of the model is equivalent to estimating the time-series of the stochastic intensity $\lambda(t)$ of the count process.

The interpretation of this estimated intensity merits additional discussion. Because our dataset contains firms that report at least one operational loss event during the sample period, the estimated intensity of losses cannot be interpreted as the intensity of losses for a firm randomly selected from the population of all firms. We will revisit this issue in Section 5.3.

4.2 Explanatory Variables

Before motivating our choice of the independent variables, we examine all events in our data sample. Table 4 provides a breakdown of the seven event types by event details and identifies the key contributing factors for the events in our sample. Contributory factors such as a lack of internal control, employee misdeeds, and management actions are quoted for many event types. These factors point to the firm-specific nature of operational risk. This is particularly true for *ET1*, *ET3*, *ET6*, and *ET7*. At the same time, market conditions are often cited as a leading contributor to event types *ET2*, *ET4*, *ET5*, and *Other*. Therefore, we hypothesize that operational risk is driven by two types of economic forces: microeconomic forces that depend on a particular firm's characteristics including size, capital structure, and profitability, and macroeconomic forces that determine the healthiness of the U.S. banking sector. We use these contributory factors to help construct our econometric models.

For the purpose of better delineating our results in the subsequent analysis, we group the seven event types into several broad categories. For example, *ET1*, *ET3*, *ET6*, and *ET7* are grouped together as internal events or Model 5, while *ET2*, *ET4*, *ET5*, and *Other* are classified as external events or Model 6. We also designate specific categories for internal fraud or fraud related losses (Models 1 and 2), because these types of losses may be particularly sensitive to the tightness of internal control. Similarly, *ET5* is designated as a separate category because losses caused by

[16] Because operational loss count data are highly right-skewed, applying the OLS estimator is inappropriate and would result in biased, inefficient, and inconsistent estimates.

natural disasters are unlikely to be effectively controlled by internal oversight. We also include as Model 8 the aggregation of all event types. These broad categories are summarized in Table 5.[17]

Among the included firm-specific variables, we use the market value of equity (MVE) to control for firm size in all models. Overall, the number of transactions and the average dollar volume of each transaction are greater at larger firms. Therefore, larger firms are likely to experience a higher frequency and severity of operational losses. At the same time, the construction of our database favors the reporting of larger, more significant, operational loss events. This may also give rise to a mechanical link between loss frequency and firm size.

Firm leverage, market-to-book ratio, and the volatility of equity returns are covariates common to the credit risk literature. Firm leverage, for example, is positively correlated with the risk of financial distress. Because imposing control and enforcing oversight is costly, firms with limited financial resources would have more frequent operational losses due to the lack of oversight. Meanwhile, employees of financially constrained firms may feel the pressure to resort to fraud or other improper business practices. Thus, firm leverage is expected to be positively associated with the frequency of operational risk events. Market-to-book ratio is commonly interpreted as a proxy for default risk, with lower values signalling distress (Fama and French (1992)). For the same reason, we expect higher market-to-book ratios to be negatively associated with the frequency of operational losses. The volatility of stock returns is calculated as the trailing standard deviation of monthly stock returns for the previous year. Firms with higher uncertainty surrounding their equity returns signal a riskier environment, thus are associated with higher operational risk. All in all, we predict a significant association between operational risk frequency and credit risk-related variables.

Return on assets (ROA) is measured as the ratio of earnings before interest, taxes, and depreciation (EBITDA) divided by the book value of total assets, and is a common measure of profitability. By the preceding argument, more profitable firms should experience fewer operational losses. However, because of moral hazard within the firm, profitability can be positively correlated with the incidence of operational losses that are "internal" to the firm. For instance, employees might be tempted to embezzle funds given "money left on the table." Alternatively, upper management can

[17] Our sub-categorization of the events is along the same line as Jarrow (2007), who suggests that events be divided broadly into two categories—one related to the operating technology of the firm, such as failed systems or transactions, and the other related to agency costs and incentives, such as internal frauds and mis-management.

look the other way at failures of internal control when profitability is high, or needs to be juiced up.

The number of employees in our models potentially proxies for the managerial power that each employee has at the firm. After controlling for firm size, the *ceteris paribus* effect of the number of employees represents the level of responsibility each employee has within the firm; when there are few employees, each person is entrusted with a high degree of control over the firm's assets and operations. This seems to suggest that operational risk would decrease with the number of employees. On the other hand, the scope for human errors or misdeeds is greater when the number of employees is larger. It is therefore possible to observe a nonlinear relation between the number of employees and the frequency of operational losses, especially those of the internal event types that involve human interaction within the firm.

We also relate the loss frequency to macroeconomic events, following the literature on predicting default probabilities from macroeconomic data. For example, Helwege and Kleinman (1997) model one-year default rates over 1981-1994 using a number of variables, including GDP growth. Duffie, Saita, and Wang (2007) predict default intensities over 1980-2004 using the 3-month Treasury rate and the 1-year return on the S&P 500 index. Other studies use changes in the consumer confidence index or the industrial production growth rate. Industry models such as CreditPortfolio View regress default rates on the GDP growth and unemployment rates. In addition, we add variables related to market risk and credit risk, as operational losses can often be triggered by market or credit events. Credit risk is measured by the Baa-Aaa spread. Market risk is measured by the annualized monthly volatility of the past 3 years of the S&P 500 index returns. This period is chosen so as to include the median duration for loss events, which is approximately 20 months.

With respect to the relation between the operational loss frequency and macroeconomic covariates, our expectations are mixed. On the one hand, we expect to see a higher incidence of fraud during good times (Povel, Singh, and Winton (2007)). On the other hand, firms have less free cash flows to allocate to improvements in internal controls during an economic downturn, which can contribute to more operational losses of all types. Also, a higher unemployment rate may result in increased instances of external fraud.

Table 6 summarizes the definitions of the dependent and independent variables that will be used in our empirical study.

5 Frequency and Severity Analysis

5.1 Frequency Models

We use a maximum likelihood estimator for the frequency models. The dependent variable is the monthly aggregated event count for each firm. We include firm dummy variables in the initial modeling in order to capture possible firm-level fixed effects. However, the coefficients are mostly insignificant. We therefore drop firm-level fixed effects from all models.

Table 7 summarizes the results from the regression analysis. We first focus on the interpretation of the firm-specific covariates. As predicted, firm size coefficients are positive and highly significant—firms with a greater market value of equity experience a larger number of operational losses. These coefficients are also quite stable across different categories of loss events. Again, this may indicate that larger firms are likely to have a higher number of losses, regardless of the source or nature of the events. Alternatively, this result may partly be caused by smaller firms being overlooked by public scrutiny.[18]

We have hypothesized that variables proven to predict financial distress would be related to operational risk frequency. Our model supports this hypothesis: for all event type categories except Model 7, the coefficient on firm leverage appears positive and statistically significant at the 1% level. Furthermore, the magnitude of this coefficient appears to be much bigger for internal events than for external events. For example, for internal frauds (Model 1A), the coefficient on leverage is 7.03, while for all events excluding internal frauds (Model 3A), it is only 3.24. Similarly, the leverage coefficient is 6.22 for internal events (Model 5A) and only 2.56 for external events (Model 6A). This suggests that firm characteristics play a more important role in the determination of internal operational losses. On the other hand, the fact that firm leverage remains significant even for external events gives us hope that better internal control can help mitigate the effect of losses that originate outside the firm.

Our results on the market-to-book ratio and equity volatility are similar to those associated with firm leverage. Namely, the frequency of losses is strongly positively (negatively) related to equity

[18] To control for a potentially nonlinear effect of firm size, we have also included the square of the firm size variable. Our results indicate that the frequency of losses initially declines with the firm size for a very small range of values, and then starts to increase monotonically. Meanwhile, our results pertaining to the other explanatory variables remain unchanged. We have also considered other firm size variables, such as total assets and net income. The market value of equity (MVE) is the most significant among the size-related variables. Hence we use it uniformly in all models.

volatility (market-to-book ratio), and the magnitude of the coefficients are larger for internal events, in particular for internal frauds. In this case, the market-to-book ratio does lose its significance for external events (Model 6).

The profitability measure (proxied by return on assets) appears to have a strongly positive link to operational loss frequency only for internal events.[19] The coefficient on ROA is significant at the 1% level for internal events (Model 5), not significant among the full sample (Model 8), and is in fact negative and insignificant for external events (Model 6). It is clear, however, that the significance of this measure derives solely from its explanatory power for internal fraud events. When internal frauds are excluded, the ROA coefficient becomes insignificant (Model 3). This evidence is consistent with ROA proxying for moral hazard in the banking industry.

Our estimation shows that the loss frequency is a concave function of the number of employees. As the number of employees increases, the operational loss frequency first increases but then falls. While this pattern is observed for all models, it is statistically significant only for internal events (e.g., Model 5 vs. Model 6). This is consistent with the human element of internal losses. When the number of employees is small, the probability of an operational loss event increases linearly with each additional employee. However, as the number of employees becomes large, each person plays only a marginal role in the operation of the firm, thus mitigating this effect.[20]

With respect to macroeconomic covariates, we obtain generally mixed results. A Wald test of the hypothesis that the coefficients on the macroeconomic covariates are jointly zero cannot be rejected for many models (event types). Many of the coefficients are statistically weak and switch signs across different models. The only macroeconomic covariate that appears to have any degree of stability and significance is GDPgr (GDP growth rate). Its negative coefficient suggests that operational losses are more frequent during economic downturns. In fact, this is the only significant macroeconomic covariate for losses aggregated across all event types (Model 8).[21]

[19] We have also experimented with other profitability measures, such as equity returns, earnings per share (EPS), and a long-term profitability measure of Tobin's Q. The strongest results are produced when we use the ROA measure defined as the ratio between EBITDA and total assets.

[20] To address the concern for multicollinearity between the number of employees and firm size, we have repeated the estimation using the number of employees scaled by the total assets of the firm. This variable has a correlation of only -0.23 with our firm size variable, and yet the nonliear effect remains qualitatively the same.

[21] In results not reported here, we have also included in our frequency models a dummy variable that equals one for the period of August 2002 to December 2003, corresponding to the part of our sample period when Sarbanes-Oxley was in effect. The coefficient for this variable is negative and significant, while the rest of our results on firm-specific and macroeconomic covariates are largely unchanged.

Incidentally, we find that the frequency of loss event type *ET5* is strongly related to all of the macroeconomic covariates. However, Table 3 shows that there are only six months during which losses of event type *ET5* occur, and Table 2 suggests that these events are likely all clustered around September 2001. For this particular case, it is possible that the inference about macroeconomic covariates (and possibly the ROA measure) reflects other events that cause losses and an economic slowdown.

To summarize the effects of firm-specific covariates, we find that the operational loss frequency is closely related to firm characteristics that proxy for financial distress, with the magnitude of the coefficients being larger for internal events than for other event types. Furthermore, we observe the same tendency among internal fraud events. This implies that better internal control is more effective at preventing internal events such as corporate fraud than external events. It is also notable that even for event types that are labeled external, many firm-specific covariates have significant explanatory power for the frequency of losses. This suggests that, if a firm is vigilant and has good internal control mechanisms, it can stop externally inflicted losses before they become damaging. Overall, our frequency analysis confirms the view that operational risk is largely firm-specific; the market-level economic environment has a lesser effect on the frequency of operational losses.

5.2 Severity Models

For severity models, we use the ordinary least squares (OLS) estimator. For each firm, we divide the monthly aggregated loss amount (if non-zero) by the monthly aggregated loss count to obtain an average loss amount per event (usually there is only one event during the month of occurrence), and then scale it by the total assets of the firm and use its logarithm as the dependent variable in the severity analysis. A dependent variable of this form represents average relative operational loss in logarithmic form, which is close to normally distributed.

Table 8 summarizes the regression results.[22] As in the frequency models, we find a significant dependence of the relative operational loss on firm-level covariates. For example, larger firms have smaller relative losses. This is to be expected. Larger firms tend to have better controls, so that losses do not grow in proportion to assets. On the other hand, the relative loss is negatively related to leverage, volatility, and the book-to-market ratio. These findings are more difficult to interpret.

[22] In our preliminary analysis, we include event duration as a covariate. However, it does not have a significant impact on the loss severity, and is therefore dropped from the subsequent analysis.

It is possible that better internal control (as proxied by lower leverage, volatility, and book-to-market) is only effective in eliminating small routine losses, leaving out large unanticipated ones. It is also possible that these results are driven by data issues such as the relatively small sample size of realized loss events.

Similar to the frequency analysis, we find that the only macroeconomic covariate consistently related to the severity of operational losses is the GDP growth rate. The estimated relation suggests that operational losses are more severe during economic downturns.

5.3 Results on All Financial Institutions

As mentioned in Section 4.1, our analysis of the determinants of operational losses makes use of financial institutions in the Algo FIRST database. Therefore, we are effectively focusing on firms in the U.S. banking industry with at least one reported operational loss event in the sample period of 1980-2003. While there is substantial cross-sectional variation in the number of loss events within this sample, an interesting issue remains as to whether the firm-specific and macroeconomic covariates used in our analysis are able to differentiate firms that have reported operational losses from those that do not. To address this issue, we extract all firms in the intersection of CRSP and Compustat with four-digit SIC codes beginning with a "6," as long as they have at least one year of coverage in both databases. This gives us a universe of 4,481 U.S. financial institutions with over 157,000 firm-months of observations, about nine times the original sample size in Table 7. If one of these firms does not have coverage in the Algo FIRST database, we assume that its number of operational loss events is zero in the sample period. We then repeat the operational loss frequency analysis using this substantially enlarged sample of financial institutions.

Table 9 reports the results of this estimation. If our conditional Poisson model has no ability in differentiating firms that have losses from those that do not, then this expansion of the sample would amount to a "random dilution" that would have caused the statistical significance of the estimated coefficients to decrease. On the other hand, if our intensity specification is successful at accounting for the differences between firms with losses and those without, then it is possible for the statistical significance of the coefficients to become even stronger in the expanded sample. Table 9 supports the latter interpretation. For example, the pseudo R^2 of the Poisson regression is uniformly larger than those in Table 7. In addition, most of the originally significant covariates, such as firm size, the number of employees, firm leverage, equity volatility, market-to-book, and GDP growth rate, remain significant with the same sign as in Table 7.

There are, however, notable changes in two of the coefficients. First, the magnitude of the coefficient on leverage has decreased considerably (yet remaining highly significant), implying that the economic significance of the effect is minimal. Therefore, while firm leverage is strongly linked with the number of loss events among firms with at least one loss, it plays a less prominent role in the unconditioned sample. Second, recall that the ROA variable explains internal operational loss events in a positive way in Table 7, suggesting a potential moral hazard effect. However, in Table 9 the coefficient is negative or insignificant, implying that more profitable firms are less likely to experience an operational loss. These subtle differences highlight the importance of separately dissecting the conditional and unconditional loss frequency distributions. Overall, however, we conclude that these covariates are able to distinguish firms with n losses from those with $n+1$ losses, whether $n \geq 1$ or $n = 0$.[23]

6 Goodness-of-Fit Tests

In this section, we carry out goodness-of-fit tests for the conditional Poisson assumption that underlies our modeling framework.

Figure 3 illustrates the fitted econometric model for the operational risk frequency of new events. Specifically, we aggregate the model-predicted intensity function across all surviving firms in our sample over time. The fitted model captures the time-series of the actual number of operational losses well. Note, however, that the actual number of losses substantially exceeds the predicted value around 1989 and 2001. We therefore conduct more formal statistical tests to evaluate the ability of our model to capture the time-series behavior of the arrival of loss events. Specifically, this is a direct test of the conditional Poisson assumption we have made throughout this study. Such an assumption implies that operational risk events occur independently from each other given the estimated intensities, and that the distribution of the inter-arrival times is exponential with the arrival rate identical to the stochastic intensity of the Poisson count process. The evidence

[23] Because the loss severity analysis involves only those firms that have experienced a loss, there is no need to repeat it for the larger sample. Moreover, because the aggregated intensity remains very close to what we would compute based on the estimated intensity of Table 7 (shown later in Figure 3), the goodness-of-fit tests conducted in Section 6 would yield similar results as well. Hence we do not repeat these additional tests for the expanded sample.

in Section 5 casts doubt on the homogeneity of the intensity rate over time. In light of the effect of microeconomic and macroeconomic conditions on loss frequencies, a conditional Poisson model appears to be a reasonable choice.

Because we are dealing with monthly aggregated data rather than daily data, carrying out tests for the inter-arrival times would not be possible. Therefore, we perform the χ^2 test, also called the Fisher dispersion test (Cochran (1952, 1954)).

We define the null and the alternative hypotheses as follows:

- H_0: Conditional on the firm-specific and macroeconomic environment at time t, the occurrence of operational risk events has a Poisson distribution with intensity $\lambda(t)$.

- H_A: Conditional on the firm-specific and macroeconomic environment at time t, the occurrence of operational risk events does not have a Poisson distribution with intensity $\lambda(t)$.

We develop a two-step procedure for the χ^2 goodness-of-fit test:[24]

- Step 1: Test of the conditional Poisson assumption at the individual firm level.

- Step 2: Test of the joint conditional Poisson assumption for all firms.

A rejection of the null at the individual firm level (Step 1) would be evidence that our estimated intensity does not provide a good fit to the time-series of loss arrivals for a particular firm. On the other hand, if the conditional Poisson assumption at the individual firm level is not violated, one can then proceed to Step 2 and test the joint conditional Poisson assumption across all such firms. This second step relies on the following property of the Poisson distribution:

Theorem 1 *Let $N_{t1}, N_{t2}, \ldots, N_{tM}$ be M independent Poisson random variables with respective intensity rates $\lambda_1(t), \lambda_2(t), \ldots, \lambda_M(t)$. Then, $\sum_{l=1}^{M} N_{tl}$ is also a Poisson random variable with intensity rate $\sum_{l=1}^{M} \lambda_l(t)$.*

An inability to reject the joint conditional Poisson assumption at the aggregate level serves as evidence of the idiosyncratic nature of operational risk events across firms after all of the risk

[24] Our procedure differs slightly from a similar test applied to corporate defaults by Das, Duffee, Kapadia, and Saita (2007). Because default is typically modeled as a single-jump process, it is not possible to test the conditional Poisson hypothesis for the default process of each firm individually. In contrast, a typical financial firm can experience multiple operational losses in our sample period. This makes it feasible to conduct Step 1 of our test procedure.

factors have been adequately controlled for. On the other hand, a rejection of the joint conditional Poisson assumption, given that the conditional Poisson assumption is not violated at the individual firm level, suggests that the event arrivals are not conditionally independent across firms.

It is notable that if the conditional Poisson assumption is violated at the firm level, it would not necessarily imply that this same condition will be violated at the aggregate level. For example, if a missing common factor is responsible for the violation at individual firms, and if the associated factor loading is random across firm, then it is entirely possible that the aggregate arrival process can still be conditionally Poisson (Jacod (1975)). Therefore, we can conduct Step 2 of the goodness-of-fit test either across all firms, or across a subset of the firms for which the conditional Poisson assumption is not rejected individually. Of course, the latter test procedure allows us to more clearly identify the source of the problem should there be a violation of the joint conditional Poisson hypothesis.

For testing a standard Poisson process with rate one, our test statistic is calculated as

$$W^2 = \sum_{k=1}^{K} \frac{(n_k - c)^2}{c}, \tag{9}$$

where n_k is the observed frequency of arrivals in bin k, and c is the size of the bins (expected frequency under the null) given a total of K bins. Under the null, the limiting distribution of the statistic is χ^2_{K-1} with $K-1$ degrees of freedom. A low p-value, defined as the probability that a χ^2_{K-1} random variable exceeds W^2, would suggest rejection of the null hypothesis in favor of the alternative.

6.1 Time-Scale Transformation

The χ^2 test statistic as defined in Equation (9) cannot be directly applied to Poisson processes of a non-homogeneous nature. One alternative is to transform a non-homogeneous process into a process with a homogeneous intensity rate via a time-scale transformation. The time-scale transformation method is due to the following result:[25]

Theorem 2 $t_1, t_2, \ldots$ *are the points in a non-homogeneous Poisson process with a continuous cumulative intensity function* $\Lambda(t)$ *if and only if* $t'_1 = \Lambda(t_1), t'_2 = \Lambda(t_2), \ldots$ *are the points in a homo-*

[25] See Brémaud (1981, pp. 41) or Çinlar (1975, pp. 98-99).

geneous Poisson process with intensity rate one. The cumulative intensity function on the interval $[0, t]$ *is defined as* $\Lambda(t) := \int_0^t \lambda(s)ds$.

In essence, this property guarantees that any non-homogeneous Poisson process can be transformed into a standard Poisson process (i.e., a Poisson process with intensity rate one) by appropriately speeding up the clock during periods of high intensity and slowing down the clock during periods of low intensity.

We implement this time-scale transformation on the monthly loss frequency data using the estimated loss intensity function after all insignificant covariates are dropped. During the first stage of this procedure, we apply the time-scale transformation to individual firms, using the estimated intensities to generate a firm-specific set of bins of size c. We then compute the test statistic in Equation (9) for each firm. In the second stage, we compute the sum of the intensities for all firms remaining in our sample each month, and construct the bins of size c using this aggregate intensity. The test statistic in (9) then allows us to test whether the (time-scale transformed) aggregate arrival of operational loss events follows a standard Poisson process. For ease of comparison across the individual and aggregate level tests, we opt for bin sizes 1, 2, 3, 4, 5, and 8.

6.2 Firm-Level Results

Table 10 presents the results of the χ^2 test on the time-scale transformed frequency data. The χ^2 test determines whether the doubly-stochastic Poisson assumption is valid for each firm and each event category individually, conditional on the firm-specific and macroeconomic covariates. The table reports the proportion of firms for which the p-values are smaller than $\alpha = \{0.10, 0.05, 0.01\}$. Larger proportions indicate widespread violation of the conditional Poisson assumption at the individual firm level.

We obtain mixed evidence regarding the validity of the conditional Poisson frequency assumption. At bin size 1, only a quarter to a third of the firms violate the assumption. The proportion of violations generally increases with the bin size. Overall, between one and three quarters of the firms have failed the Poisson goodness-of-fit test at various bin sizes and significance levels.

Why does our conditional Poisson model fail to describe the arrivals of operational loss events for a significant portion of the firms? In an effort to draw strong inferences from a limited number of loss events, we have included only a parsimonious set of covariates. As in the literature on bankruptcy prediction, the objective is not about achieving the highest possible R^2 or cumulative accuracy ratio per se, but identifying important economic covariates that help explain the incidence of a broad sample of loss events. Although we have identified several explanatory covariates in Section 4 that are highly relevant to the frequency and severity of operational losses, it is likely that missing covariates or unobserved "frailty" factors have contributed to the poor performance of the model among a subset of the firms.

6.3 Aggregate-Level Results

Following the second step of our procedure, we test the joint conditional Poisson assumption. First, we test whether the conditional Poisson assumption holds for all firms jointly.

Panel A of Table 11 describes the results of the χ^2 test of the joint conditional Poisson assumption for all firms in our sample. It reports near-zero p-values for almost all of the models, indicating a resounding rejection of the null hypothesis.

Of course, one reason for this result is perhaps that the firms for which the conditional Poisson assumption is rejected individually are included in the test. In an attempt to address this issue, we exclude all firms whose individual p-values from the firm-level χ^2 test (Table 10) are less than $\alpha = \{0.10, 0.05, 0.01\}$ for bins of size 1, sizes 1 and 2, and sizes 1, 2, and 3. The case for which we exclude firms with individual p-values less than 0.10 is the most restrictive, less than 0.05 moderately restrictive, and less than 0.01 the least restrictive filtering of firms. Similarly, the case for which we exclude firms that failed the conditional Poisson test for bins of size 1 is the least restrictive, bins of sizes 1 and 2 moderately restrictive, and bins of sizes 1, 2, and 3 the most restrictive filtering of firms. Intuitively, by using the most stringent form of filtering, we are screening for firms whose operational loss arrivals best conform to a conditional Poisson model. Among this subset, a violation of the joint conditional Poisson assumption is undoubtedly because event arrivals are not conditionally independent across firms. For less stringent filterings, it is possible for firm-level fitting errors to "contaminate" the results of the joint conditional Poisson test.

As shown in Panels B-D of Table 11, the null hypothesis cannot be rejected for a majority of the cases under Model 3 (all except internal fraud), Model 6 (external events), Model 7 (physical disasters), and Model 8 (all events). This result holds even at larger bin sizes, and is not particularly sensitive to the restrictiveness of the filtering criteria. While this suggests that the conditional Poisson assumption works reasonably well among a subset of the firms, there are significant differences across the event type models. For example, the null hypothesis is rejected for many cases involving Model 1 (internal fraud), Model 2 (internal and external fraud), and Model 5 (internal events), particularly at larger bin sizes.

6.4 Discussions

What could be responsible for the lack of conditional independence of operational loss arrivals across firms? One obvious possibility is that the time-series of bin counts based on the time-changed joint loss arrival process may not be serially independent. This can, for instance, result from a missing macroeconomic covariate that is serially correlated over time. If most firms have exposures to this missing factor, then it would not be surprising that the bin counts are autocorrelated. We therefore estimate the time-series of bin counts as an AR(1) process for each event type category.

In results not reported here, we find that none of the estimated autoregressive coefficients is statistically significant. However, these coefficients are uniformly negative, indicating that periods with many losses are interspersed with periods with few losses. In particular, those categories of loss events that have previously shown a higher tendency for rejecting the joint conditional Poisson assumption (Models 1, 2, and 5) have autoregressive coefficients that are larger in magnitude.

We also regress the bin counts on a large set of current and lagged macroeconomic covariates. While the explanatory power of these variables is generally weak, there is some evidence that the bin counts are negatively related to the current change in the industrial production index, as well as a time dummy variable that equals 1 for the sub-sample period in which the Sarbanes-Oxley Act was in effect (August 2002 to December 2003).[26] Moreover, for internal fraud we find that the bin counts are negatively related to the industry-wide count of loss events (of all types) that have been settled in the past three years. This is consistent with firms learning from the industry-wide loss experience in the recent past. We leave a systematic search for additional explanatory covariates to future research.

[26] In documenting the claimant of each loss event, we find that about 30 percent of the cases were reported by regulatory agencies such as the SEC, NASD, and NYSE. However, in 2002 and 2003 this percentage was in the range of 50-60 percent, consistent with "regulatory pressure" causing a reduction in the incidence of operational losses.

7 Conclusions

In this paper, we use 24 years of operational loss data to investigate the determinants of operational risk in U.S. financial institutions. Our sample consists of operational losses reported by third parties such as SEC press reports, the NYSE, court decisions, and affected customers. As such, they are a good representation of unexpected losses that could no longer be hidden from the public. We use a number of firm-specific and market-level covariates to explain the operational loss frequency and severity for a variety of loss event types. Specifically, we make the following contributions:

1. Our results demonstrate that operational risk has a prominent firm-specific dependence. Factors such as scale, capital structure, profitability, and volatility play a significant part in operational loss formation. In particular, we show that internal events such as frauds are closely linked to proxies for the internal control environment of the firm. Furthermore, we find a significant and nonlinear relation between operational risk and the human factor: as the number of employees increases, the frequency of internal loss events first increases, and then falls.

2. The market-related factors play a weaker and mixed role in determining the levels of operational risk in firms. A persistent finding is that the frequency and severity of operational losses are negatively related to the GDP growth rate, indicating that they tend to rise during economic downturns. It is possible, however, that this result is primarily driven by the damage to physical assets during September 2001 for this sample.

3. Even after conditioning on the time-varying arrival intensities, we are unable to find sufficient support for the hypothesis that the conditional Poisson framework fully accounts for the clustering of loss events over time and across firms. Our goodness-of-fit tests demonstrate that the assumption is violated for a significant portion of the firms in our sample. More importantly, the joint conditional Poisson assumption is often rejected even among firms whose operational risk events do not violate the conditional Poisson assumption at the individual firm level. These rejections appear to be concentrated among internal or fraud-related loss events.

The last result suggests that there are unidentified factors driving operational risk across financial institutions, causing a significant clustering of specific types of loss events. This is likely to be of interest to risk managers and bank regulators, who must assess existing models or invent new ones to effectively cope with operational risk from a portfolio perspective.

References

[1] Allen, L., and T. Bali, 2007, Cyclicality in catastrophic and operational risk measurements, *Journal of Banking and Finance* 31, 1191-1235.

[2] Arnold, M., P. T. Larsen, P. Hollinger, J. O'Doherty, and R. Milne, 2008, How Kerviel exposed lax controls at Société Générale, *Financial Times*, February 7.

[3] BCBS, 1998, Operational risk management, Bank of International Settlements, Basel, Switzerland.

[4] BCBS, 1999, A new capital adequacy framework, Bank of International Settlements, Basel, Switzerland.

[5] BCBS, 2006a, International convergence of capital measurement and capital standards, Bank of International Settlements, Basel, Switzerland.

[6] BCBS, 2006b, Observed range of practice in key elements of advanced measurement approaches (AMA), Bank of International Settlements, Basel, Switzerland.

[7] Brémaud, P., 1981, *Point processes and queues: Martingale dynamics*, Springer-Verlag, New York.

[8] Carey, M., and R. Stulz, 2006, Introduction, in *The risk of financial institutions*, M. Carey and R. Stulz, eds., University of Chicago Press, 1-25.

[9] Çinlar, E., 1975, *Introduction to stochastic processes*, Prentice-Hall, Englewood Cliffs, New Jersey.

[10] Chavez-Demoulin, V., P. Embrechts, and J. Nešlehová, 2006, Quantitative models for operational risk: Extremes, dependence, and aggregation, *Journal of Banking and Finance* 30, 2635-2658.

[11] Cochran, W. G., 1952, The χ^2 test of goodness of fit, *Annals of Mathematical Statistics* 23, 315-345.

[12] Cochran, W. G., 1954, Some methods for strengthening the common χ^2 tests, *Biometrics* 10, 417-451.

[13] Collin-Dufresne, P., R. S. Goldstein, and J. Helwege, 2003, Is credit event risk priced? Modeling contagion via the updating of beliefs, Working paper, Carnegie-Mellon University.

[14] Cummins, J. D., C. M. Lewis, and R. Wei, 2006, The market value impact of operational risk events for U.S. banks and insurers, *Journal of Banking and Finance* 30, 2605-2634.

[15] Daniélsson, J., P. Embrechts, C. Goodhart, C. Keating, F. Muennich, O. Renault, and H. S. Shin, 2001, An academic response to Basel II, Working paper, London School of Economics.

[16] Das, S., D. Duffie, N. Kapadia, and L. Saita, 2007, Common failings: How corporate defaults are correlated, *Journal of Finance* 62, 93-117.

[17] Dechow, P. M., R. G. Sloan, and A. P. Sweeney, 1996, Causes and consequences of earnings manipulation: An analysis of firms subject to enforcement actions by the SEC, *Contemporary Accounting Research* 13, 1-36.

[18] de Fontnouvelle, P., V. DeJesus-Rueff, J. Jordan, and E. Rosengren, 2006, Capital and risk: New evidence on implications of large operational losses, *Journal of Money, Credit, and Banking* 38, 1819-1846.

[19] de Fontnouvelle, P., E. Rosengren, and J. Jordan, 2006, Implications of alternative operational risk modeling techniques, in *The risk of financial institutions*, M. Carey and R. Stulz, eds., University of Chicago Press, 475-512.

[20] Duffie, D., A. Eckner, G. Horel, and L. Saita, 2006, Frailty correlated default, Working paper, Stanford University.

[21] Duffie, D., L. Saita, and K. Wang, 2007, Multi-period corporate default prediction with stochastic covariates, *Journal of Financial Economics* 83, 635-665.

[22] Duffie, D., and K. J. Singleton, 1999, Modeling term structures of defaultable bonds, *Review of Financial Studies* 12, 687-720.

[23] Efendi, J., A. Srivastava, and E. P. Swanson, 2007, Why do corporate managers misstate financial statements? The role of option compensation and other factors, *Journal of Financial Economics* 85, 667-708.

[24] Fama, E. F., and K. R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427-465.

[25] Helwege, J., and P. Kleinman, 1997, Understanding aggregate default rates of high yield bonds, *Journal of Fixed Income* 7(1):55–62.

[26] Jacod, J., 1975, Two dependent Poisson processes whose sum is still a Poisson process, *Journal of Applied Probability* 12, 170-172.

[27] Jarrow, R. A., 2007, Operational risk, Working paper, Cornell University, *Journal of Banking and Finance*, forthcoming.

[28] Jarrow, R. A., and S. M. Turnbull, 1995, Pricing derivatives on financial securities subject to credit risk, *Journal of Finance* 50, 53-85.

[29] Jarrow, R. A., and F. Yu, 2001, Counterparty and the pricing of defaultable securities, *Journal of Finance* 56, 1765-1800.

[30] Jorion, P., and G. Zhang, 2007, Good and bad credit contagion: Evidence from credit default swaps, Working paper, University of California-Irvine, *Journal of Financial Economics*, forthcoming.

[31] Lando, D., 1998, On Cox processes and credit risky securities, *Review of Derivatives Research* 2, 99-120.

[32] Perry, J., and P. de Fontnouvelle, 2005, Measuring reputational risk: The market reaction to operational loss announcements, Working paper, Federal Reserve Bank of Boston.

[33] Povel, P., R. Singh, and A. Winton, 2007, Booms, busts, and fraud, *Review of Financial Studies* 20, 1219-1254.

[34] Rosenberg, J. V., and T. Schuermann, 2006, A general approach to integrated risk management with skewed, fat-tailed risks, *Journal of Financial Economics* 79, 569-614.

[35] Yu, F., 2007, Correlated defaults in intensity-based models, *Mathematical Finance* 17, 155-173.

Table 1. Distribution of event duration by event type, 1980-2005.
This table reports the distribution of the duration of individual operational risk events in our original sample from 1980 to 2005. Primary figures are presented in months and figures in parentheses are in years. N is the number of observed events. The event type categories are defined as: ET1-internal fraud, ET2-external fraud, ET3-employment practices and workplace safety, ET4-clients, products, and business practices, ET5-damage to physical assets, ET6-business disruption and system failures, ET7-execution, delivery, and process management, and Other.

Percentile	ET1	ET2	ET3	ET4	ET5	ET6	ET7	Other	Full sample
25	5.51	1.03	0.13	12.17	0.03	0.07	0.07	3.89	3.00
	(0.46)	(0.09)	(0.01)	(1.01)	(0.003)	(0.01)	(0.01)	(0.32)	(0.25)
50	19.97	12.20	12.23	29.37	0.03	0.20	23.13	6.33	20.33
	(1.66)	(1.02)	(1.02)	(2.45)	(0.003)	(0.02)	(1.93)	(0.53)	(1.69)
75	45.22	36.57	48.73	60.90	0.03	12.20	48.73	19.22	48.73
	(3.77)	(3.05)	(4.06)	(5.08)	(0.003)	(1.02)	(4.06)	(1.60)	(4.06)
80	52.61	40.38	50.72	64.30	0.03	18.27	63.52	24.40	58.87
	(4.38)	(3.37)	(4.23)	(5.36)	(0.003)	(1.52)	(5.29)	(2.03)	(4.91)
85	60.87	48.73	66.36	73.07	0.12	23.83	73.10	45.03	68.06
	(5.07)	(4.06)	(5.53)	(6.09)	(0.01)	(1.52)	(6.09)	(3.75)	(5.67)
90	73.03	59.27	76.89	91.97	2.74	39.17	103.43	48.73	79.03
	(6.09)	(4.94)	(6.41)	(7.66)	(0.23)	(3.26)	(8.62)	(4.06)	(5.87)
95	103.20	94.11	120.27	111.64	4.98	59.33	117.18	57.12	109.60
	(8.60)	(7.84)	(10.02)	(9.30)	(0.42)	(4.94)	(9.76)	(4.76)	(9.13)
99	180.58	148.63	194.33	182.47	15.93	73.07	163.79	161.36	177.76
	(15.05)	(12.39)	(16.19)	(15.21)	(1.33)	(6.09)	(13.65)	(13.45)	(14.81)
N	213	146	129	670	28	40	106	87	1,419

Table 2. Annual frequency and severity of operational risk events by event type, 1980-2003.
This table reports the operational risk distribution by event type over our final sample from 1980 to 2003 for all events that originated prior to 2004. Panel A reports the cumulative annual frequency (at event "start date") of operational risk events. Panel B reports the cumulative annual severity (at event "end date") of operational losses. The event type categories are defined as: ET1-internal fraud, ET2-external fraud, ET3-employment practices and workplace safety, ET4-clients, products, and business practices, ET5-damage to physical assets, ET6-business disruption and system failures, ET7-execution, delivery, and process management, and Other.

Panel A: Frequency of operational risk events

Year	ET1	ET2	ET3	ET4	ET5	ET6	ET7	Other	Full sample
1980	3	0	0	5	0	0	8	0	16
1981	2	0	0	0	0	0	0	0	2
1982	3	2	0	4	0	0	0	0	9
1983	0	1	0	0	0	1	1	0	3
1984	2	1	2	6	0	0	1	0	12
1985	6	2	1	10	0	1	1	0	21
1986	1	2	1	6	0	0	3	1	14
1987	4	3	0	12	0	1	1	1	22
1988	1	3	1	4	0	0	3	1	13
1989	8	1	0	25	1	0	7	0	42
1990	3	3	0	13	0	1	0	0	20
1991	4	2	5	9	0	2	2	0	24
1992	6	1	3	12	0	1	3	2	28
1993	6	5	8	12	0	0	0	0	31
1994	9	7	3	22	0	1	4	1	47
1995	9	9	4	29	0	1	1	0	53
1996	10	1	8	23	0	2	5	1	50
1997	13	13	10	34	0	3	5	6	84
1998	17	6	10	42	0	3	2	5	85
1999	23	8	10	70	0	4	3	2	120
2000	11	19	9	48	0	6	6	9	108
2001	26	11	13	63	16	4	8	19	160
2002	7	3	8	57	5	2	14	26	122
2003	6	7	8	43	0	1	7	1	73
Total	180	110	104	549	22	34	85	75	1,159
	(15.5%)	(9.5%)	(9%)	(47.4%)	(1.9%)	(2.9%)	(7.3%)	(6.5%)	(100%)

Panel B: Severity of operational losses (USD'000)

Year	ET1	ET2	ET3	ET4	ET5	ET6	ET7	Other	Full sample
1980	33,257	0	0	24,300	0	0	0	0	57,557
1981	44,466	0	0	0	0	0	0	0	44,466
1982	0	0	0	0	0	0	0	0	0
1983	0	588,141	0	0	0	0	0	0	588,141
1984	9,722	0	373	19,265	0	0	45,268	0	74,628
1985	73,429	148,000	0	3,005	0	9,247	939	0	235,190
1986	2,724	0	0	78,822	0	0	51,291	0	132,837
1987	83,551	0	0	429,564	0	39,470	0	0	552,586
1988	18,614	8,418	0	654,540	0	137,000	90,189	0	908,572
1989	634,871	12,321	49,702	428,658	14,980	0	203,200	226,000	1,569,726
1990	430,400	16,235	0	194,213	0	0	44,687	0	685,535
1991	24,194	300,120	2,277	719,079	0	3,113	0	0	1,048,782
1992	93,106	4,631	3,792	73,927	0	0	156,725	0	332,181
1993	23,358	7,730	76,791	683,401	0	4,207	48,994	0	844,481
1994	115,064	145,610	248,121	1,616,431	0	402,000	6,212	0	2,534,392
1995	4,346	58,233	8,727	3,230,676	0	0	15,138	0	3,313,372
1996	136,120	187,221	5,248	537,060	0	188,000	147	0	1,053,795
1997	94,466	104,328	41,578	358,978	0	12,715	27,062	0	639,665
1998	141,693	62,999	288,483	3,376,947	0	228	8,222	1,562,641	5,433,867
1999	38,562	14,157	45,548	512,635	0	2,132	49,178	1,197,321	1,859,532
2000	132,602	46,197	305,753	1,714,936	0	53,819	519,473	0	2,773,461
2001	239,787	93,206	37,363	9,662,150	3,971,099	10,052	72,479	2,486,664	16,567,354
2002	193,966	237,627	90,904	5,542,439	2,510,000	0	139,517	7,651,464	16,360,215
2003	837,072	34,996	62,440	2,436,360	0	0	86,627	0	3,456,725
Total	3,405,369	2,070,170	1,267,100	32,297,385	6,496,079	861,983	1,565,348	13,124,090	61,067,061
	(5.6%)	(3.4%)	(2.1%)	(52.9%)	(10.6%)	(1.4%)	(2.6%)	(21.5%)	(100%)
Average	18,919	18,820	12,184	58,829	295,276	25,352	18,416	174,988	52,689

Table 3. Summary statistics of monthly loss frequency and severity.

This table summarizes the monthly aggregated operational loss data during months with non-zero loss entries from 1980 to 2003 by event type. Panel A reports the descriptive statistics of loss frequency during months in which at least one operational risk event has started (at "start date"). Panel B reports the descriptive statistics of loss severity during months in which at least one operational risk event has ended (at "end date"). The total number of months in the sample period is 288. The event type categories are defined as: ET1-internal fraud, ET2-external fraud, ET3-employment practices and workplace safety, ET4-clients, products, and business practices, ET5-damage to physical assets, ET6-business disruption and system failures, ET7-execution, delivery, and process management, and Other.

Panel A: Loss frequency

	ET1	ET2	ET3	ET4	ET5	ET6	ET7	Other	Full sample
N	73	44	46	107	6	27	43	28	164
Min	1	1	1	1	1	1	1	1	1
Mean	2.47	2.50	2.26	5.13	3.67	1.26	1.98	2.68	7.07
StDev	2.85	2.68	2.24	9.70	5.13	0.53	1.90	4.38	14.67
Median	1	1	1	1	1.5	1	1	1	2
Max	17	11	9	53	14	3	11	22	86

Panel B: Loss severity

	ET1	ET2	ET3	ET4	ET5	ET6	ET7	Other	Full sample
N	69	38	36	112	4	14	39	14	159
Mean	49,353	54,478	35,197	288,370	1,624,020	61,570	40,137	937,435	384,070
StDev	108,261	105,788	60,791	754,869	1,958,851	113,495	82,495	1,559,451	1,205,198
Min	53	422	107	67	1,121	84	2	5,250	2
25p	2,167	5,591	2,227	7,432	8,050	2,048	2,279	34,303	7,601
50p	12,136	14,481	9,177	43,037	1,262,490	9,649	13,896	258,260	44,466
75p	43,116	60,297	36,781	226,823	3,239,989	53,587	44,687	1,480,961	236,645
99p	588,000	580,000	250,000	4,003,094	3,969,978	402,000	481,759	5,868,818	8,121,361
Max	588,000	580,000	250,000	5,245,247	3,969,978	402,000	481,759	5,868,818	10,200,000

Table 4. Description of events and major contributory factors.
Our final data sample consists of publicly reported operational loss events from the financial industry from 1980 to 2003. This table summarizes the descriptions and key contributory factors of the events.

Events Details	Major Contributory Factors
ET1: Internal Fraud	
Credit fraud (34%)	Lack of control (31%)
Theft, embezzlement, robbery (23%)	Employee action/inaction (25%)
Unauthorized transactions (12%)	Management action/inaction (15%)
Insider trading (9%)	Omissions (e.g., failure to supervise employees, inadequate due diligence efforts) (14%)
Misappropriation of assets (8%)	Organizational structure: Excessive concentration of power (5%)
Intentional mismarking of position (7%)	Changes in market conditions (e.g., M&A, regulatory pressure) (4%)
Bribes (<1%)	Strategy flaws (<1%)
Other (e.g., tax evasion, forgery) (5%)	Corporate governance (<1%)
	Other or unspecified (5%)
ET2: External Fraud	
Theft, fraud, forgery, and robbery (56%)	Omissions (e.g., inadequate due diligence efforts) (48%)
System security and hacking (12%)	Lax security (18%)
Other (e.g., loan fraud) (32%)	Lack of internal control (10%)
	Management action/inaction (4.5%)
	Employee inaction/inaction (3%)
	Changes in market conditions (e.g., new technology) (3%)
	Strategy flaws (<1%)
	Other or unspecified (13%)
ET3: Employment Practices & Workplace Safety	
Employment discrimination (56.2%)	Management action/inaction (26%)
Compensation, benefit, termination issues (42%)	Lack of control, insufficient compliance measures (22%)
Safety of environment (<1%)	Employee action/inaction (17%)
Organized labor activity (<1%)	Staff selection and compensation (11%)
Other or unspecified (<1%)	Omissions (9%)
	Changes in market conditions (e.g., M&A, regulatory pressure) (3%)
	Strategy flaws (3%)
	Corporate governance (e.g., Sarbanes-Oxley violations) (<1%)
	Other or unspecified (7%)
ET4: Clients, Products, and Business Practices	
Suitability, disclosure, and fiduciary (e.g., disclosure issues, lender liability, fiduciary breaches) (51%)	Lack of control (29%)
Improper business and market practices (e.g., unlicensed activity, money laundering, market manipulation, improper trade, antitrust) (44%)	Omissions (e.g., lack of proper training procedures, inadequate due diligence efforts, failure to supervise employees) (22%)
Other (e.g., misuse of confidential information, advisory activities) (5%)	Changes in market conditions (e.g., M&A, regulatory pressure) (17%)
	Management action/inaction (16%)
	Employee action/inaction (7%)
	Strategy flaws (3%)
	Corporate governance (e.g., Sarbanes-Oxley violations) (1%)
	Organization structure, excessive concentration of power (<1%)
	Other (e.g., new technology, failure to comply with established policies) (5%)

ET5: Damage to Physical Assets	
Terrorism, vandalism (e.g., "9/11") (95%) Natural disasters (5%)	Omissions (12%) Changes in market conditions (7%) Strategy flaws (8%) Employment action/inaction (2%) Other and unspecified (71%)
ET6: Business Disruptions & System Failures	
Software failures (65%) Hardware failures (7%) Telecommunications (3%) Utility outage/disruptions (1%) Other technology failures (24%)	Strategy flaw: inadequate technology planning (45%) Lack of internal control (18%) Omissions (e.g., failure to test equipment) (15%) Employee inaction/inaction (8%) Changes in market conditions (6%) Management action/inaction (3%) Other (5%)
ET7: Execution, Delivery, & Process Management	
Transaction execution and maintenance (e.g., accounting error, data entry error) (54%) Failed or inaccurate mandatory reporting (31%) Customer/client account mismanagement (11%) Other (4%)	Lack of control (e.g., poor documentation, lax security, insufficient compliance measures, failure to test for data accuracy) (31%) Omissions (e.g., failure to supervise employees, inadequate due diligence efforts) (23%) Management action/inaction, poor execution (18%) Employee action/inaction, misdeeds, errors (10%) Change in market conditions (e.g., M&A, regulatory pressure, financial reporting) (4%) Strategy flaw (4%) Other (8%)
Other	
Losses due to new market regulations Strategy failures Enron-related M&A Other	Changes in market conditions (41%) Omissions (e.g., inadequate due diligence efforts) (14%) Strategy flaws (8%) Lack of control (6%) Management action/inaction (4%) Other or unspecified (27%)

Table 5. Model descriptions.
The event type categories are defined as: ET1-internal fraud, ET2-external fraud, ET3-employment practices and workplace safety, ET4-clients, products, and business practices, ET5-damage to physical assets, ET6-business disruption and system failures, ET7-execution, delivery, and process management, and Other.

Model 1	Internal Fraud (ET1)
Model 2	Fraud (ET1 and ET2)
Model 3	All Except Internal Fraud
Model 4	All Except Fraud
Model 5	Internal Events (ET1, ET3, ET6, and ET7)
Model 6	External Events (ET2, ET4, ET5, and Other)
Model 7	Physical Disasters (ET5)
Model 8	All Events

Table 6. Description of model variables.

Variable Name	Description	Measurement Units	Source
$Numloss^{start}_i$, $i=1,2,\ldots,8$	Monthly aggregated event count at "start date," for a given event type category	Count	Algo FIRST
$Numloss^{end}_i$, $i=1,2,\ldots,8$	Monthly aggregated event count at "end date," for a given event type category	Count	Algo FIRST
$Loss_i$, $i=1,2,\ldots,8$	Monthly aggregated loss amount at "end date," for a given event type category	USD'000	Algo FIRST
$Duration_i$, $i=1,2,\ldots,8$	Average duration of events, for a given event type category	Months	Algo FIRST
LogMVE	Logarithm of market value of equity	Log USD'000	COMPUSTAT quarterly: DATA61*DATA14
TA	Total assets	USD'000	COMPUSTAT quarterly: DATA44
Leverage	Leverage	ratio	COMPUSTAT quarterly: DATA54/(DATA54 + DATA59)
ROA	Return on assets	ratio	COMPUSTAT quarterly: DATA21/DATA44
Retsd	Trailing standard deviation of monthly stock returns over the past year	Decimal, annualized	CRSP monthly: holding period return
Market-to-book	Market-to-book ratio	ratio	COMPUSTAT quarterly: (DATA54+DATA61*DATA14)/DATA44
LogEmpl	Logarithm of number of employees	Log '000	COMPUSTAT annual: DATA29
LogSpread	Logarithm of yield spread estimated as U.S. Baa corporate bond yield minus Aaa corporate bond yield, middle rate	Log USD	Moody's monthly
Unemplr	Civilian unemployment rate	Decimal, annualized	U.S. Department of Labor database, monthly
Tbill3mr	3-month trailing Treasury Bill rate (secondary market)	Decimal, annualized	Board of Governors of the Federal Reserve System, monthly
S&P1mr	S&P 500 1-month return	Decimal, annualized	S&P's website, monthly
S&P1mrsd	Trailing standard deviation of S&P 500 1-month returns over the last 3 years	Decimal, annualized	S&P's website, monthly
GDPgr	GDP growth rate	Decimal, annualized	Datastream, quarterly

Table 7. Regression results for operational risk frequency by event type with microeconomic and macroeconomic covariates. The dependent variable is the monthly aggregated loss count *Numloss*^{start}. Columns denoted by "A" indicate models with only firm-specific covariates. Columns denoted by "B" indicate models with firm-specific and macroeconomic covariates. The panel data Poisson regression models are estimated with the method of maximum likelihood. All models except Model 7 include a January dummy; estimates are omitted for brevity. The *t*-statistics (in parentheses) are based on heteroskedasticity-robust standard errors. ***, **, and * denote statistical significance of the coefficients at the 1, 5, and 10 percent levels, respectively.

	Model 1 Internal Fraud		Model 2 Fraud		Model 3 All Except Internal Fraud		Model 4 All Except Fraud	
	A	B	A	B	A	B	A	B
Const	-24.8825	-27.7912	-24.3660	-27.6231	-19.6371	-16.9998	-19.5623	-15.9756
	(-10.20)***	(-8.55)***	(-12.55)***	(-9.96)***	(-18.37)***	(-10.53)***	(-17.46)***	(-9.43)***
LogMVE	0.5839	0.6957	0.5303	0.6714	0.5865	0.5281	0.6014	0.5139
	(5.77)***	(5.76)***	(6.02)***	(6.18)***	(9.84)***	(8.23)***	(9.65)***	(7.87)***
Leverage	7.0324	6.3604	7.0414	6.3836	3.2378	3.1290	2.9355	2.8536
	(3.49)***	(3.22)***	(4.26)***	(4.00)***	(3.67)***	(3.55)***	(3.25)***	(3.15)***
Market-to-book	-0.9702	-1.1183	-0.8540	-1.0368	-0.4268	-0.4314	-0.3915	-0.3834
	(-2.19)**	(-2.24)**	(-2.47)***	(-2.75)***	(-1.87)*	(-1.86)*	(-1.72)*	(-1.70)*
ROA	27.0562	25.9866	21.3273	20.3205	5.8282	6.1541	5.4281	6.1464
	(2.12)**	(2.02)**	(1.81)*	(1.67)*	(0.75)	(0.78)	(0.68)	(0.76)
Retsd	0.6469	0.6845	0.5817	0.6440	0.4390	0.3977	0.4345	0.3813
	(4.99)***	(4.88)***	(4.63)***	(4.86)***	(6.24)***	(5.41)***	(6.05)***	(5.09)***
LogEmpl	0.6933	0.5873	1.0683	0.9487	0.2775	0.3050	0.1618	0.2135
	(1.78)*	(1.54)	(2.79)***	(2.48)**	(1.71)*	(2.03)**	(1.05)	(1.51)
LogEmpl^2	-0.1160	-0.1127	-0.1516	-0.1489	-0.0534	-0.0509	-0.0409	-0.0387
	(-2.25)**	(-2.14)**	(-3.11)***	(-2.92)***	(-2.39)**	(-2.33)**	(-1.88)*	(-1.83)*
LogSpread		-1.2408		-0.9772		0.1294		0.2518
		(-2.31)**		(-2.45)**		(0.51)		(0.92)
Unemplr		27.4634		23.4099		-16.9483		-22.7627
		(1.65)*		(1.81)*		(-2.10)**		(-2.59)***
Tbill3mr		2.8489		3.6043		-0.6026		-1.3754
		(1.69)*		(2.97)***		(-0.78)		(-1.61)
S&P1mr		0.2819		0.2521		-0.0757		-0.1145
		(1.56)		(1.65)*		(-0.62)		(-0.88)
S&P1mrsd		0.7725		0.5287		-0.7700		-0.9588
		(0.62)		(0.56)		(-1.06)		(-1.18)
GDPgr		-1.7782		-0.8830		-1.9666		-2.3451
		(-0.84)		(-0.63)		(-2.16)**		(-2.29)**
Num. Obs.	17,290	17,266	17,290	17,266	17,290	17,266	17,290	17,266
χ^2 macro		12		16		9		12
		(0.0684)*		(0.0161)**		(0.1647)		(0.0544)*
Pseudo R^2	0.2533	0.2605	0.3242	0.3321	0.3691	0.3723	0.3472	0.3523

	Model 5 Internal Events		Model 6 External Events		Model 7 Physical Disasters		Model 8 All Events	
	A	B	A	B	A	B	A	B
Const	-23.0795	-24.7160	-19.4907	-15.7568	-32.2745	11.9474	-19.8661	-18.0602
	(-15.45)***	(-11.18)***	(-15.90)***	(-8.58)***	(-8.28)***	(0.78)	(-19.71)***	(-11.92)***
LogMVE	0.5025	0.6281	0.6210	0.5068	1.6871	1.1305	0.5881	0.5532
	(5.36)***	(7.52)***	(9.86)***	(6.96)***	(5.84)***	(3.60)***	(10.41)***	(9.28)***
Leverage	6.2224	5.7583	2.5596	2.4673	-3.2329	-2.0604	3.6600	3.5056
	(5.22)***	(5.25)***	(2.55)**	(2.43)**	(-1.51)	(-1.08)	(4.33)***	(4.18)***
Market-to-book	-0.5399	-0.6204	-0.4555	-0.4420	-1.5311	-1.1671	-0.5039	-0.5197
	(-1.99)**	(-2.25)**	(-1.58)	(-1.49)	(-1.20)	(-0.84)	(-2.28)**	(-2.29)**
ROA	27.2857	28.7157	-2.6010	-3.5748	-41.9166	-31.2136	8.9007	9.0971
	(3.66)***	(3.99)***	(-0.26)	(-0.35)	(-2.78)***	(-3.09)***	(1.26)	(1.27)
Retsd	0.6225	0.6670	0.3854	0.3115	0.2025	0.0229	0.4705	0.4408
	(7.03)***	(7.28)***	(4.72)***	(3.48)***	(0.46)	(0.04)	(6.91)***	(6.07)***
LogEmpl	1.0369	0.9374	0.0820	0.1656	0.2555	0.1315	0.3245	0.3327
	(3.48)***	(3.24)***	(0.53)	(1.16)	(0.29)	(0.20)	(2.07)**	(2.29)**
$LogEmpl^2$	-0.1533	-0.1520	-0.0279	-0.0270	-0.1179	-0.0427	-0.0610	-0.0584
	(-4.15)***	(-3.94)***	(-1.25)	(-1.25)	(-1.01)	(-0.59)	(-2.87)***	(-2.79)***
LogSpread		-0.3717		0.1448		4.7687		-0.0650
		(-1.02)		(0.52)		(3.14)***		(-0.27)
Unemplr		10.6121		-22.8820		-364.13		-10.3126
		(0.94)		(-2.40)**		(-2.66)***		(-1.28)
Tbill3mr		0.8292		-0.6791		-25.8650		-0.1366
		(0.80)		(-0.79)		(-2.91)***		(-0.19)
S&P1mr		0.2130		-0.1645		-2.7310		-0.0219
		(1.54)		(-1.19)		(-4.13)***		(-0.20)
S&P1mrsd		-0.2280		-0.8148		-21.8201		-0.5348
		(-0.24)		(-1.04)		(-3.05)***		(-0.78)
GDPgr		-2.2918		-1.6866		-62.8225		-1.9511
		(-1.75)*		(-1.63)		(-5.78)***		(-2.24)**
Num. Obs.	17,290	17,266	17,290	17,266	17,290	17,266	17,290	17,266
χ^2 macro		8		11		108		8
		(0.2096)		(0.0783)*		(0.0000)***		(0.2672)
Pseudo R^2	0.2838	0.2877	0.3733	0.3776	0.2403	0.5427	0.3720	0.3748

Table 8. Regression results for operational loss severity by event type with microeconomic and macroeconomic covariates. The dependent variable is the *logarithm* of the average loss percentage, *Loss/(Numloss^{end} *TA)*, calculated as the average of the monthly aggregated loss amount divided by total assets, given that the loss amount is non-zero. Columns denoted by "A" indicate models with only firm-specific covariates. Columns denoted by "B" indicate models with firm-specific and macroeconomic covariates. The econometric model is panel data OLS regression. The *t*-statistics (in parentheses) are based on heteroskedasticity-robust standard errors. ***, **, and * denote statistical significance of the coefficients at the 1, 5, and 10 percent levels, respectively.

	Model 1 Internal Fraud		Model 2 Fraud		Model 3 All Except Internal Fraud		Model 4 All Except Fraud	
	A	B	A	B	A	B	A	B
Const	22.0942	22.0846	20.7398	20.7483	7.5786	7.8136	6.8996	7.1288
	(6.08)***	(5.67)***	(7.46)***	(6.67)***	(5.70)***	(5.23)***	(5.03)***	(4.56)***
LogMVE	-0.8467	-0.8238	-0.8825	-0.8442	-0.7244	-0.7226	-0.6880	-0.6869
	(-6.32)***	(-5.87)***	(-8.56)***	(-8.13)***	(-11.66)***	(-11.19)***	(-10.51)***	(-10.08)***
Leverage	-16.1008	-15.4351	-14.9100	-14.5183	-5.6707	-5.4645	-5.6434	-5.3249
	(-6.32)***	(-5.64)***	(-6.15)***	(-5.95)***	(-4.77)***	(-4.55)***	(-4.51)***	(-4.18)***
Market-to-book	-2.2094	-1.6558	-1.1630	-0.9847	0.6191	0.6295	0.5921	0.6001
	(-1.02)	(-0.73)	(-2.00)**	(-1.65)	(2.23)**	(2.27)**	(2.12)**	(2.16)**
Retsd	-1.1040	-1.0781	-1.1093	-1.0532	-0.3057	-0.3801	-0.1973	-0.2899
	(-1.70)*	(-1.62)	(-2.33)**	(-2.14)**	(-1.41)	(-1.59)	(-0.86)	(-1.15)
LogSpread		0.4342		0.4270		0.1019		0.1045
		(0.67)		(0.82)		(0.28)		(0.25)
S&P1mrsd		-2.0361		-1.8825		0.0703		0.2550
		(-1.05)		(-1.35)		(0.06)		(0.20)
GDPgr		-4.5323		-2.6278		-3.9731		-5.4836
		(-1.02)		(-0.79)		(-1.90)*		(-2.43)***
Num. Obs.	90	89	143	142	405	405	352	352
F macro		1		1		2		3
		(0.4182)		(0.3987)		(0.1601)		(0.0517)*
R^2	0.5287	0.5375	0.5117	0.5183	0.3559	0.3642	0.3467	0.3617

	Model 5 Internal Events		Model 6 External Events		Model 7 Physical Disasters		Model 8 All Events	
	A	B	A	B	A	B	A	B
Const	9.5302	9.2822	8.6339	8.5195	7.3193	-1.9039	8.7400	9.2780
	(4.06)***	(3.56)***	(5.83)***	(5.16)***	(0.56)	(-0.11)	(6.88)***	(6.52)***
LogMVE	-0.8631	-0.7804	-0.7109	-0.7268	-0.7130	-0.7943	-0.7453	-0.7337
	(-9.09)***	(-8.17)***	(-10.68)***	(-10.33)***	(-1.91)*	(-2.90)**	(-13.20)***	(-12.76)***
Leverage	-6.4836	-6.5632	-6.6014	-6.0921	-2.7603	-1.6671	-6.6454	-6.4272
	(-2.52)**	(-2.56)**	(-6.00)***	(-5.55)***	(-0.24)	(-0.19)	(-5.81)***	(-5.57)***
Market-to-book	0.8142	0.8791	0.5038	0.5059	4.7873	3.7650	0.5494	0.5687
	(5.93)***	(6.59)***	(1.17)	(1.18)	(1.47)	(1.75)	(1.91)*	(1.98)**
Retsd	-0.2439	-0.1091	-0.3817	-0.4925	-4.9921	-3.5019	-0.2908	-0.3559
	(-0.61)	(-0.25)	(-1.61)	(-1.89)*	(-1.90)*	(-2.44)*	(-1.45)	(-1.61)
LogSpread		0.9004		-0.1173		7.2762		0.3218
		(2.27)**		(-0.26)		(5.74)***		(1.03)
S&P1mrsd		-2.1616		0.8227		18.6900		-0.6377
		(-1.30)		(0.70)		(0.93)		(-0.65)
GDPgr		-1.6919		-3.8167		-65.8283		-5.1049
		(-0.54)		(-1.62)		(-3.31)**		(-2.70)***
Num. Obs.	170	169	326	326	14	14	479	478
F macro		2		2		30		4
		(0.0969)*		(0.1886)		(0.0005)***		(0.0142)**
R^2	0.3948	0.4070	0.3784	0.3876	0.6995	0.9254	0.3784	0.3918

Table 9. Regression results for operational risk frequency by event type with microeconomic and macroeconomic covariates for all financial firms.

The dependent variable is the monthly aggregated loss count *Numloss*^{start}. Columns denoted by "A" indicate models with only firm-specific covariates. Columns denoted by "B" indicate models with firm-specific and macroeconomic covariates. The panel data Poisson regression models are estimated with the method of maximum likelihood. All models except Model 7 include a January dummy; estimates are omitted for brevity. The *t*-statistics (in parentheses) are based on heteroskedasticity-robust standard errors. ***, **, and * denote statistical significance of the coefficients at the 1, 5, and 10 percent levels, respectively.

	Model 1 Internal Fraud		Model 2 Fraud		Model 3 All Except Internal Fraud		Model 4 All Except Fraud	
	A	B	A	B	A	B	A	B
Const	-25.4926 (-16.59)***	-25.0164 (-9.11)***	-24.7549 (-18.94)***	-26.1240 (-11.63)***	-22.2656 (-26.25)***	-19.5343 (-13.00)***	-22.2752 (-25.48)***	-18.5728 (-11.84)***
LogMVE	0.9777 (9.88)***	0.9740 (8.88)***	0.9244 (10.73)***	0.9797 (9.96)***	0.8524 (14.09)***	0.7601 (11.12)***	0.8498 (13.66)***	0.7247 (10.36)***
Leverage	0.0081 (6.06)***	0.0085 (5.58)***	0.0079 (6.85)***	0.0085 (6.82)***	0.0044 (3.03)***	0.0041 (2.63)***	0.0037 (1.97)**	0.0032 (1.52)
Market-to-book	-0.6560 (-2.51)**	-0.6871 (-2.42)**	-0.5073 (-2.65)***	-0.5549 (-2.79)***	-0.1574 (-1.50)	-0.1370 (-1.39)	-0.1327 (-1.26)	-0.1067 (-1.11)
ROA	4.6138 (0.69)	5.0382 (1.14)	-1.8573 (-3.88)***	-1.8534 (-3.60)***	-1.1109 (-5.93)***	-1.0421 (-5.50)***	-1.0166 (-4.61)***	-0.9318 (-4.10)***
Retsd	0.4281 (7.91)***	0.4232 (8.47)***	0.4146 (7.85)***	0.4134 (8.48)***	0.2570 (6.06)***	0.2353 (5.63)***	0.2444 (5.64)***	0.2214 (5.18)***
LogEmpl	0.4239 (1.49)	0.4003 (1.48)	0.5712 (2.10)**	0.4865 (1.92)*	0.2174 (2.65)***	0.2847 (3.35)***	0.1842 (2.36)**	0.2767 (3.30)***
LogEmpl^2	-0.0703 (-1.24)	-0.0654 (-1.17)	-0.0669 (-1.35)	-0.0584 (-1.21)	-0.0030 (-0.20)	-0.0043 (-0.26)	-0.0015 (-0.10)	-0.0032 (-0.20)
LogSpread		-0.8318 (-1.40)		-0.7646 (-1.81)*		0.0221 (0.08)		0.1231 (0.44)
Unemplr		-0.9773 (-0.05)		7.2650 (0.50)		-17.0989 (-1.88)*		-22.5109 (-2.31)**
Tbill3mr		1.8514 (0.98)		3.4210 (2.87)***		-0.7034 (-0.88)		-1.6334 (-1.85)*
S&P1mr		0.3114 (1.58)		0.1532 (0.97)		-0.0644 (-0.53)		-0.0610 (-0.48)
S&P1mrsd		-0.3616 (-0.31)		-0.0679 (-0.07)		-0.1819 (-0.24)		-0.2553 (-0.32)
GDPgr		-3.7303 (-1.52)		-2.1453 (-1.36)		-2.3465 (-2.35)**		-2.6418 (-2.42)**
Num. Obs.	157,350	157,118	157,350	157,118	157,350	157,118	157,350	157,118
χ^2 macro		10 (0.1367)		13 (0.0388)**		13 (0.0494)**		18 (0.0066)***
Pseudo R^2	0.3261	0.3362	0.4014	0.4094	0.4358	0.4410	0.4135	0.4205

	Model 5 Internal Events		Model 6 External Events		Model 7 Physical Disasters		Model 8 All Events	
	A	B	A	B	A	B	A	B
Const	-23.7767 (-18.01)***	-23.4781 (-12.01)***	-22.7178 (-24.99)***	-19.1178 (-11.27)***	-35.2176 (-11.18)***	3.2649 (0.24)	-22.2671 (-27.99)***	-19.7965 (-14.01)***
LogMVE	0.8604 (9.11)***	0.9044 (10.91)***	0.8701 (13.37)***	0.7335 (9.51)***	1.6427 (7.12)***	1.3421 (5.21)***	0.8673 (15.30)***	0.7856 (12.54)***
Leverage	0.0059 (5.51)***	0.0063 (5.82)***	0.0043 (2.13)**	0.0041 (1.96)*	-0.1240 (-4.97)***	-0.1838 (-7.57)***	0.0052 (4.65)***	0.0051 (4.38)***
Market-to-book	-0.1941 (-1.08)	-0.2323 (-1.47)	-0.2207 (-1.78)*	-0.1992 (-1.66)*	-0.8338 (-0.76)	-0.7266 (-0.81)	-0.2022 (-1.93)*	-0.1841 (-1.84)*
ROA	1.4371 (0.20)	4.1411 (0.73)	-1.1703 (-6.08)***	-1.0966 (-5.70)***	-2.5340 (-1.69)*	-2.1587 (-1.73)*	-1.1439 (-5.40)***	-1.0754 (-4.92)***
Retsd	0.3661 (6.42)***	0.3761 (7.59)***	0.2541 (5.24)***	0.2197 (4.34)***	0.0449 (0.07)	0.2916 (1.13)	0.2790 (7.23)***	0.2604 (6.89)***
LogEmpl	0.7598 (2.92)***	0.7011 (2.90)***	0.1255 (1.88)*	0.2210 (2.94)***	-0.3606 (-1.93)*	-0.2493 (-1.02)	0.2301 (2.76)***	0.2881 (3.41)***
LogEmpl^2	-0.0966 (-2.30)**	-0.0921 (-2.22)**	0.0079 (0.60)	0.0075 (0.50)	0.0030 (0.08)	-0.0021 (-0.04)	-0.0093 (-0.60)	-0.0101 (-0.60)
LogSpread		-0.1309 (-0.34)		-0.0127 (-0.04)		4.2410 (2.91)***		-0.0841 (-0.32)
Unemplr		-6.4116 (-0.50)		-20.6904 (-1.93)*		-324.33 (-2.59)***		-15.0667 (-1.64)
Tbill3mr		0.0041 (0.00)		-0.6740 (-0.75)		-25.1386 (-3.38)***		-0.3974 (-0.51)
S&P1mr		0.2294 (1.65)*		-0.1516 (-1.09)		-3.0309 (-4.01)***		-0.0155 (-0.14)
S&P1mrsd		-0.3624 (-0.39)		-0.2723 (-0.34)		-21.2640 (-2.88)***		-0.2242 (-0.32)
GDPgr		-3.1675 (-2.13)**		-2.2234 (-1.99)**		-62.7461 (-5.83)***		-2.5508 (-2.64)***
Num. Obs.	157,350	157,118	157,350	157,118	157,350	157,118	157,350	157,118
χ^2 macro		7 (0.3197)		14 (0.0341)**		182 (0.0000)***		13 (0.0510)*
Pseudo R^2	0.3592	0.3684	0.4358	0.4403	0.2850	0.5600	0.4387	0.4442

Table 10. Firm-level χ^2 test of the conditional Poisson assumption of loss arrivals.
The null hypothesis is that the conditional Poisson frequency assumption is valid. For each firm, the test statistic is calculated as $W^2 = \sum_{k=1}^{K} \frac{(n_k - c)^2}{c}$ that has a limiting χ^2 distribution with $K-1$ degrees of freedom, where K is the total number of bins of size c. We opt for bin sizes 1, 2, 3, 4, 5, and 8. A low *p*-value, defined as $\Pr\left(\chi^2 \geq W^2\right)$, suggests rejection of the null hypothesis. The test is conducted after a time-scale transformation is applied to the monthly aggregated frequency data of each firm, using the frequency estimates produced from the Poisson regression analysis. A separate test is performed for each event type model. The reported figures represent proportions of firms for which the *p*-values are less than or equal to $\alpha = \{0.10, 0.05, 0.01\}$. Larger proportions indicate widespread violations of the null hypothesis. N/A indicates scenarios in which the number of observations is insufficient to perform a test with the specified bin size.

Bin size	α	Model 1 Internal Fraud	Model 2 Fraud	Model 3 All Except Internal Fraud	Model 4 All Except Fraud	Model 5 Internal Events	Model 6 External Events	Model 7 Physical Disasters	Model 8 All Events
1	*0.10*	0.2424	0.3250	0.2987	0.2532	0.2391	0.2346	0	0.2785
	0.05	0.2424	0.2750	0.2468	0.2278	0.2391	0.1852	0	0.2658
	0.01	0.2121	0.2250	0.2208	0.2025	0.1739	0.1605	0	0.1899
2	*0.10*	0.3158	0.4333	0.4194	0.3226	0.3415	0.4098	0	0.3846
	0.05	0.2632	0.3667	0.2742	0.2581	0.2927	0.2623	0	0.2923
	0.01	0.2632	0.3000	0.2097	0.1935	0.2195	0.1967	0	0.2462
3	*0.10*	0.6666	0.5909	0.5000	0.5294	0.4138	0.5400	N/A	0.4821
	0.05	0.4444	0.5000	0.4423	0.4118	0.2069	0.3600	N/A	0.3929
	0.01	0.3333	0.2727	0.2500	0.2353	0.1724	0.1800	N/A	0.2500
4	*0.10*	0.6000	0.4667	0.5000	0.5625	0.5455	0.6222	N/A	0.6275
	0.05	0.6000	0.4000	0.4800	0.4375	0.4091	0.4889	N/A	0.4706
	0.01	0.4000	0.2667	0.2600	0.3333	0.2273	0.2444	N/A	0.2745
5	*0.10*	1.0000	0.6667	0.6739	0.6000	0.6250	0.6667	N/A	0.5833
	0.05	0.5000	0.5333	0.5217	0.4000	0.4375	0.5385	N/A	0.5000
	0.01	0.5000	0.2667	0.3913	0.2889	0.1875	0.3333	N/A	0.3333
8	*0.10*	N/A	0.8000	0.6970	0.7500	0.7500	0.8333	N/A	0.7297
	0.05	N/A	0.8000	0.5455	0.6071	0.5000	0.7083	N/A	0.6486
	0.01	N/A	0.6000	0.3939	0.4286	0.5000	0.4583	N/A	0.4324

Table 11. χ^2 test of the joint conditional Poisson assumption of loss arrivals.
The null hypothesis is that the conditional Poisson frequency assumption is valid. For each firm, the test statistic is calculated as $W^2 = \sum_{k=1}^{K} \frac{(n_k - c)^2}{c}$ that has a limiting χ^2 distribution with $K-1$ degrees of freedom, where K is the total number of bins of size c. We opt for bin sizes 1, 2, 3, 4, 5, and 8. A low p-value, defined as $\Pr(\chi^2 \geq W^2)$, suggests rejection of the null hypothesis. The test is conducted after a time-scale transformation is applied to the monthly aggregated frequency data across all surviving firms, using the sum of the frequency estimates produced from the Poisson regression analysis. A separate test is performed for each event type model. The table reports the p-values associated with the χ^2 test. Panel A considers all firms (no restriction) regardless of whether the conditional Poisson assumption is violated or not at the firm level (Table 9). Panels B, C, and D consider only those firms whose p-value in the firm-level conditional Poisson test is greater than α, i.e., no violation of the conditional Poisson assumption at the firm-level at significance level α: Panel B uses α = 0.01 (least restrictive scenario), Panel C uses α = 0.05 (moderately restrictive scenario), and Panel D uses α = 0.10 (most restrictive scenario). Furthermore, within each case we use 3 further restriction levels according to which firms are selected for the joint test: firms whose individual p-values are greater than α for a bin size of 1 (least restrictive), bin sizes of 1 and 2 (moderately restrictive), and bin sizes of 1, 2, and 3 (most restrictive). ***, **, * indicate violation of the joint conditional Poisson assumption at the 1, 5, and 10 percent levels, respectively. N/A indicates scenarios in which the number of observations is insufficient to perform a test with the specified bin size.

Panel A: All firms

Bin size	Model 1 Internal Fraud	Model 2 Fraud	Model 3 All Except Internal Fraud	Model 4 All Except Fraud	Model 5 Internal Events	Model 6 External Events	Model 7 Physical Disasters	Model 8 All Events
1	0***	0***	0.0029***	0***	0***	0***	0***	0.4172
2	0***	0***	0***	0***	0***	0***	0***	0***
3	0***	0***	0***	0***	0***	0***	0***	0***
4	0***	0***	0***	0***	0***	0***	0***	0***
5	0***	0***	0***	0***	0***	0***	0.0001***	0***
8	0***	0***	0***	0***	0***	0***	0.0004**	0***

Panel B: Least restrictive scenario (firms with p-value > 0.01 for specified bin sizes)

Bin size	Bin sizes for which p-value >0.01 in individual Poisson test	Model 1 Internal Fraud	Model 2 Fraud	Model 3 All Except Internal Fraud	Model 4 All Except Fraud	Model 5 Internal Events	Model 6 External Events	Model 7 Physical Disasters	Model 8 All Events
1	*1*	0.9472	0.7709	0.9983	0.7241	0.8603	0.9985	0.4451	0.9813
	1 ,2	0.5892	0.7761	0.9325	0.9809	0.6419	0.9990	0.5712	0.9767
	1, 2, 3	0.9126	0.7892	0.9936	0.9853	0.8719	0.9679	N/A	0.9993
2	*1*	0.2478	0.0363**	0.6251	0.3540	0.1012	0.5910	0.1159	0.4582
	1 ,2	0.5013	0.2708	0.8386	0.5072	0.0421**	0.6237	0.2722	0.7462
	1, 2, 3	0.4168	0.1827	0.6931	0.7468	0.1231	0.2463	N/A	0.8144
3	*1*	0.3135	0.0628*	0.2883	0.0495**	0.0578*	0.3249	0.0533*	0.1335
	1 ,2	0.0868*	0.0582*	0.3648	0.4259	0.0025***	0.2839	0.0960*	0.4994
	1, 2, 3	0.5271	0.0674*	0.3864	0.1053	0.0893*	0.2518	N/A	0.4266
4	*1*	0.0142**	0.0003***	0.3529	0.4472	0.0008***	0.3011	0.1163	0.2986
	1 ,2	0.1833	0.0201**	0.2869	0.0408**	0.0171**	0.2681	0.1243	0.3170
	1, 2, 3	0.6947	0.0474**	0.3720	0.3542	0.0623*	0.2553	N/A	0.2483
5	*1*	0.0584*	0.0102**	0.2010	0.0353**	0.0001***	0.2266	0.3872	0.1643
	1 ,2	0.3013	0.0082***	0.2606	0.1688	0.0038***	0.1852	N/A	0.3172
	1, 2, 3	0.8207	0.0570*	0.2722	0.1813	0.0549*	0.2484	N/A	0.1828
8	*1*	0.0313**	0.0040***	0.0324**	0.2349	0.0671*	0.1279	N/A	0.0867*
	1 ,2	0.0888*	0.0091***	0.1886	0.0368**	0.0028***	0.0259**	N/A	0.0483**
	1, 2, 3	0.6062	0.0050***	0.1454	0.1708	0.0120**	0.1117	N/A	0.2526

Panel C: Moderately restrictive scenario (firms with *p*-value > 0.05 for specified bin sizes)

Bin size	Bin sizes for which p-value >0.05 in individual Poisson test	Model 1 Internal Fraud	Model 2 Fraud	Model 3 All Except Internal Fraud	Model 4 All Except Fraud	Model 5 Internal Events	Model 6 External Events	Model 7 Physical Disasters	Model 8 All Events
1	*1*	0.9494	0.9156	0.9462	0.4140	0.8425	0.9985	0.4451	0.9143
	1 ,2	0.5892	0.5392	0.9864	0.9945	0.6423	0.9877	0.5712	0.9737
	1, 2, 3	0.5832	0.5582	0.7918	0.6335	0.9463	0.9417	N/A	0.6373
2	*1*	0.2414	0.0997*	0.4054	0.0026***	0.1922	0.5977	0.1159	0.4168
	1 ,2	0.5013	0.0189**	0.8441	0.7753	0.0096***	0.4868	0.2722	0.4102
	1, 2, 3	0.3343	0.1985	0.8061	0.0961*	0.6474	0.3050	N/A	0.0758*
3	*1*	0.0533*	0.0087***	0.0820*	0.0001***	0.1741	0.3273	0.0533*	0.0668*
	1 ,2	0.0868*	0.0299**	0.2290	0.1712	0.0797*	0.4926	0.0960*	0.4003
	1, 2, 3	0.6488	0.0873*	0.6750	0.2819	0.2143	0.4977	N/A	0.3533
4	*1*	0.0951*	0.0255**	0.0404**	0.1552	0.0096***	0.3119	0.1163	0.0494**
	1 ,2	0.1833	0.0173**	0.5087	0.3447	0.0153**	0.2583	0.1243	0.2322
	1, 2, 3	0.6318	0.0432**	0.4343	0.2409	0.1203	0.3365	N/A	0.1340
5	*1*	0.0432**	0.0150**	0.1901	0.2506	0.0109**	0.2384	0.3872	0.2312
	1 ,2	0.3013	0.0011***	0.3477	0.2780	0.0188**	0.2141	N/A	0.2692
	1, 2, 3	0.9020	0.0275**	0.5427	0.1531	0.0841*	0.2921	N/A	0.0525*
8	*1*	0.0184**	0.0030***	0.1064	0.0643*	0.0032***	0.1355	N/A	0.0759*
	1 ,2	0.0888*	0.0078***	0.5393	0.1899	0.0049***	0.0804*	N/A	0.2946
	1, 2, 3	0.6305	0.0222**	0.4096	0.1349	0.0243**	0.1809	N/A	0.4847

Panel D: Most restrictive scenario (firms with *p*-value > 0.1 for specified bin sizes)

Bin size	Bin sizes for which p-value >0.1 in individual Poisson test	Model 1 Internal Fraud	Model 2 Fraud	Model 3 All Except Internal Fraud	Model 4 All Except Fraud	Model 5 Internal Events	Model 6 External Events	Model 7 Physical Disasters	Model 8 All Events
1	*1*	0.9494	0.8511	0.9865	0.7990	0.8425	0.9214	0.4451	0.8779
	1 ,2	0.7546	0.8160	0.5259	0.8769	0.9765	0.9884	0.5712	0.7090
	1, 2, 3	0.3156	0.5047	0.5393	0.8991	0.8213	0.7920	N/A	0.9771
2	*1*	0.2414	0.2438	0.8295	0.1163	0.1922	0.4446	0.1159	0.6770
	1 ,2	0.0813*	0.2295	0.0458**	0.2230	0.2608	0.5972	0.2722	0.2917
	1, 2, 3	0.4524	0.2195	0.0847*	0.4338	0.3843	0.5029	N/A	0.5243
3	*1*	0.0533*	0.0440**	0.4268	0.0084***	0.1741	0.5056	0.0533*	0.0229**
	1 ,2	0.1087	0.0652*	0.3086	0.3513	0.0597*	0.6540	0.0960*	0.5213
	1, 2, 3	0.4644	0.1242	0.4920	0.3441	0.2453	0.5077	N/A	0.3000
4	*1*	0.0951*	0.0144**	0.3350	0.3117	0.0096***	0.0288**	0.1163	0.3353
	1 ,2	0.2034	0.0185**	0.2017	0.2688	0.0704*	0.5496	0.1243	0.3076
	1, 2, 3	0.2327	0.0632*	0.7340	0.2510	0.1253	0.3298	N/A	0.3679
5	*1*	0.0432**	0.0041***	0.2514	0.0618*	0.0109**	0.2897	0.3872	0.2346
	1 ,2	0.3087	0.0185**	0.3235	0.2256	0.0100***	0.1795	N/A	0.3904
	1, 2, 3	0.4547	0.2268	0.5302	0.1906	0.1671	0.6988	N/A	0.0820*
8	*1*	0.0184**	0.0016***	0.2432	0.0821*	0.0032***	0.0916*	N/A	0.1204
	1 ,2	0.2814	0.0043***	0.2522	0.1324	0.0148**	0.5275	N/A	0.1065
	1, 2, 3	0.1227	0.0679*	0.7676	0.2209	0.0483**	0.1296	N/A	0.3220

Figure 1. Frequency histogram of event duration for operational risk events, 1980-2005.

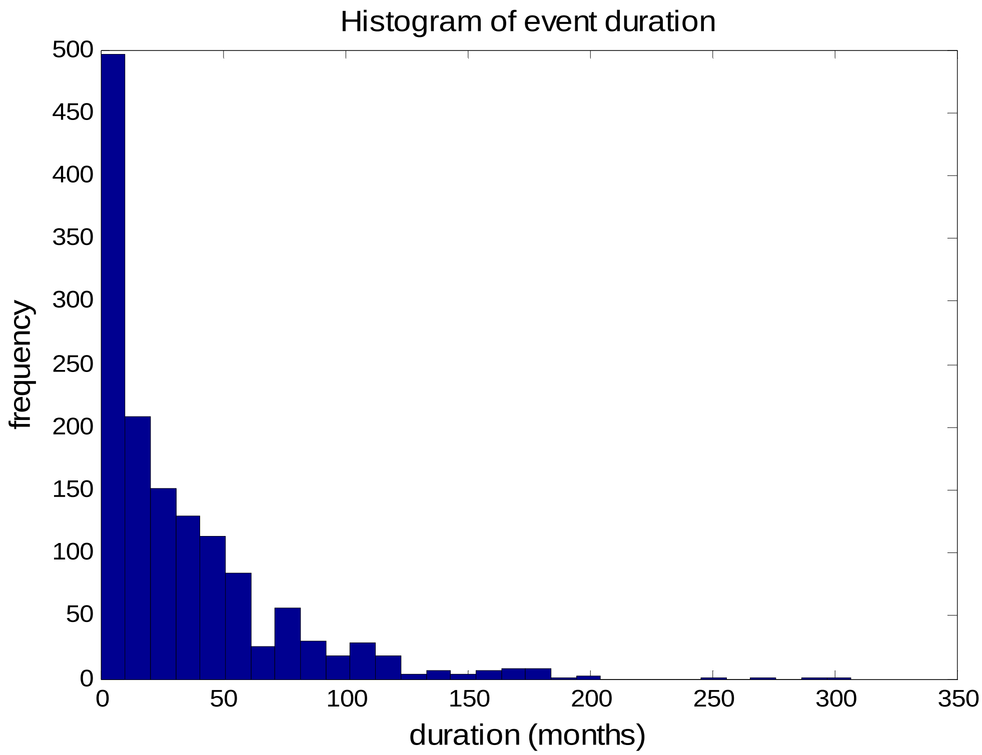
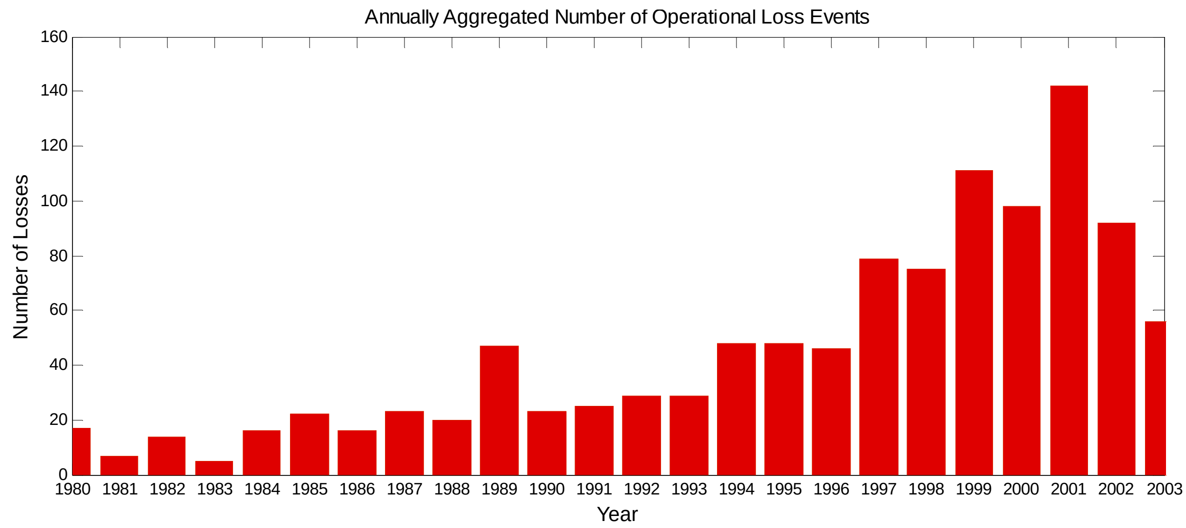
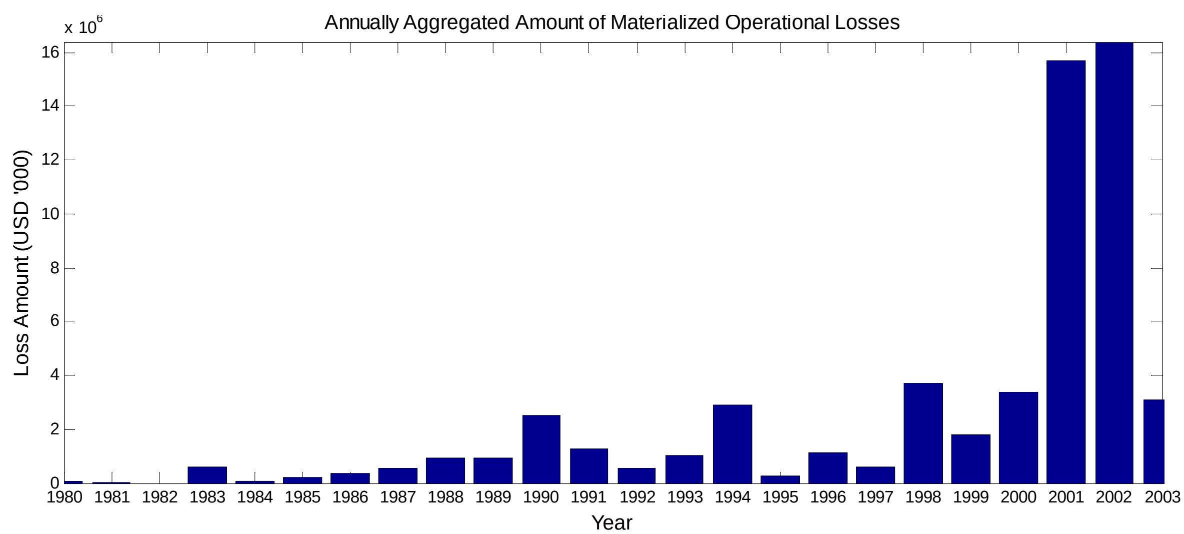


Figure 2. Frequency and severity of annually aggregated operational losses, 1980-2003.

Panel A: Operational loss frequency



Panel B: Optional loss severity



Panel C: Default frequency

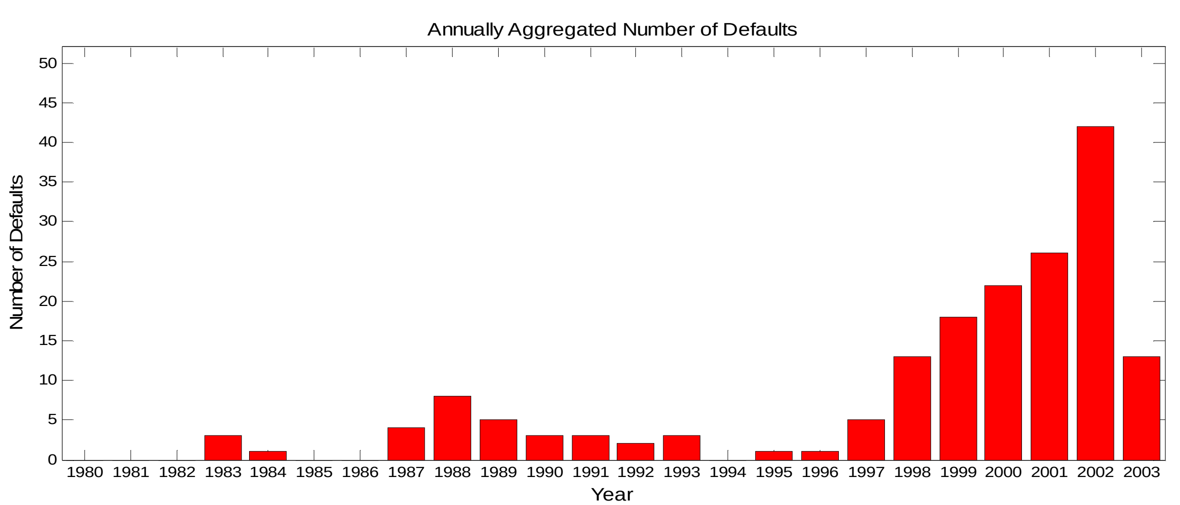
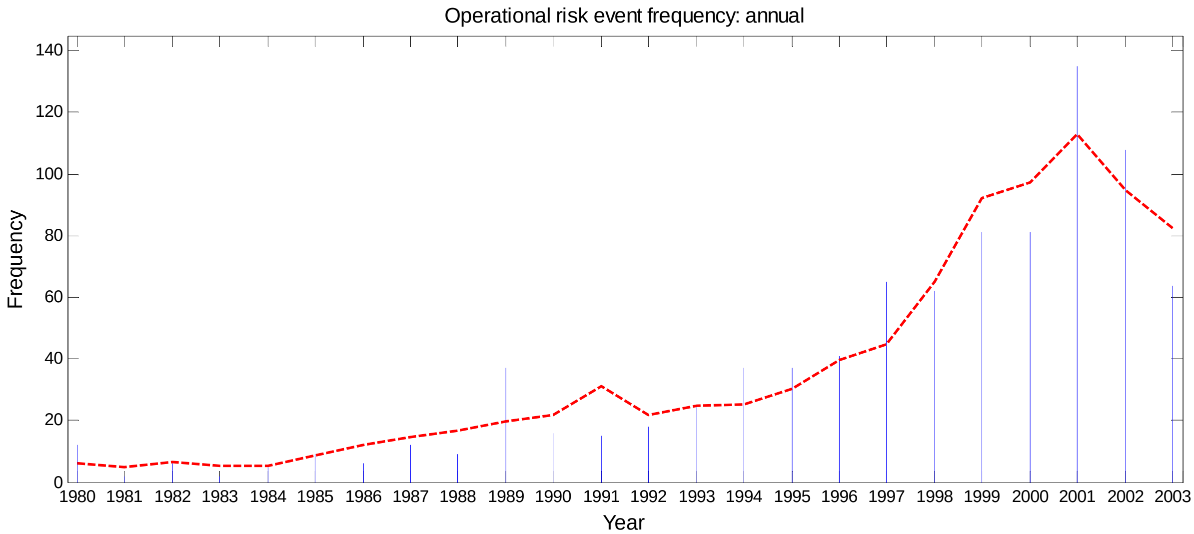


Figure 3. Actual frequency versus predicted frequency of events.
The dashed line represents the predicted frequency and the vertical bars represent the actual frequency of events. Panel A illustrates losses of all types. Panel B focuses on fraud-related events.

Panel A: Model 8 (All Events)



Panel B: Model 2 (Fraud)