Volume Title: The Risks of Financial Institutions

Volume Author/Editor: Mark Carey and René M. Stulz, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-09285-2

Volume URL: http://www.nber.org/books/care06-1

Conference Date: October 22-23, 2004

Publication Date: January 2007

Title: Implications of Alternative Operational Risk Modeling Techniques

Author: Patrick de Fontnouvelle, Eric Rosengren, John Jordan

URL: http://www.nber.org/chapters/c9617

10

Implications of Alternative Operational Risk Modeling Techniques

Patrick de Fontnouvelle, Eric S. Rosengren, and
John S. Jordan

10.1 Introduction

Large operational losses as a result of accounting scandals, insider fraud, and rogue trading, to name just a few, have received increasing attention from the press, the public, and policymakers. The frequency of severe losses, with more than 100 instances of losses at financial institutions exceeding $100 million, has caused many financial institutions to try to explicitly model operational risk to determine their own economic capital. As financial institutions have begun to comprehensively collect loss data and use it to manage operational risk, bank regulators have increased their expectations for measuring and modeling operational risk. Under the current U.S. rules proposal for implementing the Basle Accord, large, internationally active banks will be expected to use internal models to estimate capital for unexpected operational losses. A criticism of this proposal has been that the tools for modeling operational risk are in their infancy, making estimating capital problematic.

This paper uses data supplied by six large, internationally active banks to determine if the regularities in the loss data will make consistent modeling of operational losses possible. We find that there are similarities in the results of models of operational loss across institutions, and that our results are consistent with publicly reported operational risk capital estimates produced by banks' internal economic capital models.

This paper was prepared for the NBER Project on the Risks of Financial Institutions. It was substantially completed while John Jordan was with the Federal Reserve Bank of Boston. We thank our colleagues in the Federal Reserve System and in the Risk Management Group of the Basel Committee for the many fruitful interactions that have contributed to this work. However, the views expressed in this paper do not necessarily reflect their views, those of the Federal Reserve Bank of Boston, or those of the Federal Reserve System.

We begin the analysis by considering tail plots of each bank's loss data by business line and event type. Three findings clearly emerge from this descriptive analysis.[1] First, loss data for most business lines and event types may be well modeled by a Pareto-type distribution, as most of the tail plots are linear when viewed on a log-log scale. Second, the severity ranking of event types is consistent across institutions. Clients, products, and business practices is the highest severity event type, while external fraud and employment practices are the lowest severity event types. Third, the tail plots suggest that losses for certain business lines and event types are very heavy tailed. This last finding highlights that while basic measurement approaches such as the tail plot are easy to implement and are intuitively appealing, overly simplistic approaches may yield implausible estimates of economic capital. A main contribution of this paper is to show how quantitative modeling can result in more plausible conclusions regarding tail thickness and economic capital.

We next attempt to model the distribution of loss amounts using a "full-data" approach, whereby one fits all of the available loss data with a parametric severity distribution. We consider nine commonly used distributions, four of which are light-tailed and five of which are heavy-tailed. We fit each of these distributions by business line and event type at each of the six institutions considered. The heavy-tailed distributions provide consistently good fits to the loss data, which confirms our findings based on visual inspection of the tail plots. The light-tailed distributions do not generally provide good fits. However, we find that some parameter estimates for the heavy-tailed distributions can have implausible implications for both tail thickness and economic capital.

Extreme Value Theory (EVT) is an alternative to the full-data approach that is increasingly being explored by researchers, by financial institutions, and by their regulators. However, it is well-known that EVT techniques yield upward-biased tail estimates in small samples. Huisman, Koedijk, Kool, and Palm (2001) have proposed a regression-based EVT technique that corrects for small-sample bias in the tail parameter estimate. Applying their technique (hereafter HKKP) to the six banks in our sample, we obtain estimates that are both plausible and consistent with earlier estimates using purely external data (de Fontnouvelle et al., 2003).

It is important to stress that the statistical analysis of operational loss data is a new field, and that this paper's results should be viewed as preliminary. This is particularly true given that we only have data for one year from each bank. The paper also raises several technical issues that should be addressed in future research as a longer time series becomes available.

1. Suppose one has a series of observations (x_i) with a cumulative empirical distribution function denoted by $F(x)$. A tail plot is obtained by plotting $\log(1 - F[x_i])$ on the vertical axis against $\log(x_i)$ on the horizontal axis.

The most significant such issue is that even though the data appear to be heavy-tailed, we cannot formally reject the hypothesis that they are drawn from a light-tailed distribution, such as the lognormal. To investigate this possibility, we propose a threshold analysis of the lognormal distribution that, to our knowledge, is new to this subject. This technique also provides a reasonable characterization of the tail behavior of operational losses.

We also examine the frequency of operational losses. We consider both the Poisson distribution and the Negative Binomial distribution as potential models for the number of losses that a bank could incur over the course of one year. Using Monte Carlo simulation to combine the frequency and severity distributions, we obtain an estimate for the distribution of total annual operational losses. The quantiles of this aggregate loss distribution are interpreted as economic capital estimates for operational risk. These estimates should be viewed with several significant cautions. First, we are assuming that the data are complete: however, banks have moved to more comprehensive data collection platforms, which may improve the loss capture. Second, we are only using internal data for one year, and banks will be required to have three years of comprehensive data. Third, analysis of internal loss data will not be the sole determinant of capital for operational risk; banks will also be required to demonstrate that their risk estimates reflect exposures that are not captured in internal loss data.[2] Given these qualifications, the estimates should be viewed as a preliminary indication (and most probably a lower bound) for the amount of capital needed.

Despite these caveats, the estimates implied by the modeling of the internal loss data are consistent with capital estimates using purely external data (de Fontnouvelle et al., 2003). The results imply that for a variety of plausible assumptions regarding the frequency and severity of operational losses, the level of capital needed for operational risk for the typical (median) bank in our sample would be equivalent to 5–9 percent of the bank's current minimum regulatory capital requirement. This range also seems consistent with the 12–15 percent of minimum regulatory capital that most banks are currently allocating to operational risk, given that the banks' models tend to have a broader set of model inputs than those used in this analysis, including external data, scenarios, and qualitative risk assessments.[3] Our results thus confirm that operational risk is a material risk faced by financial institutions.

The remainder of the paper is organized as follows. The next section

2. The proposed Basel Accord requires banks to measure losses to which they are exposed, but that have not actually occurred (via analysis of scenarios and external data). Banks would also be required to measure exposures that have arisen since the data collection period (via analysis of business environment and control factors).

3. See page 26 of Basel Committee on Banking Supervision (2001). Given the uncertainties in evaluating the relative merits of different techniques and estimators using only limited data, we would consider results to be consistent if they are within an order of magnitude of each other.

(10.2) provides a description of the data. Section 10.3 reviews related literature on the measurement of operational risk in financial institutions. Section 10.4 discusses some commonly used continuous distributions, and discusses their potential relevance to modeling the severity of operational losses. Section 10.5 presents visual analyses of the loss data, and draws preliminary conclusions regarding which distributions may be appropriate for modeling loss severity. Section 10.6 explores full-data approaches to modeling operational losses, and formally compares the alternative severity distributions. Section 10.7 explores EVT-based approaches to modeling the loss data. Section 10.8 compares alternative frequency distributions. Section 10.9 provides the implied capital numbers from estimating different loss distributions using Monte Carlo simulations. The final section provides conclusions on using these techniques for quantifying operational risk.

10.2 Data

The 2002 Operational Risk Loss Data Collection Exercise (LDCE) was initiated by the Risk Management Group (RMG) of the Basel Committee on Banking Supervision in June 2002. The LDCE asked participating banks to provide information on individual operational losses exceeding €10,000 during 2001, among various other data items. Banks were also asked to indicate whether their loss data were complete. The LDCE data include 47,269 operational loss events reported by eighty-nine banks from nineteen countries in Europe, North and South America, Asia, and Australasia. For additional information and summary statistics regarding the LDCE, readers can refer to Risk Management Group (2003).

Based on the information provided in the LDCE, and on our knowledge of the banks involved, we identified a list of institutions whose data submissions seem relatively complete. Due to practical considerations, we limit our sample to loss data from six of these banks. This paper presents results for these six banks on a bank-by-bank basis (with the exception of the operational risk exposure figures reported in table 10.5). However, the results are presented in a way that makes it impossible to identify the individual banks. Focusing on a cross-sectional study of banks enables us to determine whether the same statistical techniques and distributions apply across institutions that may have very different business mixes and risk exposures.

The LDCE categorizes losses into eight business lines and seven event types. To protect the confidentiality of banks participating in the LDCE, we present results only for those business lines and event types when three or more banks reported sufficient data to support analysis. The business lines presented are: trading and sales, retail banking, payment and settlement, and asset management. The loss types presented are: internal fraud, external fraud, employment practices and workplace safety, clients, products, and business practices, and execution, delivery, and process management.[4]

10.3 Related Literature

Moscadelli (2003) also analyzes data from the 2002 LDCE, and performs a thorough comparison of traditional full-data analyses and extreme value methods for estimating the operational loss severity distribution. He finds that extreme value theory outperforms the traditional methods in all eight Basel Business Lines. He also finds that the severity distribution is very heavy-tailed, and that there is a substantial difference in loss severity across business lines.

There are several differences between the current paper and Moscadelli (2003). First, Moscadelli (2003) aggregates the data across all banks in the LDCE sample. In this paper, we analyze data at the individual bank level in order to determine whether the same quantitative techniques work for a variety of banks with different business mixes, control infrastructures, and geographic exposures. We believe that doing so provides a useful test of the techniques under consideration, and also yields an indication of their ultimate applicability at individual banks. Second, the current paper explores the newly developed technique of Huisman et al. (2001) to correct for potential bias in the tail parameter estimate. Third, we explore several models of the loss frequency distribution, which allows us to obtain indicative estimates of economic capital for operational risk.

10.4 Distributions for Operational Loss Data

We begin our empirical analysis by exploring which of various empirical approaches best fits the data. In principle, we are willing to consider any distribution with positive support as an acceptable candidate for modeling operational loss severity. To keep the size of our tables within reason, however, we will focus on nine commonly used distributions. This section discusses the salient features of each. In section 10.6, we consider how well these distributions describe the statistical behavior of losses in our database.

Table 10.1 lists each distribution we consider, together with its density function and its maximal moment (discussed at the end of this section). We begin our discussion with the exponential distribution, which is one of the simplest statistical distributions—both analytically and computationally. The exponential distribution is frequently used to analyze duration data

4. The following business lines were omitted: corporate finance, commercial banking, agency services, and retail brokerage. The following event types were omitted: damage to physical assets, business disruption, and system failure. To preserve confidentiality, we do not report the cutoff that was used for inclusion of business lines and event types.

Table 10.1 Parametric distributions used for modeling operational loss severity

Distribution name	Density, $f(x)$	Maximal moment
Exponential	$(1/b)\exp(-x/b)$	∞
Weibull	$(\beta x^{\beta-1}/\eta^{\beta})\exp(-(x/\eta)^{\beta})$	∞
Gamma	$(x/b)^{c-1}[\exp(-x/b)]/[b\Gamma(c)]$	∞
LogGamma	$[\log(x)/b]^{c-1}x^{-1/b-1}/[b\Gamma(c)]$	$1/b$
Pareto	$\xi^{-1}x^{-1/\xi-1}$	$1/\xi$
GPD	$\beta^{-1}(1+\xi x/\beta)^{-1/\xi-1}$	$1/\xi$
Burr	$(\tau/\beta)x^{\tau-1}(1+\xi x^{\tau}/\beta)^{-1/\xi-1}$	τ/ξ
Lognormal	$(2\pi x^2\sigma^2)^{-1/2}\exp\{-[\log(x)-\mu]^2/(2\sigma^2)\}$	∞
LogLogistic	$\alpha x^{1/b-1}/[b(1+\alpha x^{1/b})^2]$	$1/b$

(e.g., time to failure of a machine part), and is the only continuous distribution characterized by a "lack of memory." In the duration context, lack of memory means that the time until the occurrence of an event (failure) does not depend on the length of time that has already elapsed (time since installation). In the operational loss context, lack of memory implies that the distribution of excess losses over a threshold does not depend on the value of the threshold. So if half of all losses exceeding \$1 are less than \$10, then half of all losses exceeding \$1 million will be less than \$1,000,010 (\$1,000,000 + \$10). Such a result does not seem plausible. However, the exponential distribution arises in the context of EVT as a possible limiting distribution for excess losses above high thresholds. For this reason (and also because it can be transformed into other interesting distributions), we include it in our analysis.

The Weibull distribution is a two-parameter generalization of the exponential that allows the time-until-event occurrence to depend on the amount of time that has already elapsed. Thus, the Weibull can capture phenomena such as "burn in," in which the failure rate is initially high but decreases over time. In the context of operational risk, the Weibull may be appropriate for modeling a business line exposed to many small losses but only a few large losses. The gamma distribution is another two-parameter generalization of the exponential. A gamma-distributed random variable arises as the sum of n exponentially distributed random variables. Thus, a machine's failure time is gamma distributed if the machine fails whenever n components fail, and if each component's failure time is exponentially distributed. Like the Weibull distribution, the gamma also allows the time until event occurrence to depend on the amount of time that has already elapsed.

Another generalization of the exponential distribution can be obtained by exponentiating an exponentially distributed random variable. The resulting distribution is called a Type I Pareto, and can also be referred to as a log-exponential or power-law distribution. The lack of memory of the exponential distribution manifests itself as scale invariance in the Pareto distribution. Roughly speaking, scale invariance means that data "look the same" no matter what the unit of measure (e.g., hundreds of dollars versus millions of dollars). So in the earlier example, where half of all losses exceeding $1 were less than $10, half of all losses over $1 million would be less than $10 million. Power-law behavior has been observed in phenomena as disparate as city sizes, income distributions, and insurance claim amounts, and has been an important research topic for those interested in the behavior of complex systems (i.e., systems consisting of agents linked via a decentralized network rather than via a market or social planner).[5] A variation of the Pareto distribution can be obtained by exponentiating a gamma-distributed random variable instead of an exponentially distributed random variable. The result is referred to as the Loggamma distribution.

The Pareto distribution also arises in EVT as another limiting distribution of excesses over a high threshold. In this case, the limiting distribution is given by a two-parameter variant of the Pareto, which is known as the Generalized Pareto Distribution (GPD). One commonly used transformation of the GPD is obtained by raising a GPD-distributed variable to a power. The result is called the Burr distribution.

Another distribution that we consider is the Lognormal, which is so widely used that little discussion is required here. However, it is worth noting that the normal distribution is appropriate for modeling variables that arise as the sum of many different components. It is also a worthwhile exercise to consider which types of operational losses may be characterized in this manner. Consider, for example, losses arising from workplace safety lapses. One could argue that the severity of these losses may be approximated by the lognormal distribution, as it is influenced by many factors, including weather, overall health of the injured party, physical layout of the workplace, and the type of activity involved. The final distribution that we consider is the loglogistic, which is obtained by exponentiating a logistic-distributed random variable. The Loglogistic is similar to the Lognormal, but may be more appropriate for modeling operational loss data because it has a slightly heavier tail.

We conclude this section by classifying the distributions discussed previously according to their tail thickness. This will facilitate interpretation of the estimation results, as the relevance of a particular distribution to modeling operational losses will be suggestive of the relevance of other distributions with similar tail thickness. There is no commonly agreed-upon definition of what constitutes a heavy-tailed distribution. However, one

5. See Embrechts, Klüppelberg, and Mikosch (1997), Gabaix (1999), and references therein.

such definition can be based on a distribution's maximal moment, which is defined as $\sup(r : E[x^r] < \infty)$. Maximal moments for the distributions under consideration are reported in table 10.1. In this paper, we will call a distribution light tailed if it has finite moments of all orders, and heavy tailed otherwise. Under this definition, four of the distributions being considered are light tailed (exponential, Weibull, gamma, and lognormal), and the remaining five distributions are heavy tailed (loggamma, Pareto, GPD, Burr, and loglogistic).

10.5 Descriptive Analysis

This section considers several tools that provide a visual characterization of the loss data. Suppose one has a series of observations (x_i) with a cumulative empirical distribution function denoted by $F(x)$. A tail plot is obtained by plotting $\log(1 - F[x_i])$ on the vertical axis against $\log(x_i)$ on the horizontal axis. Figures 10.1 and 10.2 present tail plots of the six banks' loss data by Basel event type and Basel business line, respectively.

Many of the tail plots show linear behavior. This is quite interesting, as a linear tail plot implies that the data are drawn from a power-law distribution. Furthermore, the slope of the plot provides a heuristic estimate of the tail parameter, as $\log(1 - F[x_i]) = -a \log(x_i) + c$, where c denotes a constant.

Another feature of these plots is that the slopes associated with the seven Basel event types preserve roughly the same ordering across banks. For example, client's products and business practices is one of the heaviest-tailed event types for all of the banks where it is plotted separately. Employment practices and workplace safety is always one of the thinnest-tailed event types. While the tail plots by business line also suggest power-law tail behavior, there is no evident consistent cross-bank ordering of business lines. We interpret this as initial evidence that risk may be better ordered by event type, but will revisit this issue later in the paper.

Each of the tail plots also indicates a reference line with slope of –1. Many of the plots lie near or above this line, thus implying heuristic tail-parameter estimates of 1 or higher. These estimates highlight the shortcomings of using an overly simplistic approach to measuring operational risk: tail parameters exceeding 1 suggest that the expected loss is infinite for many business lines and event types, and that the capital required for operational risk alone could exceed the amount of capital that large banks are currently allocating to all risks.[6] We will argue in this paper that the distribution of operational losses is not as heavy tailed as it first appears, and

6. The LDCE data suggest that a \$100 billion bank could experience 500 operational losses (exceeding \$10,000) per year. If these follow a Pareto distribution with a tail parameter equal to 1, then Monte Carlo simulation of the aggregate loss distribution indicates capital of \$5 billion at the 99.9 percent soundness level. Tail parameters of greater than 1 would imply capital levels several times larger than this figure.

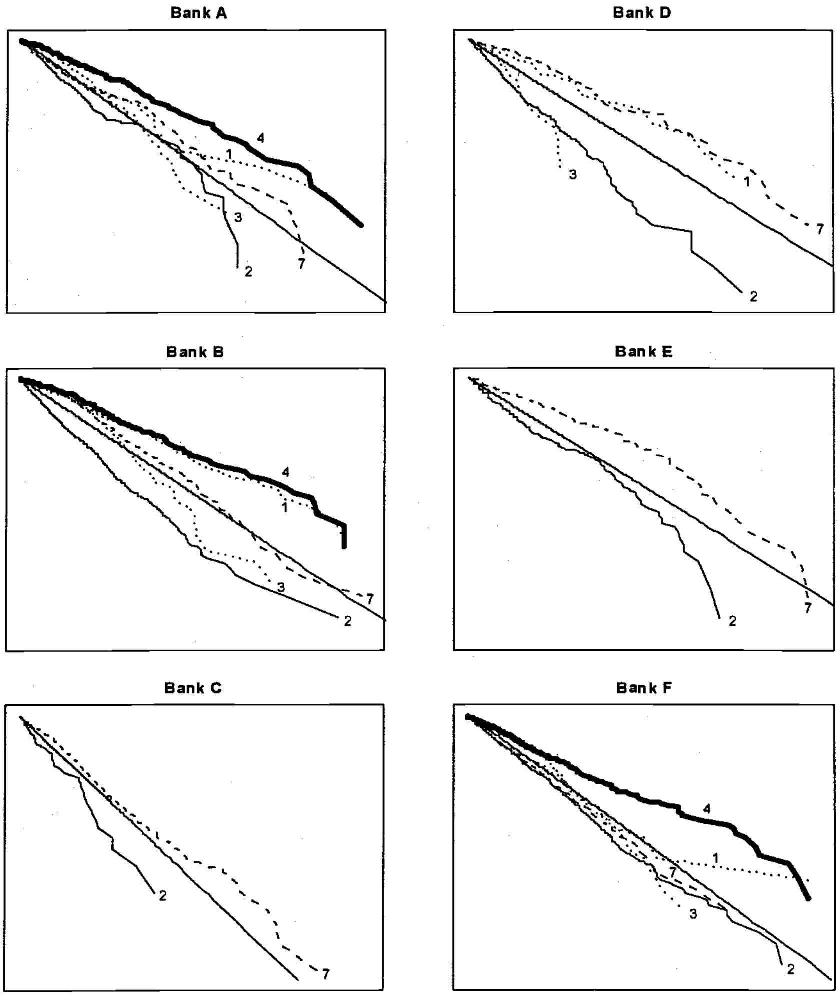

Fig. 10.1 Tail plots of loss data by Basel Event Type

Notes: Event Types are labeled as follows. 1—Internal Fraud. 2—External Fraud. 3—Employment Practices and Workplace Safety. 4—Clients, Products, and Business Practices. 7—Execution, Delivery, and Process Management.

that it is possible to obtain more plausible estimates of regulatory capital for operational risk.

Another useful diagnostic tool is the mean excess plot. The mean excess for a given threshold is defined as the average of all losses exceeding the threshold, minus the threshold value. The mean excess plot reports the mean excess as a function of the threshold value. The shape of the mean excess plot varies according to the type of distribution underlying the data.

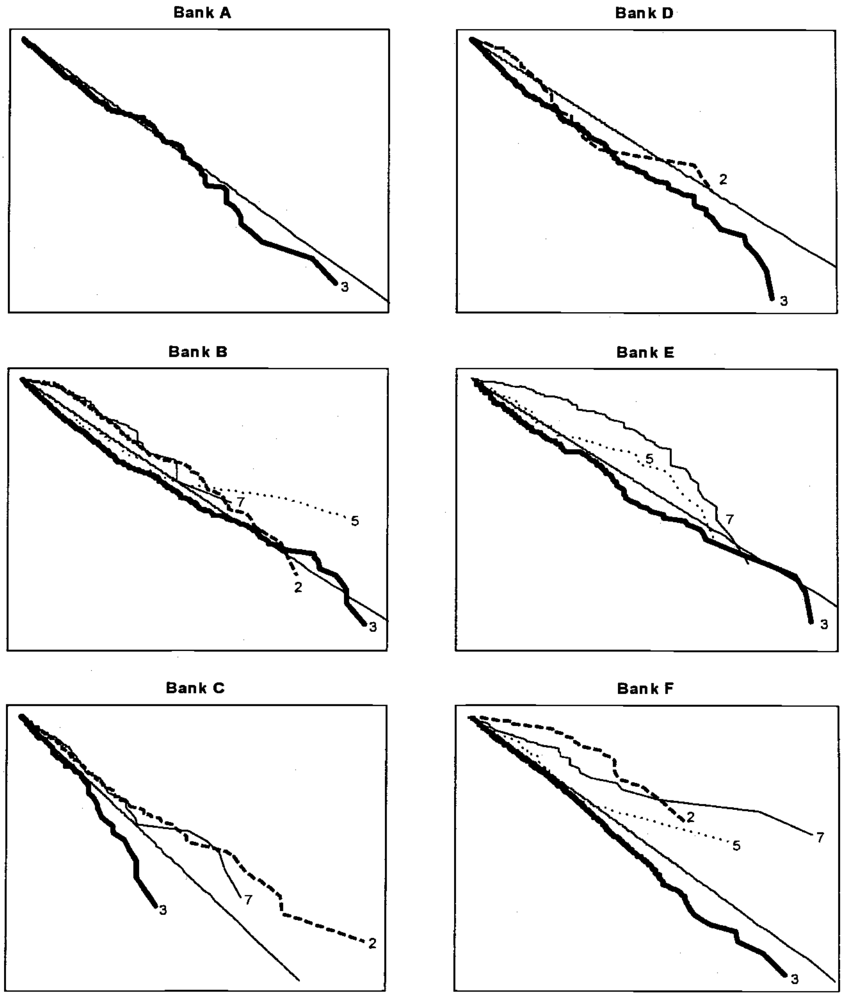

Fig. 10.2 Tail Plots of loss data by Basel Business Line

Notes: Business lines are labels as follows. 2—Trading and Sales. 3—Retail Banking. 5—Payment and Settlement. 7—Asset Management.

For example, a Pareto distribution implies a linear, upward-sloping mean excess plot; an exponential distribution implies a horizontal linear mean excess plot, and a lognormal distribution implies a concave, upward-sloping mean excess plot.

Figures 10.3 and 10.4 present mean excess plots for loss data by event type and business line, respectively. (Each curve has been rescaled in order to display the different business lines and event types together on one plot. Thus, these plots cannot be used to risk-rank business lines or event types.)

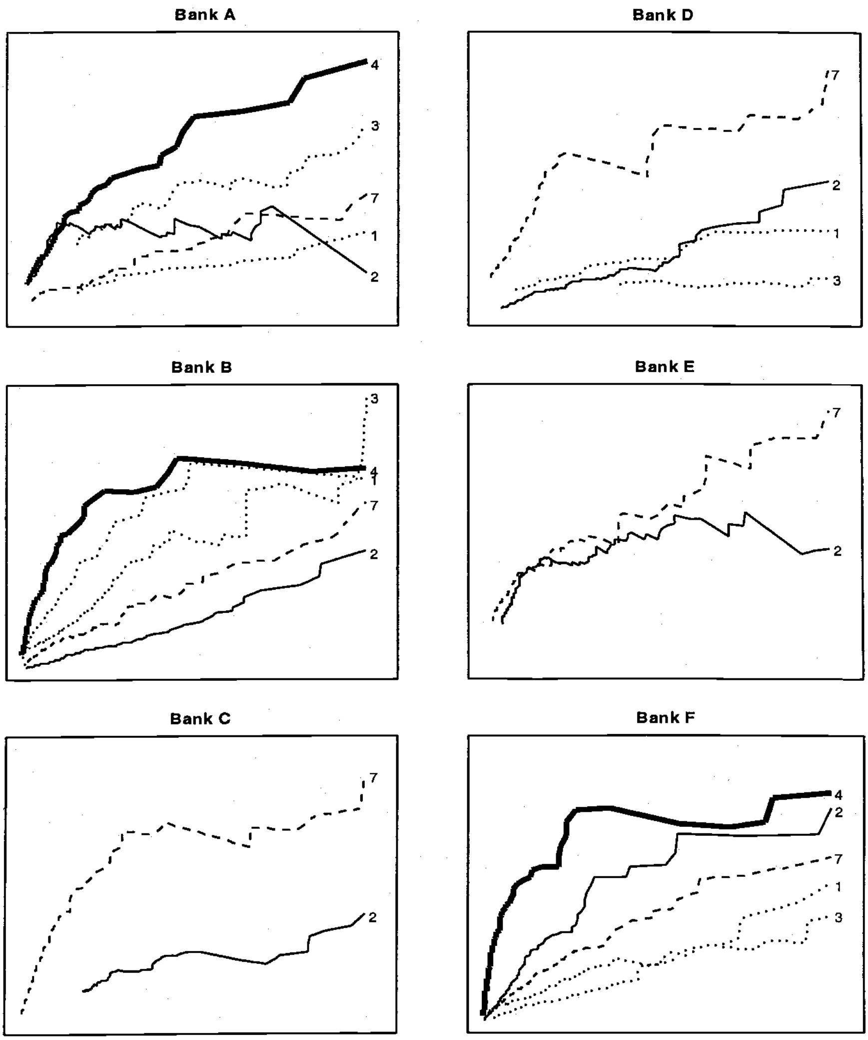

Fig. 10.3 Mean excess plots by Basel Event Type
Note: See fig. 10.1.

Nearly all of the plots slope upward, which indicates tails that are heavier than exponential. Some of the plots are linear (e.g., event type 7 for bank B), which suggests a Pareto-like distribution. Some are concave, which suggests a lognormal or Weibull-like distribution. It is also difficult to establish a consistent pattern across either business line or event type. Potentially, this issue would be less severe with more data.

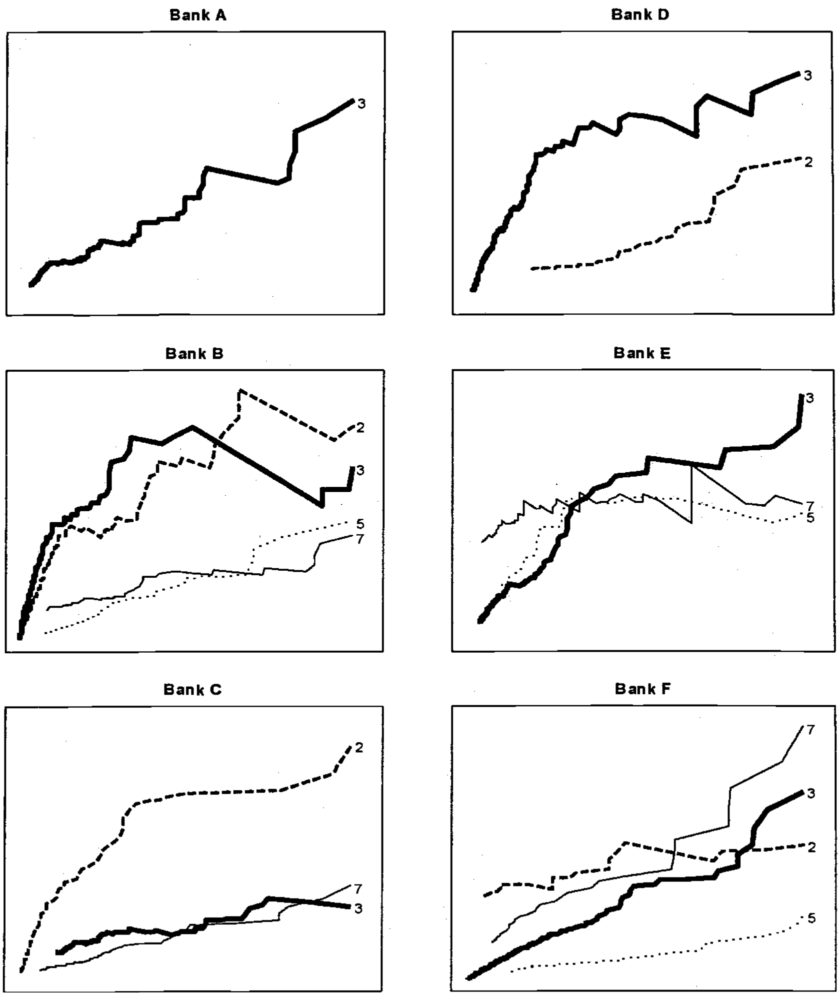

Fig. 10.4 Mean excess plots by Basel Business Line
Note: See fig. 10.2.

10.6 Fitting the Distributions

In this section, we fit each of the distributions listed in table 10.1 to the LDCE data via maximum likelihood. Results are reported separately for each bank under consideration, and are also broken down by business line and event type.

Table 10.2 reports probability values for Pearson's χ^2 goodness-of-fit

Table 10.2 **Goodness of fit across Basel Business Lines and Event Types (%)**

Distribution	Bank A	Bank B	Bank C	Bank D	Bank E	Bank F
			All observations			
Burr	0.0	6.4	72.0	23.3	13.6	0.1
Exponential	0.0	0.0	0.0	0.0	0.0	0.0
Gamma	0.0	0.0	0.0	0.0	1.5	0.0
LogGamma	0.0	1.5	64.1	33.9	1.4	0.7
LogLogistic	0.0	6.4	79.4	23.8	2.1	0.7
Lognormal	0.0	0.2	51.8	0.0	3.5	0.2
GPD	0.0	4.5	75.8	25.7	1.6	0.8
Weibull	0.0	0.0	0.0	0.0	54.7	0.0
			Event type 1—Internal fraud			
Burr	31.7	86.1		99.1		13.0
Exponential	0.0	0.0		0.1		0.0
Gamma	12.2	0.0		73.4		0.0
LogGamma	32.7	85.6		98.1		18.9
LogLogistic	31.8	87.4		98.1		13.6
Lognormal	35.0	86.4		98.4		13.7
GPD	33.1	87.6		95.1		13.3
Weibull	74.9	0.4		40.9		0.0
			Event type 2—External fraud			
Burr	0.0	10.8	6.4	13.2	1.7	0.0
Exponential	0.0	0.0	0.1	0.0	0.0	0.0
Gamma	0.0	0.0	0.7	0.0	1.8	0.0
LogGamma	0.3	6.5	7.6	7.3	2.8	0.0
LogLogistic	0.0	5.1	5.8	9.3	2.7	0.0
Lognormal	0.0	0.0	6.6	4.8	3.0	0.0
GPD	0.0	10.5	6.1	13.9	1.7	0.1
Weibull	0.0	0.0	5.9	0.1	9.9	0.0
			Event type 3—Employment practices and workplace safety			
Burr	87.2	36.7		85.0		23.0
Exponential	1.0	0.0		29.0		0.0
Gamma	66.5	0.0		86.9		0.2
LogGamma	88.1	7.8		74.7		0.2
LogLogistic	91.7	59.5		92.0		24.5
Lognormal	95.5	57.1		86.7		24.8
GPD	92.9	64.4		82.0		35.2
Weibull	87.0	0.2		85.8		7.6
			Event type 4—Clients, products, and business practices			
Burr	98.9	58.0				37.0
Exponential	0.0	0.0				0.0
Gamma	0.6	0.0				0.0
LogGamma	80.1	77.2				42.2
LogLogistic	80.8	58.6				39.0
Lognormal	81.4	36.5				40.3
GPD	76.7	57.4				34.7
Weibull	50.1	0.0				0.0

Table 10.2 (continued)

Distribution	Bank A	Bank B	Bank C	Bank D	Bank E	Bank F
Event type 7—Execution, delivery, and process management						
Burr	3.0	1.1	78.9	24.7	78.1	72.6
Exponential	0.0	0.0	0.0	0.0	0.0	0.0
Gamma	0.2	0.0	0.0	0.0	19.8	0.0
LogGamma	0.8	0.4	54.7	22.3	26.6	67.1
LogLogistic	0.1	0.0	76.8	47.7	89.4	77.3
Lognormal	0.1	0.0	51.7	52.1	68.7	0.0
GPD	2.6	0.0	77.6	39.8	83.6	89.6
Weibull	12.0	0.0	0.0	0.6	47.1	7.1
Business line 2—Trading and sales						
Burr		1.6	68.6	88.1		58.4
Exponential		0.0	0.0	1.7		12.4
Gamma		0.0	0.0	1.1		27.4
LogGamma		0.0	65.1	70.6		42.1
LogLogistic		0.0	69.7	65.3		91.8
Lognormal		0.0	67.0	18.8		86.9
GPD		0.0	70.6	25.1		58.0
Weibull		0.0	0.0	2.3		18.3
Business line 3—Retail banking						
Burr	0.1	12.5	32.3	8.5	0.9	1.7
Exponential	0.0	0.0	0.3	0.0	0.0	0.0
Gamma	0.0	0.0	8.1	0.0	0.1	0.0
LogGamma	0.0	0.2	43.0	1.3	5.8	2.4
LogLogistic	0.0	0.0	35.2	0.2	5.6	3.7
Lognormal	0.0	0.0	46.9	0.0	5.5	3.8
GPD	0.1	12.5	32.2	9.0	2.4	4.7
Weibull	0.0	0.0	15.5	0.0	14.7	0.0
Business line 5—Payment and settlement						
Burr		48.5			11.0	69.2
Exponential		0.0			0.0	0.0
Gamma		0.0			7.2	1.7
LogGamma		66.7			40.2	62.0
LogLogistic		49.4			22.7	
Lognormal		63.0			38.5	63.4
GPD		45.3			11.1	66.8
Weibull		0.3			13.3	52.4
Business line 7—Asset management						
Burr		64.9	84.4		30.1	20.2
Exponential		6.4	0.0		3.6	0.0
Gamma		32.3	0.0		43.4	0.0
LogGamma		31.3	79.9		15.9	17.6
LogLogistic		63.1	63.6		44.9	17.4
Lognormal		45.9	62.6		69.8	18.2
GPD		67.5	64.2		61.1	20.8
Weibull		25.7	4.5		44.8	2.3

Notes: This table reports goodness of fit for each of the distributions under consideration. The test was based on a standard chisquare procedure, except for the rounding adjustment discussed in section 10.6. The reported figures are probability values, so that a value of 5 percent or less indicates a poor fit.

statistic.[7] In general, the heavy-tailed distributions (Burr, loggamma, log-logistic, and Pareto) seem to fit the data quite well. The reported probability values exceed 5 percent for many business lines and event types, which suggests that we cannot reject the null that data are in fact drawn from the distribution under consideration. Conversely, most of the light-tailed distributions rarely provide an adequate fit to the data. This is not surprising, as the tail plots suggested that most of the data are heavy tailed. What is somewhat surprising is the degree to which the lognormal distribution fits the data. In fact, this light-tailed distribution fits the loss data for roughly as many business lines and event types as many of the heavier-tailed distributions.

Table 10.3 presents parameter estimation results for the GPD and lognormal distributions. To preserve bank confidentiality, we present only the estimate of the tail parameter ξ for the GPD and only the value of $\mu + \sigma^2/2$ for the lognormal distribution. While the χ^2 statistics presented in table 10.2 suggested that these two distributions provide a reasonable fit to the data, the parameter estimates generally suggest the opposite. Panel A reports estimates of the GPD tail parameter ξ. The parameter estimates are at or above 1 for many business lines and event types, and also above 1 when data is pooled across business lines and event types. Note that a tail parameter of 1 or higher has implausible implications for both expected losses and regulatory capital. Panel B of table 10.3 reports the estimated value of $\mu + \sigma^2/2$ for the lognormal distribution, which enables one to calculate the average loss severity via the formula $\exp(\mu + \sigma^2/2)$. While estimates of the average loss vary by business line and event type, one can see that it is less than exp(0) dollars for multiple business lines and event types. Thus, neither the Pareto nor the lognormal distribution consistently yields plausible parameter estimates.

Because of space considerations, we do not provide parameter estimates for the other distributions that were estimated. However, the GPD is of special interest because of its role in EVT and the lognormal is of special interest because it is the only light-tailed distribution that seems to fit the data (according to the χ^2 test). Parameter estimates for other heavy-tailed distributions were qualitatively similar to those of the GPD, in that they had implausible implications for tail thickness of the aggregate loss distribution.

10.6.1 For Which Business Lines and Event Types Can Full Data be Fit?

In this subsection, we ask whether there seem to be particular event types for which the full-data approach might work. Losses due to employment practices and workplace safety (event type 3) are well fit by most of

7. We calculated χ^2 goodness-of-fit tests because tests based on the empirical distribution function can be sensitive to data rounding, which is prevalent in the LDCE data. One can accommodate rounding within the χ^2 test by choosing bin values appropriately.

Table 10.3 Parameter estimates for the Generalized Pareto and lognormal distributions

	Bank A	Bank B	Bank C	Bank D	Bank E	Bank F
	A. Estimates of the tail parameter ξ for the GPD					
All BL & ET	1.28 (0.08)	0.87 (0.03)	0.99 (0.08)	0.92 (0.07)	0.97 (0.11)	1.01 (0.03)
ET1-IntFrd	1.24 (0.36)	1.31 (0.18)		1.10 (0.38)		1.02 (0.14)
ET2-ExtFrd	1.17 (0.12)	0.79 (0.05)	0.63 (0.19)	0.69 (0.07)	0.86 (0.14)	0.93 (0.03)
ET3-EP&WS	0.50 (0.16)	0.42 (0.05)		–0.15 (0.22)		0.50 (0.06)
ET4-CPBP	1.36 (0.21)	1.25 (0.15)				1.46 (0.13)
ET7-EDPM	1.42 (0.16)	0.71 (0.05)	0.94 (0.08)	1.00 (0.18)	0.96 (0.17)	0.93 (0.09)
BL2-T&S		0.68 (0.06)	1.18 (0.13)	0.49 (0.18)		0.42 (0.28)
BL3-RetBnk	1.15 (0.10)	1.09 (0.05)	0.55 (0.17)	0.94 (0.07)	0.99 (0.14)	0.93 (0.03)
BL5-P&S		1.06 (0.23)			1.07 (0.35)	1.03 (0.29)
BL7-AsstMgt		0.49 (0.20)	0.96 (0.21)		0.37 (0.18)	1.64 (0.40)
	B. Estimates of $\mu + \sigma^2/2$ for the lognormal distribution					
All BL & ET	–6.08	>0	>0	–21.27	>0	–9.23
ET-IntFrd	–9.85	>0		>0		–6.49
ET2-ExtFrd	–8.35	–22.68	>0	–5.84	>0	–22.31
ET3-EP&WS	>0	>0		>0		>0
ET4-CPBP	>0	>0				–5.85
ET7-EDPM	–9.64	>0	>0	>0	>0	–21.51
BL2-T&S		>0	–3.73	>0		>0
BL3-RetBnk	–13.32	–7.45	>0	–5.24	–11.77	–9.61
BL5-P&S		–12.09			>0	–2.01
BL7-AsstMgt		>0	>0		>0	–2.12

the heavy-tailed distributions as well as the lognormal. Furthermore, the parameter estimates for both the GPD and lognormal are plausible. There are two event types (internal fraud, and clients, products, and business practices) where several banks' data are well fit by multiple distributions, but where the resulting parameter estimates are not plausible. External fraud losses are not consistently well fit by any distribution on a cross-bank basis. Results for execution, delivery, and process management are less consistent across banks, with two institutions failing the goodness-of-fit tests, but the others having good fits and (perhaps) plausible parameter estimates.

The results are broadly similar in the case of estimation by business line. There are two business lines (agency services, and asset management) that pass the goodness-of-fit tests, and yield plausible parameter estimates for several banks. Another business line (retail banking) fails the fit tests at most banks, and the final business line (payment and settlement) yields implausible parameter estimates.

10.6.2 What Might Individual Banks Do?

Our discussion to this point has searched for features of operational loss data that hold across all of the six banks in our sample. However, the measurement of operational risk will ultimately take place at individual banks, who may not have the luxury of seeing whether their choices and assumptions are also valid at other institutions. We begin our discussion by focusing on bank F. Bank staff might begin by fitting one statistical distribution across all business lines and event types, but poor goodness-of-fit statistics would quickly lead them to alternate approaches. They might consider fitting a separate loss-severity distribution to each of the seven event types. However, they would find that losses from the most frequent event type (external fraud) were not well modeled by any of the distributions. The next most frequent event type (clients, products, and business practices) is modeled quite well by several heavy-tailed distributions. However, they would be quite surprised to find tail-parameter estimates exceeding 1, and might conclude that this was not a reasonable way to model operational risk. If they next attempted to fit separate loss-severity distributions for each business line, they would discover that loss data for the most common business line (retail banking) were not well-modeled by any of the distributions considered.

Bank F was chosen at random for discussion. If presented with their bank's results from tables 10.2 and 10.3, risk management staff from the other five institutions might reach similar conclusions. They would discover that for many of the important business lines and event types, none of the statistical distributions considered adequately captured the behavior of operational losses. They would also discover that some business lines and event types were well modeled by heavy-tailed distributions, but that the resulting parameter estimates had implausible implications for their overall operational risk exposure.

10.7 Threshold Analysis of Loss Data

The previous section's results suggest that it may be difficult to fit parametric loss-severity distributions over the entire range of loss amounts, even if separate analyses are conducted for each business line and event type. In this section, we focus on the largest losses, as these are most relevant for determining a bank's operational risk exposure. The main theoretical result underlying this "Peaks Over Threshold" (POT) approach is that if the distribution of excess losses converges to a limiting distribution as the threshold increases, then this limiting distribution is either the exponential distribution or the generalized Pareto distribution.

Implementation of the POT approach begins with choosing an estimator for the tail index parameter ξ, the most common being the Hill estimator. The appeal of this estimator derives from its conceptual and computational simplicity. For a set of losses exceeding a given threshold, the Hill estimator equals the average of the log of the losses minus the log of the threshold. If the underlying loss distribution is a Type I Pareto, then the Hill estimator is the maximum likelihood estimate of the tail thickness parameter. This property is quite useful, as it enables one to conduct likelihood ratio tests of various hypotheses.

Let k denote the number of observations exceeding a given threshold value. The quantity k is often referred to as the number of *exceedances.* Figure 10.5 presents plots of the Hill estimator for the six banks under consideration. The solid black line represents the Hill estimator calculated across all business lines and all event types for various values of k between 1 and 200. Traditionally, the final estimate of the tail index parameter has depended heavily on the choice of k. However, Huisman, Koedijk, Kool, and Palm (2001; hereafter HKKP) have recently proposed a regression-based enhancement to the Hill estimator that minimizes the role of threshold selection. HKKP note that the Hill estimator is biased in small samples, and that the bias is approximately linear in k, so that

$$E(\gamma(k)) = \xi + ck, \tag{1}$$

where $\gamma(k)$ denotes the Hill estimator calculated using k exceedances, and ξ denotes the true value of the tail index parameter. HKKP use equation (1) to motivate the following regression

$$\gamma(k) = \beta_0 + \beta_1 k + \varepsilon(k), \tag{2}$$

which is estimated for k in $(1, \ldots, K)$. The estimate of β_0 is interpreted as a bias-corrected estimate of ξ. This method also requires the researcher to choose the number of exceedances to include in the analysis. However, HKKP conclude that the estimate of β_0 is robust to the choice of k.

We apply the HKKP technique to the six Hill plots presented in figure 10.5. The results are presented in table 10.4.

The second column reports the number of exceedances (K) that were used to estimate the above regression. HKKP suggest setting K equal to half the sample size N, and also note that the function $\gamma(k)$ should be approximately linear over the range $k = (1, \ldots, K)$. In results not reported, we found that setting $K = N/2$ would not be appropriate, as none of the six Hill plots were linear over such a wide range.[8] However, each of the plots in figure 10.5 do indicate a range of k over which $\gamma(k)$ is approximately linear. We have chosen K accordingly.

The third column of table 10.4 reports the estimate of β_0 that was obtained using the optimal K. The estimates vary between 0.50 and 0.86, which implies that the maximal moment $\alpha = 1/\xi$ varies between 1.16 and 2.00. These findings confirm the intuition that operational losses have a heavy-tailed severity distribution. The parameter estimates in the third column of table 10.4 are used in simulations of the aggregate loss distribution reported in section 10.9, table 10.5.

8. To preserve the banks' confidentiality, we do not report Hill plots using either N or $N/2$ exceedances, as doing so would reveal the number of losses at each bank.

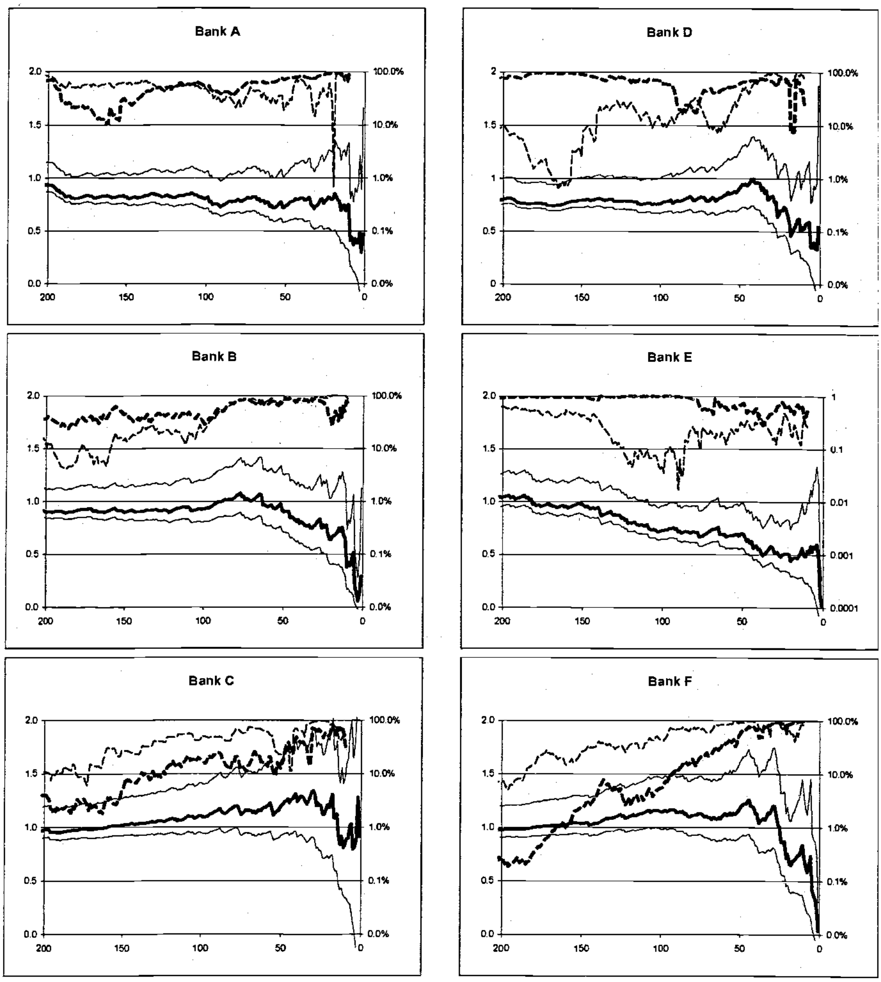

Fig. 10.5 Hill plots of the tail index parameter

Notes: The following are Hill plots of the tail index parameter for the six banks under construction. The thick dark line indicates the point estimates of the tail parameter as the number of exceedances varies between 1 and 200. The thin dark lines indicate 95 percent confidence intervals for the point estimates. The thick, medium gray (light gray) line indicates *P*-values for the Likelihood Ratio test of the hypothesis that the tail parameter is constant across business lines (event types).

The last row of the table reports results obtained for a sample consisting of all six banks. Interestingly, the resulting parameter estimate of 0.68 is consistent with the results of de Fontnouvelle et al. (2003), who reported tail-parameter estimates of about 0.65. This consistency is remarkable, given that de Fontnouvelle et al. (2003) used external, publicly reported

Table 10.4 Tail parameter estimates based on the HKKP method

		No. of exceedances used in estimation			
Bank ID	Optimal K	K	$0.75K$	$0.5K$	$0.25K$
A	180	0.823	0.794	0.817	0.717
		(0.016)	(0.020)	(0.030)	(0.042)
B	80	0.628	0.591	0.565	0.313
		(0.020)	(0.022)	(0.029)	(0.016)
C	30	0.859	0.824	0.952	1.032
		(0.085)	(0.097)	(0.182)	(0.353)
D	50	0.498	0.405	0.456	0.415
		(0.019)	(0.015)	(0.028)	(0.039)
E	200	0.552	0.534	0.558	0.488
		(0.003)	(0.008)	(0.013)	(0.018)
F	50	0.633	0.538	0.536	0.342
		(0.030)	(0.019)	(0.038)	(0.026)
All	140	0.681	0.554	0.419	0.305
		(0.014)	(0.012)	(0.008)	(0.015)

Notes: This table reports tail index estimates calculated under the HKKP regression algorithm. The optimal number of exceedances (K) is chosen to correspond to the linear portion of the Hill plot. Standard errors are reported in parentheses.

loss data (rather than internal data), as well as substantially different empirical techniques than the current paper.

The final three columns of table 10.4 report tail-index estimates obtained using different numbers of exceedances in the regression procedure of HKKP. For all six banks, the results do not change materially when the number of exceedances is reduced from K to $0.75K$ or $0.5K$. The results change more when $0.25K$ exceedances are used. Overall, table 10.4 confirms that the estimation results are not highly sensitive to the choice of K.

10.7.1 Estimation by Business Line and Event Type

Our Hill plot analyses have thus far taken place at the "top of the house" level, where data are aggregated across both business line and event type. However, one might ask whether this approach is appropriate, or whether the tail behavior of the loss-severity distribution might vary by business line and event type. To investigate this issue, we calculated for each value of k (the number of exceedances) separate Hill estimators for each business line and event type. For each k, we then calculated likelihood ratio test statistics for the hypothesis that the tail index is constant across business lines and that it is constant across event types. The probability values for these statistics are reported graphically in figure 10.5. The results indicate that both hypotheses can sometimes be rejected at the 10 percent level when k is near 200. However, neither hypothesis can be rejected at the 10 percent level for values of k where the Hill estimator is constant (banks A, C, and

E) or decreasing (banks B, D, and F). Because choosing a small k provides a less biased value of the Hill estimator, segregating the analysis by business line or event type does not seem to be called for. This finding does not mean that tail behavior of operational losses is constant across business lines and event types. Rather, the ability of statistical estimation techniques to meaningfully differentiate tail behavior across business lines is hindered by a lack of data on large losses using only internal data for one year, and by the concentration of these data in one or two business lines and event types.

10.7.2 On the Possibility of Thin-Tailed Severity Distributions

The results presented in table 10.4 suggest that loss-severity distributions at the six banks under consideration have tail indexes ranging between 0.50 and 0.86. The reported standard errors also seem to exclude the possibility that $\xi = 0$, which would indicate a thin-tailed loss distribution. However, the Hill estimator is designed for situations where $\xi > 0$. Thus, it cannot be used to reject the hypothesis of a thin-tailed loss distribution. This is an interesting hypothesis, because thin-tailed distributions, such as the lognormal, could have significantly different implications for capital than fatter-tailed distributions, such as the Pareto.

Dekkers, Einmahl, and de Haan (1989) show how to extend the Hill estimator so that it is valid for any ξ in $\Re$. The graph of this estimator as k varies is commonly referred to as a DEdH plot. Figure 10.6 reports DEdH plots for the six banks under consideration. These plots indicate that for the low values of k for which the Hill estimator was constant or decreasing, we cannot reject the null of a thin-tailed severity distribution at any of the six banks. This is problematic. The choice of fat versus thin tailed loss severity distribution will have significant impact on the capital calculation, yet based on limited data for only one year, available statistical techniques provide little guidance on which choice is more appropriate. We expect that as banks accumulate more data on large losses, the DEdH plots will either be able to reject the null of $\xi = 0$ or will indicate tail estimates close enough to zero that the choice does not matter so much. For now, we explore the empirical consequences of assuming a thin-tailed loss severity distribution.

Extreme value theory suggests that the exponential distribution is an appropriate choice for modeling loss severity under the thin-tailed assumption. Thus, we wish to construct a threshold plot showing how the exponential parameter varies as the threshold increases (k decreases). Because the maximum likelihood estimate of this parameter is given by the mean excess, these threshold plots would be identical to the mean excess plots already presented in figures 10.3 and 10.4. As discussed earlier, the mean excess plots suggest that the exponential distribution does not provide an accurate description of the tail behavior of operational losses. All six banks' excess plots are concave and increasing, whereas exponentially distributed data imply a linear and horizontal excess plot.

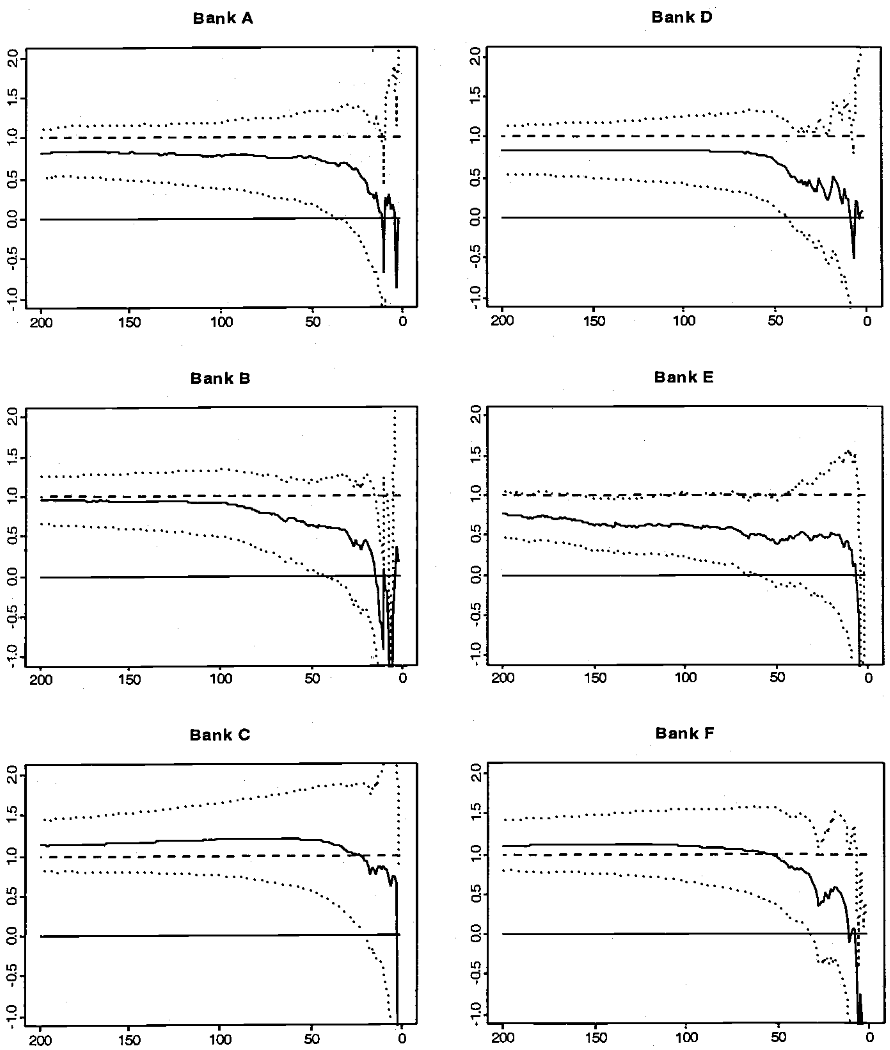

Fig. 10.6 DEdH plots of the tail index parameter

Notes: The following are DEdH plots of the tail index parameter for the six banks under consideration. The solid line indicates the point estimates of the tail parameter as the number of exceedances varies between 1 and 500. The dotted lines indicate 95 percent confidence intervals for the point estimates.

Since the DEdH plots do suggest that tail behavior of operational losses might be modeled with a light-tailed distribution, we consider whether some other such distribution provides a better fit to the data than the exponential. Because the log-normal was the one light-tailed distribution investigated in section 10.6 that provided a good fit across multiple banks, business lines, and event types, we investigate whether it might also provide a useful description of the tail behavior of operational losses.

Figures 10.7 and 10.8 present threshold plots for the six banks, under the assumption that losses above high thresholds follow a (truncated) lognormal distribution. For each value of k (the number of exceedances), estimates

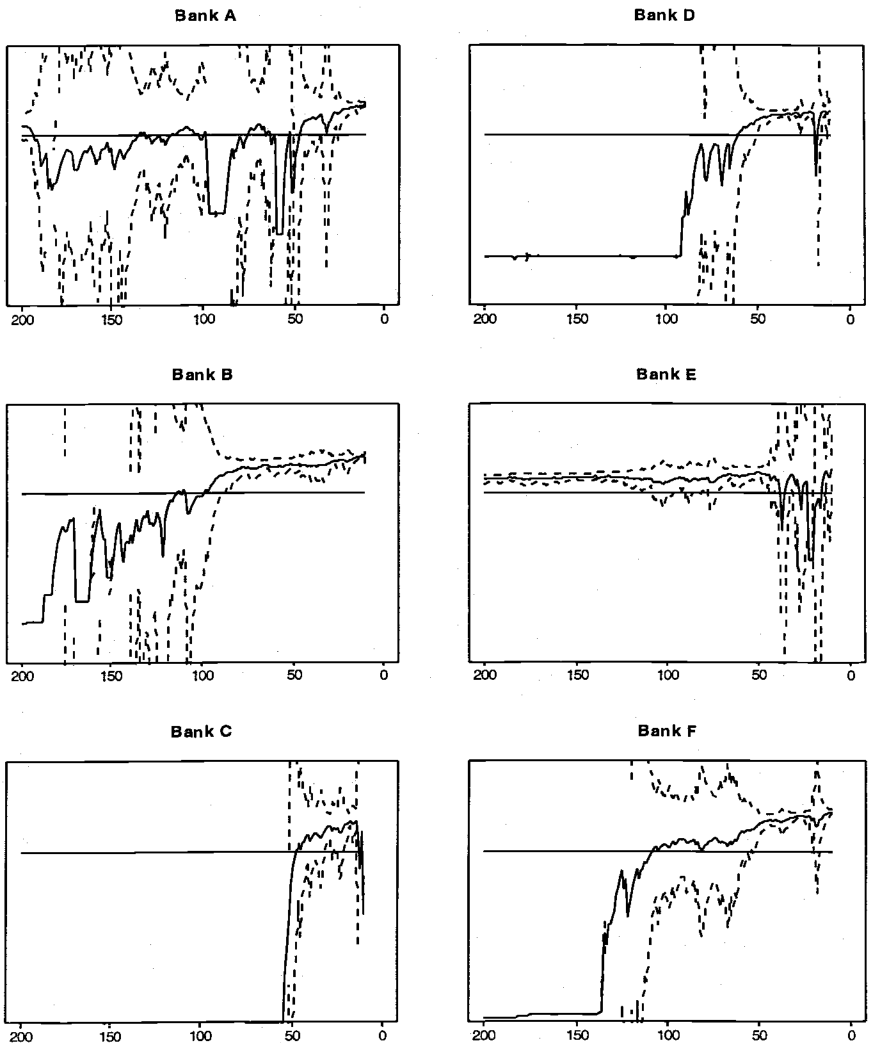

Fig. 10.7 Threshold plots of the lognormal parameter μ

Notes: The following are threshold plots of the lognormal parameter μ for the six banks under consideration. The solid line indicates the point estimates of μ as the number of exceedances varies between 10 and 200. The dotted lines indicate 95 percent confidence intervals for the point estimates. Labels are omitted from the vertical axis to preserve confidentiality.

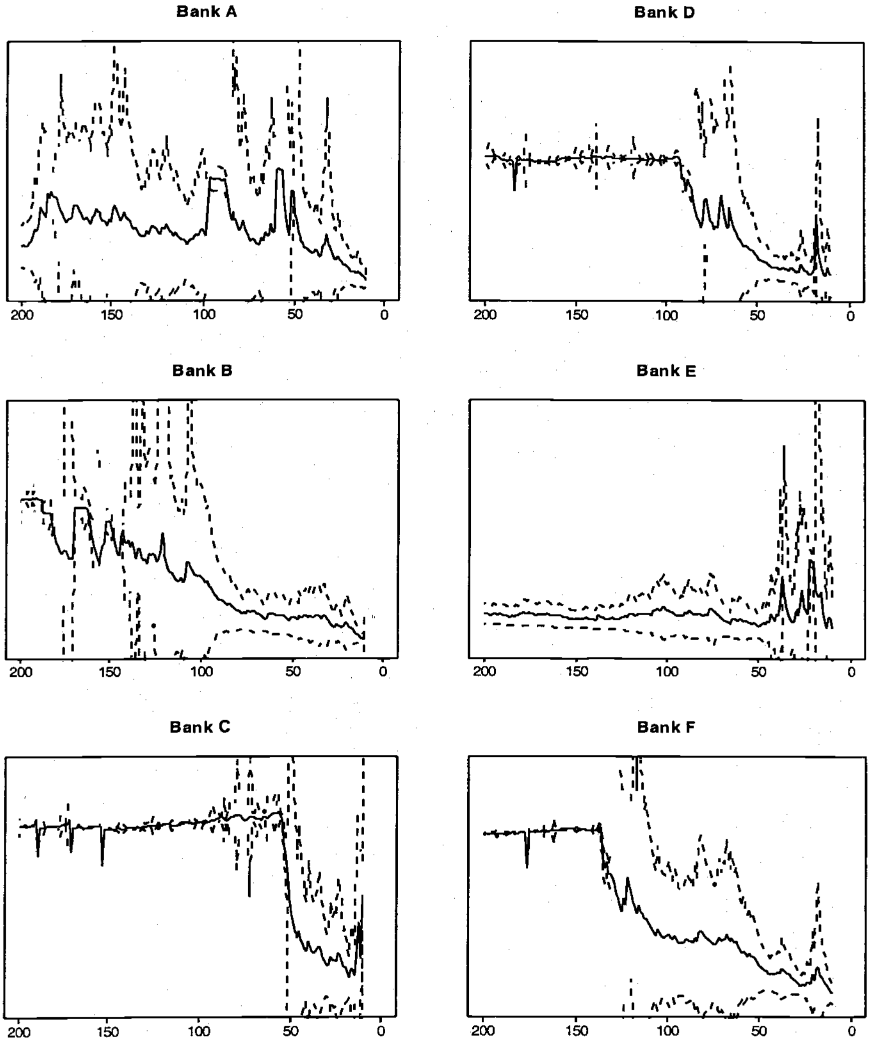

Fig. 10.8 Threshold plots of the lognormal parameter σ

Notes: The following are threshold plots of the lognormal parameter σ for the six banks under consideration. The solid line indicates the point estimates of σ as the number of exceedances varies between 10 and 200. The dotted lines indicate 95 percent confidence intervals for the point estimates. Labels are omitted from the vertical axis to preserve confidentiality.

of the lognormal parameters were obtained via maximum likelihood. (Vertical axis scales have been omitted to protect data confidentiality. However, a reference line in figure 10.7 indicates the location of $\mu = 0$.) One can discern a common pattern in the estimates of μ and σ across all six banks. For example, consider the plots for bank B. Both suggest that the lognormal is not a good fit for more than 100 exceedances, in that the estimates are unstable as k varies and the point estimate for μ is less than 0. We have already argued that this is not a reasonable characterization of operational loss data. However, parameter estimates are both stable and reasonable when thirty to seventy exceedances are used for estimation. The μ estimate lies between 4 and 8, while the σ estimate lies between 0 and 2.[9] Three of the other banks display a similar pattern, with stable (and reasonable) parameter estimates emerging over high thresholds. The two remaining banks' (C and E) POT plots become unstable for small numbers of exceedances.

In results not reported, we chose a specific number of exceedances K for each of the six banks reported in figures 7 and 8, such that the estimates of μ and σ are stable for $k \leq K$.[10] (This procedure follows the traditional analysis of the Hill plot as discussed in Embrechts, Klüppelburg, and Mikosch, 1997). The resulting estimates of μ and σ are used in simulations of the aggregate loss distribution reported in section 10.9, table 10.5.

10.8 The Operational Loss Frequency Distribution

We have thus far focused on the loss severity distribution, which describes the potential size of an operational loss, given that the loss has occurred. Operational risk capital will also depend on the loss frequency distribution, which describes how many losses might actually occur over a given time period. The Poisson distribution is a natural starting point for modeling loss frequency because it arises whenever the loss occurrence rate is constant over time. We thus begin by modeling frequency at bank i by the following:

$$n_i \sim \text{Po}(\lambda_i) \tag{3}$$

That the Poisson distribution has only one parameter makes it particularly attractive in the current context. The LDCE does not provide information regarding the date of an event, beyond the knowledge that all losses occurred sometime during the year 2001. Thus, we have enough information to estimate the Poisson parameter, but not enough to estimate multiparameter frequency distributions. Maximum likelihood estimates of the parameter λ are given by the annual number of loss events.[11]

An interesting property of a Poisson variable is that the mean and variance are equal. So if a LDCE bank were to report 10,000 loss events for the year 2001, we would expect (with 95 percent probability) it to report between 9,800 and 10,200 events the following year. On an intuitive level, this

9. The actual range of variation is significantly narrower, but has not been reported, so as to protect data confidentiality.

10. These results were not reported because estimates of μ and σ would reveal confidential information regarding the loss-severity distribution at individual institutions.

11. To preserve confidentiality, we have not reported the number of loss events.

seems like a very narrow range, and one might ask whether frequency should be modeled via a distribution permitting more variability than the Poisson. One such distribution is the negative binomial, which is a commonly used generalization of the Poisson.

As was discussed earlier, the LDCE data do not support estimation of two-parameter frequency distributions at the individual bank level. In order to model excess dispersion in the loss frequency distribution, we take a cross-sectional approach. That is, we estimate the following regression:

$$n_i \sim F(X_i, \mathbf{b}), \quad (4)$$

where $F(\cdot)$is a discrete nonnegative-valued distribution, X_i is an observable characteristic of bank i (e.g., asset size), and **b** is a parameter vector. Because our data set is purely cross-sectional (i.e., there is no time series element), we cannot estimate any fixed effects. Fixed effects represent bank-specific variation in the frequency of operational losses, which could arise from factors such as the quality of an individual bank's risk control environment. However, it is worth nothing that equation (3) can be interpreted as a fixed effects model. Seen in this light, equations (3) and (4) are different but complementary ways of treating the fixed effects issue. Under the latter, the expected number of events is purely a function of a bank's observable characteristics, whereas under the former, the expected number of events is purely bank specific.

We begin by estimating equation (4) under the assumption that $F(\cdot)$ is the Poisson distribution, so that $n_i \sim \text{PO}(\text{mean} = bX_i)$. Setting each X_i as bank i's total assets as of year-end 2001, we obtain an estimate of 8.2 for the parameter b. This indicates that banks in our sample reported on average 8.2 operational events for every billion dollars in assets. Next, we estimate equation (4) under the assumption that $F(\cdot)$ is the negative binomial distribution, so that $n_i \sim \text{NB}(\text{mean} = b_1X_i, \text{dispersion} = b_2)$. We obtain an estimate of 7.4 for b_1 and 0.43 for b_2.

10.9 The Aggregate Loss Distribution

In this section, we combine the severity results of section 10.7 with the frequency results of section 10.8 in order to estimate economic capital for operational risk, which is specified as the 99.9th percentile of the aggregate loss distribution. We explore two alternate assumptions regarding the loss frequency distribution: the Poisson and negative binomial distributions, as estimated in section 10.8. We also explore three different assumptions regarding the tail of the loss-severity distribution: the Pareto as estimated in section 10.7 (table 10.4), the lognormal as estimated in section 10.7b, and the empirical distribution.

We use Monte Carlo simulation to derive an estimate of the aggregate loss distribution as follows. In the case of the empirical severity distribution, the number of loss events in year i is drawn at random from the frequency distribution, and is denoted N_i. Then, N_i individual losses ($l[1], \ldots, l[N_i]$) are drawn from the empirical distribution. The N_i losses are summed to obtain the aggregate loss for year i. This process is repeated for one million simulated years in order to obtain the aggregate loss distribution.

Monte Carlo simulation for the Pareto (lognormal) severity distribution proceeds similarly, except that losses in ($l[1], \ldots, l[N_i]$) greater than or equal to the relevant threshold value are replaced with random draws from the Pareto (lognormal) distribution estimated in section 10.6.[12] The N_i losses are then summed to obtain the aggregate loss for year i, and the process is repeated for one million simulated years in order to obtain the aggregate loss distribution. The use of Monte Carlo techniques in the current context has already been extensively documented, and we refer readers interested in further details to Klugman, Panjer, and Wilmot (1998) and Embrechts, Kaufmann, and Samorodnitsky (2002), and to their references.

10.9.1 Simulations Based on a Poisson Frequency Distribution

In this subsection, we assume that the frequency of operational losses follows a Poisson distribution with a fixed effects specification, as in equation (3). We make three different assumptions for loss severity: the Pareto, the lognormal, and the empirical distribution. Results are presented in panel A of table 10.5. To preserve the confidentiality of the banks in the sample, we scaled each percentile for each bank by that bank's assets. The cross-bank median for each percentile is then reported.

In 2001, the Basel Committee conducted a quantitative impact study covering 140 banks in twenty-four countries. The committee reported that the median (mean) ratio of reported operational risk capital to minimum regulatory capital was 12.8 percent (15.3 percent), and concluded that "a reasonable level of the overall operational risk capital charge would be about 12 percent of minimum regulatory capital."[13] If one estimates minimum regulatory capital to be 5 percent of a bank's assets, then a reasonable benchmark value for operational risk capital would be 0.6 percent of assets. The median value of 0.468 percent reported in Panel A (for the 99.9th percentile) seems roughly consistent with this benchmark. It is also worth noting that our estimation is based solely on internal loss data for one year, providing limited data to estimate high-severity losses. Banks are also using external loss data and scenario analysis to provide additional information on the tail where they have insufficient high-severity losses in

12. For the lognormal distribution, the relevant threshold is the same as that used for estimation of the tail parameter. For the Pareto distribution, the relevant threshold is the largest observed loss value. This is because by construction, the HKKP tail-parameter estimate β_0 corresponds to zero exceedances.

13. See page 26 of Basel Committee on Banking Supervision (2001).

Table 10.5 **Quantiles of the simulated aggregate loss distribution**

	Percentiles of the aggregate loss distribution (%)		
Severity distribution	95	99	99.9
A. Poisson frequency distribution—fixed effects model			
Pareto	0.066	0.117	0.468
Lognormal	0.047	0.056	0.070
Empirical	0.047	0.053	0.058
B. Poisson frequency distribution—cross-sectional model			
Pareto	0.106	0.148	0.362
Lognormal	0.089	0.101	0.121
Empirical	0.086	0.093	0.102
C. Negative binomial frequency distribution—cross-sectional model			
Pareto	0.166	0.237	0.400
Lognormal	0.143	0.198	0.273
Empirical	0.146	0.202	0.273

Notes: This table reports quantiles of the simulated aggregate loss distribution. To preserve the confidentiality of the banks in the sample, we scale each percentile for each bank by that bank's assets. The cross-bank median for each percentile is then reported. Panel A presents results under the assumption that loss frequency follows a Poisson distribution whose parameter is estimated separately for each bank (fixed-effects model). Panel B presents results under the assumption that loss frequency follows a Poisson distribution whose parameter is a linear function of each bank's asset size (cross-sectional model). Panel C presents results under the assumption that loss frequency follows a negative binomial distribution whose parameter is a linear function of each bank's asset size (cross-sectional model).

a particular business line. Thus, we would view the figure of 0.468 percent as a lower bound on the banks' true operational risk exposure.

The next set of simulations is conducted under the assumption that the severity of operational losses follows a lognormal distribution. The results suggest that cross-bank median of the 99.9th percentile is 0.07 percent of assets. This figure seems small in comparison with both that obtained in the Pareto-based simulations and with the 0.6 percent benchmark discussed previously.

We conducted the final set of simulations by drawing the number of loss events from a Poisson distribution and the loss amounts from the empirical severity distribution. One may think of the resulting 99.9th percentiles as a lower bound on the true capital requirement. Alternatively, one may think of these percentiles as representing the portion of capital that derives from banks' actual loss experience, rather than from their exposure—as measured by a fitted distribution function, which would also include information from external data and scenario analysis. Because the lognormal is a thin-tailed distribution, the 99.9th percentile based on the lognormal severity distribution exceeds that based on the empirical distribution by about 20 percent. Because the Pareto is a heavy-tailed distribution, the 99.9th percentile based on the Pareto severity distribution exceeds that based on the empirical distribution by a factor of eight.

10.9.2 Simulations Based on a Negative Binomial Frequency Distribution

In the previous section, we assumed that the frequency of operational losses followed a Poisson distribution. We found that assuming a Pareto severity distribution yielded capital estimates that were broadly consistent with the Basel Committee's expectation that operational risk accounts for 12 percent of minimum regulatory capital. Assuming a lognormal severity distribution yielded markedly lower capital estimates. In this section, we investigate how these results change under the assumption that the frequency of operational losses follows a negative binomial distribution, as was discussed in section 10.8.

Panels B and C of table 10.5 report quantiles of aggregate loss distributions that were simulated using cross-sectional frequency models based on the Poisson and negative binomial distributions, respectively. (Note that the cross-sectional Poisson model is included because it is not informative to directly compare the cross-sectional negative binomial results with the fixed-effects Poisson results, as differences could be due to either differences in the handling of effects or to differences in the assumed frequency distribution.) The negative binomial specification implies significantly more variability in the number of operational losses than does the Poisson specification. Thus, intuition suggests that the aggregate loss distribution should have a heavier tail under the negative binomial specification. This intuition proves correct in the case of the lognormal severity distribution. The median 99.9th percentile is about twice as large under the negative binomial as under the cross-sectional Poisson specification. However, intuition proves incorrect in the case of the Pareto distribution, for which the median 99.9th percentile is not materially different under the negative binomial than under the Poisson.[14]

Under the negative binomial specification of loss frequency, it is difficult to decide whether the Pareto or the lognormal provides the more useful characterization of the loss-severity distribution. The difference between the two sets of results is within an order of magnitude that may be considered close given the preliminary nature of the data and techniques.

10.10 Conclusion

This paper examines operational risk modeling using only internal operational loss data. By focusing on internal data, it captures the potential modeling issues faced by banking organizations that have only recently started to collect comprehensive loss data. The analysis indicates that the

14. It has been argued that intuition can be misleading if risks follow very heavy-tailed Pareto-type distributions (e.g., Embrechts, McNeill, and Straumann 2002, Rootzen and Klüppelberg 1999).

data do show statistical regularities, and that the severity ranking of event types is similar across banks. The analysis also shows that the data is reasonably fit by heavy-tailed distributions (such as the Pareto), and illustrates that certain statistical methods yield plausible tail-parameter estimates for these heavy-tailed distributions. In fact, the tail-parameter estimates for the severity distribution are quite close to the estimates based on publicly available time series of high-severity losses (de Fontnouvelle et al. 2003).

It is important to qualify our results by noting that they are based on only one year of loss data. This limited data makes it difficult to distinguish between different distributional assumptions, though some thin-tailed distributions do appear inconsistent with the data. At this point, we would conclude that a variety of threshold-based techniques seem to yield results that are consistently plausible across banks. However, we may need to await the arrival of better data before making more definitive conclusions. As banks obtain three or more years of good operational loss data, the ability to differentiate across alternative distributional assumptions should improve.

References

Basel Committee on Banking Supervision. 2001. Working Paper on the regulatory treatment of operational risk. Basel, Switzerland: Bank for International Settlements.

De Fontnouvelle, Patrick, Virginia Dejesus-Rueff, John Jordan, and Eric Rosengren. 2006. Forthcoming. Capital and risk: New evidence on implications of large operational losses. *Journal of Money, Credit and Banking.*

Dekkers, A., J. Einmahl, and L. de Haan. 1989. A moment estimator for the index of an extreme-value distribution. *Annals of Statistics* 17:1833–55.

Embrechts, Paul, Roger Kaufmann, and Gennady Samorodnitsky. 2002. Ruin theory revisited: Stochastic models for operational risk. In *Risk management for central bank foreign reserves,* ed. C. Bernadell, P. Cardon, J. Coche, F. Diebold, and S. Manganelli, 143–61. Frankfurt: European Central Bank.

Embrechts, Paul, Claudia Klüppelberg, and Thomas Mikosch. 1997. *Modelling external events for insurance and finance.* New York: Springer-Verlag.

Embrechts, Paul, Alexander McNeil, and Daniel Straumann. 2002. Correlation and dependence in risk management: Properties and pitfalls. In *Risk management: Value-at-Risk and beyond,* ed. M. A. H. Dempster, 176–223. Cambridge: Cambridge University Press.

Gabaix, Xavier. 1999. Zipf's Law for cities: An explanation. *Quarterly Journal of Economics* 114:739–67.

Huisman, Ronald, Kees Koedijk, Clemens Kool, and Franz Palm. 2001. Tail-index estimates in small samples. *Journal of Business and Economic Statistics* 19:208–16.

Klugman, Stuart, Harry Panjer, and Gordon Willmot. 1998. *Loss models.* New York: Wiley.

Moscadelli, Marco. 2004. The modeling of operational risk: The experience from the analysis of the data collected by the Risk Management Group of the Basel Committee, Working paper no. 517, Bank of Italy.

Risk Management Group. 2003. *The 2002 Loss Data Collection Exercise for Operational Risk: Summary of the data collected.* Report to the Basel Committee on Banking Supervision. Basel, Switzerland: Bank for International Settlements.
Rootzen, Holger, and Claudia Klüppelberg. 1999. A single number can't hedge against economic catastrophes. *Ambio* 28 (6): 550–55.

Comment Andrew Kuritzkes

Until recently, operational risk was hard enough to define, let alone quantify. One of the undoubted benefits of the New Basel Capital Accord (Basel II) for international bank capital regulation is that it has standardized the definition of operational risk, at least for the banking industry.[1] Basel II requires banks around the world to collect internal data on operational losses—defined to include losses resulting from the failure of "internal processes, people, or systems" or from external events—and classify losses into one of seven categories.[2] As a result of this mandatory data collection effort, it is now becoming increasingly possible to analyze the behavior of operational losses systematically, within a commonly accepted definitional framework.

Once we are in a position to define operational risk, the next question becomes whether we can measure it. The ability to quantify operational risk has important policy implications, because Basel II bases a new regulatory capital charge for operational risk on banks' internal modeling of operational losses. Under Basel II's "Advanced Measurement Approach," internationally active banks will be required to estimate their exposure to operational losses over a one-year time horizon at the 99.9th percentile level. The regulatory capital charge for operational risk will then be set equivalent to a bank's internal estimate of the tail risk at the 99.9th percentile, or a one-in-one-thousand-year outcome. Overall, the Basel Committee responsible for developing the new bank capital rules expects that this bottom-up calculation of operational risk capital will comprise about 12 percent of total bank regulatory capital. By comparison, the expected operational risk capital requirement is more than five times the regulatory capital charge for market risk that was introduced for banks in the mid-1990s.[3]

I would like to thank Mark Ames of Mercer Oliver Wyman for his help in preparing this comment. Any errors or omissions are my responsibility.

1. See Kuritzkes (2002) for a discussion of the difficulties of defining operational risk, and implications for quantification.
2. See Basel Committee on Banking Supervision (2001).
3. According to a recent study by Beverly Hirtle of the Federal Reserve Bank of New York, the median market risk capital charge for nineteen U.S. bank holding companies subject to Basel's market risk amendment ranged from 1.0 percent to 2.3 percent of required regulatory capital on a quarterly basis from 1998 through 2001. See Hirtle 2005.

Within this context, de Fontnouvelle, Rosengren, and Jordan, working together at the Federal Reserve Bank of Boston, seek to assess how bank operational losses can be modeled from internal data sets. Narrowly, their focus is on whether the techniques of Extreme Value Theory (EVT) can be successfully applied to estimate operational risk distributions for individual banks using internal loss data. The analysis follows from a previous study by the same authors and another colleague (de Fontnouvelle et al. 2003) that applied an EVT approach to two external databases of publicly reported operational losses for the banking industry. That study concluded that a generalized Pareto distribution appeared to fit an aggregate operational loss distribution well.

Significantly, in this paper, the authors extend their analysis to a confidential set of bank-level data collected by the Federal Reserve as part of Basel II's second Quantitative Impact Study (QIS 2). Through QIS 2, the authors are able to analyze the internal loss distributions of six large, internationally active banks that were deemed to have comprehensive data sets for 2001, the year of the impact study. Given confidentiality restrictions, de Fontnouvelle et al. need to protect the anonymity of the six banks in their sample, and they take care in reporting results not to reveal information that could be used to identify the institutions.

The authors make excellent use of their access to the regulatory dataset to provide a unique window on estimation techniques for modeling operational risk. Specifically, they are concerned with three main questions:

1. Are operational losses best characterized by a thin-tailed (e.g., lognormal), or fat-tailed (e.g., Pareto) distribution?
2. Are the shapes of these distributions consistent across individual banks?
3. Do the results provide a reasonable basis for allocating regulatory capital?

Affirmative answers to the first two questions lead to a tentative yes to the third: to the extent that operational losses are fat tailed and can be modeled by a Pareto distribution, then tail estimates of operational risk at the 99.9 percent level are more likely to fall within the expected range for regulatory capital. And to the extent that the same modeling approach can be shown to generate consistent results across the six banks in the sample, the more confident we can be that reliance on internal models will not lead to random differences in capital requirements for similar institutions.

In addressing these questions, de Fontnouvelle et al. need to overcome two challenges.

First, since operational risk capital is defined at a point in the tail of the loss distribution, by definition there will be a paucity of data on extreme losses (99.9 percent events) within any one institution. This is particularly true when looking at a one-year time horizon (although the problem per-

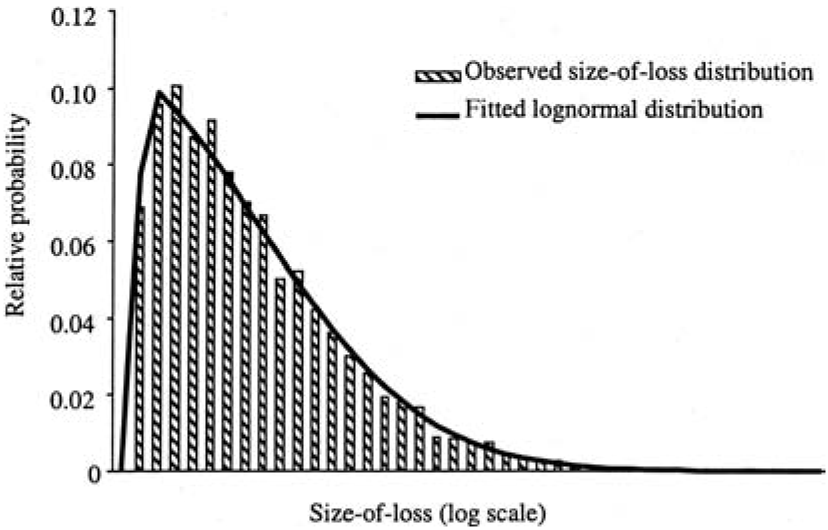

Fig. 10C.1 Lognormal versus empirical fit for a large international bank
Note: Based on five years of internal operational loss data.

sists even if the measurement period is extended to three or more years, as has been the case with many of the large banks preparing for Basel II implementation). If a bank experiences an extreme loss within the measurement period, is this a sign that it is intrinsically more prone to control breakdowns and operational failures, or has the bank just been unlucky? The measurement problem is most acute in the tail—the region of relevance for setting operational risk capital.

Second, observable operational losses are characterized by a large mode of high-frequency, low-severity events, and they appear to fit a lognormal distribution well. This is evident in Figure 10C.1, which shows a fitted lognormal distribution plotted against empirically observed losses for a large international bank, based on five years of the bank's operational loss history. (Like the authors, I cannot reveal the identity of this bank for confidentiality reasons.) A quick visual check reveals that the fit is quite good. Similarly, the authors' own data show how close a lognormal distribution comes to the empirically observed data across the six banks in their study. As summarized in table 10C.1, the lognormal estimates of severity under three different frequency-estimation techniques fall within 9 percent of the observed empirical values at the 99th percentile.

Yet if operational risk is appropriately characterized by a lognormal distribution, the resulting capital estimates will be too "low" to be reasonable. To see this, I have, in table 10C.2, normalized the authors' estimate of the 99.9 percent loss from the lognormal distribution under each of their

Table 10C.1 Lognormal severity versus empirical observations: Aggregate loss distribution across six banks (%)

Frequency estimation approach	99th percentile, empirical fit	99th percentile, lognormal estimate
Poisson fixed effects	.053	.056
Poisson cross-sectional	.093	.101
Negative binomial	.202	.198

Source: De Fontnouvelle et al., table 5.
Note: Scaled as a percentage of bank assets.

Table 10C.2 Lognormal estimates at 99.9 percent

Frequency estimation approach	99.9th percentile (as percent of regulatory capital)
Poisson fixed effects	1.40
Poisson cross-sectional	2.42
Negative binomial	5.46

Source: De Fontnouvelle et al., Table 5.
Note: Scaled relative to 5 percent total regulatory capital to assets.

frequency-estimation approaches (which the authors report as a percentage of total assets) to the 5 percent Basel II regulatory capital-to-asset ratio suggested by the authors.[4] On this basis, the implied regulatory capital charge for operational risk would range from 1.40 percent to 5.46 percent of total Basel II regulatory capital—significantly lower than the Basel Committee's expectation that operational risk capital would account for roughly 12 percent of total regulatory capital.

De Fontnouvelle et al. marshal an impressive array of counterevidence to demonstrate that operational losses should be modeled with a fat-tailed distribution. Their argument follows five main steps.

1. For capital purposes, we are not concerned with the mode of the distribution, but with the tail at the 99.9 percent level.
2. Extreme Value Theory and experience in modeling risks that are similar to operational losses, such as catastrophe risks in insurance, suggests that tails can behave very differently from the body of the distribution.

4. It would be more convenient if the authors reported the results scaled by risk-weighted assets rather than by total bank assets, since total Basel II regulatory capital is defined as a fixed percentage (8 percent) of risk-weighted assets. For consistency, I have adopted the authors' suggestion that total regulatory capital is around 5 percent of total assets in normalizing the results.

3. Profiling of operational loss data for each of the six banks in their sample by tail plots, average excess plots, and chi-square tests indicates that fat-tailed distributions generally outperform thin-tailed distributions.

4. Using the EVT-based peaks-over-threshold approach developed by Huisman, Koedijk, Kool, and Palm (HKKP 2001), a Pareto distribution appears to provide a strong fit in the tail for all six banks.

5. The results are consistent with the previous paper by de Fontnouvelle and colleagues on publicly reported operational losses.

Taken together, the evidence is persuasive that operational losses are indeed fat tailed. For practitioners interested in the behavior of the distribution at around the 99.9 percent level—in particular, bank risk managers and regulators—modeling the operational risk tail using a peaks-over-threshold approach and a generalized Pareto distribution offers a promising solution.

Turning to the paper's second question, do the estimates across individual banks converge? Here we are hampered by the aggregation of results necessary to protect the anonymity of the six banks studied. The authors are not able to report the regulatory capital estimate at the 99.9 percent level for each of the six banks in the study without risking disclosure of confidential information, so instead they report the median result across the six banks (after scaling each bank's results by total assets). This is the basis on which they conclude that "the median value of .468 percent (of assets, equivalent to 9.36 percent of total regulatory capital) reported (see the author's table 10.5) for the 99.9th percentile seems reasonable" in comparison to regulatory expectations.

But while the authors do not report individual loss estimates for the six banks, they do report the shape parameter, ξ, for the generalized Pareto distribution for each bank under the HKKP estimation. The ξ's range from a low of .498 to a high of .859, with a combined aggregate value for the six banks of .681. The difference in shape parameters actually implies a very wide range in tail estimates at the individual bank level.

To illustrate the effect of differences in the shape parameter, my colleague, Mark Ames, and I calculated regulatory capital at the 99.9 percent level across the six values of ξ for hypothetical banks whose exposures were otherwise identical. For each bank we assumed twenty operational loss exceedances per year above a \$1 MM threshold, and a value of \$0.75 MM for the GPD beta. The threshold and beta values were chosen to be broadly consistent with de Fontnouvelle and colleagues' work in their previous study (de Fontnouvelle et al. 2003). The results are reported in table 10C.3.

Differences in the shape parameter appear to have a significant impact on an individual bank's tail risk. The estimates for our hypothetical bank show that for twenty exceedances above \$1 MM, if the shape parameter

Table 10C.3 **Regulatory capital as a function of tail density**

Sample	Value of shape parameter ξ	Regulatory capital estimate at 99.9 percent in $MMs
Bank D	0.498	208
Bank E	0.552	321
Bank B	0.628	600
Bank F	0.633	625
Combined	0.681	935
Bank A	0.823	3,157
Bank C	0.859	4,320

Source: De Fontnouvelle et al., table 4.

Note: HKKP estimates of GPD shape parameter ξ for six banks.

were as low as .498 the 99.9th percentile loss estimate would be $200 MM. If the shape parameter were as high as .859 the 99.9th percentile loss estimate would be $4.3 billion. This twenty-to-one range is consistent with the findings of de Fontnouvelle and colleagues' previous study (de Fontnouvelle et al. 2003) based on externally reported operational losses. From a regulatory perspective, the key question is how confident can we be in each bank's own estimate—based solely on internal data—of its shape parameter. Do such apparently large differences in the shapes of the tail reflect true differences in banks' vulnerability to operational losses, or are they artifacts of measurement?

An order-of-magnitude range may not be that surprising for an attempt to estimate the one-in-one-thousand-year tail risk of operational loss, given that the authors were only able to work with a single year's worth of data for each of the six banks. More generally, the range in magnitude is a reflection of the early stage of development of operational risk measurement. No doubt future research across more banks and on longer datasets will help narrow the range.

In light of the differences in the shape parameter, can we say that the results provide a reasonable basis for allocating capital to individual banks? In my view, the answer is that it is too early to tell. Until we know more about the behavior of tail estimates at the individual bank level, we will not know whether differences in operational risk capital calculations across banks reflect true differences in their loss experience and control environment, or the limitations of using sparse data to forecast extreme events.

References

Basel Committee on Banking Supervision. 2001. Regulatory treatment of operational risk. Bank for International Settlements (BIS) Working Paper no. 8. Available at: www.bis.org/publ/bcbs_wp8.pdf.

de Fontnouvelle, Patrick, Virginia DeJesus-Rueff, John Jordan, and Eric Rosengren. 2003. Using loss data to quantify operational risk. Federal Reserve Bank of Boston Working Paper. Available at: www.bos.frb.org/bankinfo/oprisk/articles .htm.

Hirtle, Beverly. 2005. What market risk capital tells us about bank risk. Forthcoming. *Federal Reserve Bank of New York Economic Policy Review.*

Huisman, Ronald, Kees Koedjik, Clemens Kool, and Franz Palm. 2001. Tail-index estimates in small samples. *Journal of Business and Economic Studies* 19:208–16.

Kuritzkes, Andrew. 2002. Operational risk capital: A problem of definition. *Journal of Risk Finance* 4 (1): 47–56.

Discussion Summary

Part of the discussion revolved around the paucity of observations in the tails of loss distributions, both currently and going forward. *Eric Rosengren* noted that the tail is more populated for some banks than others and that such variation may be a source of the variation in estimated tail-index values that the authors observe. He also noted that under Basel II's advanced-measurement approach, banks are not limited to internal data, but also may use external data and scenario analysis. *Patricia Jackson* wondered whether Basel II's loss-size cutoff for data collection might be raised to reduce costs, but *Ken Abbott* observed that, in his experience, small losses may be indicative of process problems that might result in very large losses under other circumstances. Thus, there should be a role for judgment in internal reporting of small losses.

Darrell Duffie suggested that the authors might take a Bayesian approach to dealing with a potential censorship problem in their data: losses are capped at the level of a firm's capital, because only surviving firms contribute observations to operational risk-loss databases. *Casper de Vries* suggested that the authors could use bootstrap methods in determining the optimal number of observations to use in tail estimation, that they use the empirical distribution in estimating losses occurring in the body of the distribution rather than the lognormal, and that variation in constant terms may account for the variation in tail-index estimates that they observe.