Probability in the Engineering and Informational Sciences, **32**, 2018, 434–459.
doi:10.1017/S0269964817000250

BUDGET ALLOCATIONS IN OPERATIONAL RISK MANAGEMENT

YUQIAN XU
College of Business,
University of Illinois at Urbana-Champaign,
Champaign, IL 61820, USA
E-mail: yxu@stern.nyu.edu

JIAWEI ZHANG and MICHAEL PINEDO
Stern Business School, New York University,
New York, NY 10012, USA
E-mail: jzhang@stern.nyu.edu; mpinedo@stern.nyu.edu

We consider a resource allocation model to analyze investment strategies for financial services firms in order to minimize their operational risk losses. A firm has to decide how much to invest in human resources and in infrastructure (information technology). The operational risk losses are a function of the activity level of the firm, of the amounts invested in personnel and in infrastructure, and of interaction effects between the amounts invested in personnel and infrastructure. We first consider a deterministic setting and show certain monotonicity properties of the optimal investments assuming general loss functions that are convex. We find that because of the interaction effects "economies of scale" may not hold in our setting, in contrast to a typical manufacturing environment. We then consider a general polynomial loss function in a stochastic setting with the number of transactions at the firm being a random variable. We characterize the asymptotic behaviors of the optimal investments in both heavy and light trading environments. We show that when the market is very liquid, that is, it is subject to heavy transaction volumes, it is optimal for a financial firm that is highly risk sensitive to use a balanced investment strategy. Both a heavier right tail of the distribution of transaction volume and a firm's risk sensitivity necessitate larger investments; in a heavy trading environment these two factors reinforce one another. However, in a light trading environment with the transaction volume having a heavy left tail the investment will be independent of the firm's sensitivity to risk.

Keywords: asymptotic, convexity, heavy tail, operational risk, risk sensitivity, stochastic resource allocation

1. INTRODUCTION

Financial firms are subject to various types of risks, in particular credit risk, market risk and operational risk. Of these three types of risks, operational risk (OpRisk) is the type that is the most difficult to assess even though it can be devastating; see Cruz [15]. Typically, it is

© Cambridge University Press 2017 0269-9648/17 $25.00

https://doi.org/10.1017/S0269964817000250 Published online by Cambridge University Press

considered *the risk of losses due to failures of internal processes, people or systems, or due to occurrences of unexpected external events* (see the Basel Committee on Banking Supervision [1]). Compared with other risks, the cost of OpRisk management is often perceived by a firm's management as being more controllable (see Deloitte [16]). However, the literature on the modeling of the management problems associated with OpRisk is not very extensive. The initial motivation of this paper is to propose a quantitative way to manage OpRisk in practice. We formulate a resource allocation model to assist investment decision making that has as goal the minimization of the total expected OpRisk losses. This is one of the main issues that was addressed by the Basel Committee on Banking Supervision [2].

According to Basel II, there are seven different categories of OpRisk event types, namely (i) internal fraud, (ii) external fraud, (iii) employment practices and workplace safety, (iv) clients, products and business practices, (v) damage to physical assets, (vi) business disruptions and systems failures, and (vii) execution, delivery and process management. In this paper, we focus on the two last ones of these event categories, namely, business disruptions and systems failures (e.g., utility disruptions, software failures, and hardware failures), and execution, delivery, and process management failures (e.g., data entry errors, accounting errors, failed mandatory reporting, and negligent loss of client assets). The activities in any unit of a financial services firm are based on a combination of manual processes and automated processes. In Basel II, it is mentioned that it is important to study the interaction between these two types of processes. In Basel II, it is also stated that the performances of the different business units in a firm are positively correlated with one another. Therefore, instead of considering a business unit in isolation, we also consider the contagion effects of the workload of one specific business unit of a firm on the other business units. The model therefore considers multiple business units. The correlation between the different business units of a firm is one of the characteristic properties that distinguish financial services firms from firms in other services industries. The model we consider is therefore very well suited for financial services firms that focus on OpRisk losses due to events caused by excessive workload on personnel or by excessive stress on equipment. Note that we do not consider here incentives or behavioral issues, such as, for example, OpRisk losses caused by rogue traders (which would fall within the category of internal fraud).

We therefore consider in this paper two types of investments, namely investments in equipment and in personnel, and consider the impact of the randomness in the number of transactions (which would be equivalent to the trading volume) on such investments. The number of transactions here could be compared to a production quantity in a manufacturing company. However, a second characteristic property of a financial services firm is the daily randomness of this production quantity caused by market volatility. Manufacturing companies are also subject to some uncertainty in consumer demand; however, a manufacturing firm can always maintain inventory, and hence the production quantity is more stable. From this point of view, the volatility in the daily workload would be a more important cause of operational risks in a financial services firm than in a manufacturing firm. Therefore, budget allocations for personnel and equipment constitute a more difficult problem in a financial services firm. Moreover, a third characteristic property of the financial services industry is the relatively high probability of tail events, and the larger damage caused by such extreme events. Other services industries (e.g., a call center) may also have a high probability of errors in a period with an excessive workload (e.g., many calls arriving at the same time), but even if a call center does not have enough personnel to handle incoming calls, the damage caused by errors may not be that high. However, in the financial services industry, a simple mistake in a heavily traded market could cause a significant loss to the firm. For example, in 2006, a fat-finger error by a trader at Mizuho Securities in Japan caused the firm to sell, erroneously, a stock short; that trade cost the firm 40 billion Yen to

https://doi.org/10.1017/S0269964817000250 Published online by Cambridge University Press

unwind. In the finance world, human errors causing losses in the millions of dollars occur regularly. Therefore, a proper budget allocation for a financial services firm should take into account situations with high volatility in transaction levels. In what follows, we study the asymptotic behaviors of the budget allocation process and consider both the left and right tail effects of the number of transactions.

Our model is also motivated by the problems associated with such investments in financial services firms that have multiple businesses, as well as by the problems associated with the planning of future expansions of a financial services firm. Consider, for example, the operations at the trading desks of a major bank. (Banks typically have various trading desks that are dedicated to different types of securities, including equities, fixed income, and foreign exchange.) A typical foreign exchange (ForEx or FX) trading desk at a large financial firm executes tens of thousands of trades every day. A trading desk may employ easily between 50 and 100 traders. Most of the trades are executed electronically (especially the standard, plain vanilla trades), but certain trades (the more intricate ones) have to be done manually. A firm's database of open trades (i.e., trades that have not reached their maturity date yet) may contain at any point in time several million trades. A certain percentage of these trades may have errors. A mishap could have been caused by a human error (at times referred to as thick finger), or by an equipment malfunction or by a software bug. An error can easily result in a monetary loss in excess of, say, 1 million USD. Such a mishap is in the finance world referred to as an OpRisk event. Therefore, a fourth characteristic property of the financial services industry we capture is the potential damage that may be caused by the automation process or the information system infrastructure. Although nowadays, most companies in the manufacturing world are highly automated, even if a machine goes down because of a heavy workload, the damage will not be as high as the damage a similar event can cause in the finance world. For example, in August 2012 a system bug caused Knight Capital to lose 460 million USD in 30 min.

Our main results are the following. First, when the market is very active, that is, has a high expected transaction volume, it is optimal for a very risk sensitive or vulnerable financial firm to use a balanced investment strategy. Second, the optimal investments are larger when the distribution of the number of transactions has a heavier right tail and when the firm is more risk sensitive; moreover, these two factors reinforce each other in a heavy trading environment. Third, with a heavier left tail, the optimal investments will not depend on the firm's risk sensitivity or vulnerability; hence firms operating in a light trading environment (i.e., firms that operate under lower pressure) will always invest according to the same formula, independent of their size. Finally, we find that because of the effects of the interactions between personnel and infrastructure, "economies of scale" may not apply in our setting, in contrast to a typical manufacturing environment.

The contributions of this paper are threefold. First, we take an initial step toward developing a resource allocation model that specifically takes operational risk losses at a financial services firm into account. Second, we capture four characteristic properties of the financial services industry in our model, and show how these characteristics affect investment decisions. Third, we analyze both the left and right tail effects of the transaction volumes on the investment decisions. We provide managerial insights with regard to these tail effects as well.

This paper is organized as follows. In the next section, we describe the literature in three research areas that are related to our problem, namely OpRisk in financial services, the effects of workload on operational performance, and resource allocation. In Section 3, we propose a general resource allocation model with two investment decision variables, namely the amount to be invested in personnel and the amount to be invested in systems and infrastructure, and study the structural properties of the optimal investments. For

https://doi.org/10.1017/S0269964817000250 Published online by Cambridge University Press

a general formulation of our model in a deterministic setting, we show the monotonicity relationship between the optimal investment and the number of transactions as well as the bounded curvature structure of the optimal investment. To analyze the investment decision problem in a stochastic setting with the number of transactions being a random variable, we use, in Section 4, a general polynomial function as our objective function. We consider in this section a heavy trading environment with the number of transactions being unbounded. The asymptotic behavior of the allocation rules for the necessary investments in equipment and human resources in this regime are functions of the moments of the distribution of the number of transactions. This property provides us with a basis for determining the impact of the tails of various different distribution types on the amounts to be invested. We study the tail effects of three different distribution types in detail. Also, in Section 4, we characterize the nature of the optimal investments and the budget allocation rules in an environment of light trading (i.e., the number of transactions being small). In contrast to the heavy trading environment, the asymptotic behavior of the optimal investments in a light trading environment vary across several different scenarios. We present our conclusions in the last section.

2. LITERATURE REVIEW

Compared with the extensive literature on market risk and credit risk, OpRisk in financial services has not received that much attention. However, over the last decade some books have appeared; see, for example, Chernobai, Rachev, and Fabozzi [12], Cruz [15], Scharfman [35] and Gregoriou [17]. These books discuss OpRisk modeling, measurements, managerial aspects, and regulations in fairly general terms. Most of the recent papers on OpRisk tend to deal with the statistical aspects. They tend to focus on the measurement of OpRisk, that is, the statistical modeling of aggregate OpRisk loss distributions in order to be able to estimate the Operational Value at Risk (Ops-VaR); see Neil, Fenton, and Tailor [31], Bocker and Kluppelberg [6], Sparrow [36]. Only recently papers started to focus on the financial aspects of OpRisk; see Jarrow [20], Jarrow, Oxman, and Yildirim [21], Chapelle et al. [10], Cheng et al. [11], Leippold and Vanini [26]. Very few papers in the literature have considered the actual operations management aspects of OpRisk, that is, Hora and Klassen [18], see Xu, Pinedo, and Xue [43] for an overview. One goal of this paper is to take an initial step in filling the gap between OpRisk in finance and operations management.

In recent years, an increasing number of papers in the Operations Management literature have studied the relationships between workload and various operational performance measures; see Boudreau [7] and Boudreau et al. [8]. However, most of these papers consider operational performance measures in productivity terms rather than in error counts or in operational losses. For instance, Tan and Netessine [37] analyzed a large, detailed operational data set from a restaurant chain and showed an inverted-U-shaped relationship between the workload (defined as the number of tables or diners that a server has to handle at the same time) and the server's performance (measured in sales and in meal durations). Similarly, we will model our workload as the total number of transactions divided by the number of servers (either headcount or computer servers). Several papers in healthcare services have also studied the impact of workload on individual performances. Kc and Terwiesch [24] conducted a rigorous empirical analysis of the impact of workload on service times using operational data from patient transport services in cardiothoracic surgery. The studies mentioned above focus on the impact of workload on service times (productivity). Some other studies examine the impact of workload on certain dimensions of service quality, which is closer related to our work. Kuntz, Mennicken, and Scholtes [25] suggested a

https://doi.org/10.1017/S0269964817000250 Published online by Cambridge University Press

nonlinear relationship between hospital workload and mortality rates. Powell, Savin, and Savva [32] found that overworked physicians generate less revenue per patient because of a workload-induced reduction in due diligence with regard to paperwork. The results presented in Cassenti, Kelley, and Carlson [9] indicate that the human error rate increases more or less linearly with the workload. Taking into account also the effects of fatigue, we propose in our model a convex relationship between workload on personnel and operational losses. At the same time, Iyer and Rossetti [19] showed that the probability of a CPU or a system related error (the load hazard) increases nonlinearly with an increase in the workload; that is, a CPU's or a system's performance rapidly deteriorates when it operates at full capacity; the load hazard is observed to be very sensitive to system CPU utilization, the I/O rate, and interrupt rates. We assume, therefore, a similar convex relationship between workload on equipment (e.g., computer systems) and operational losses. Motivated by this literature on workload and operational performance, we consider in our model specifically the OpRisk losses that are caused by excessive workloads on personnel and on equipment.

A third research area that is also related to our problem concerns both static and dynamic resource allocation problems; see Luss [27] for an overview. A number of papers have been written on stochastic resource allocation problems and stochastic assignment problems; see, for example, [30], [33], and [42]. The capacity investment decision problem and its variants are well studied in the operations management literature; see Van Mieghem [40] for a overview of this class of problems. A number of papers have appeared that focus on variants of investment decision problems; for examples of recent ones, see Van Mieghem and Dada [41], Thomas, Teneketzis, and Mackie-Mason [38], Bish and Wang [3], Chod and Rudi [13], Bish et al. [4], Royset and Szechtman [34], Karger, Oh, and Shah [22], and Karsten, Slikker, and Van Houtum [23]. Among the previous studies, those similar to ours are Meier, Christofides, and Salkin [29] and Cooper [14]. Meier et al. [29] study explicitly optimal capital budgeting in scenarios that are subject to an uncertain future. Cooper [14] study the asymptotic behavior of an allocation policy that is derived for a deterministic revenue optimization problem. Moreover, even though resource allocation models are widely used in finance, most of the allocation models in finance deal with the management of investment risks. In particular, a number of papers consider capital allocation issues with respect to market risk and credit risk, for example, Markowitz [28], Black and Litterman [5]. However, not many applications have been considered so far with regard to OpRisk. A second goal of our paper is to introduce a new resource allocation application in OpRisk management.

The main goal of our paper is to establish a stronger tie between OpRisk in finance and operations management; we analyze budget allocations that have to cover personnel and equipment expenses in settings where the expected OpRisk losses caused by excessive workloads have to be minimized. We consider the uncertainty that is introduced by the randomness in the number of transactions that constitute the workload, and study the asymptotic behaviors of the optimal investments when only the moments of the distribution of the number of transactions are known. We analyze the impacts of the tails of specific distributions in more detail.

3. PROBLEM FORMULATION AND PRELIMINARY RESULTS

There is a strong belief in the finance world that an excessive workload on both personnel and equipment (mainly on the computer systems) has an impact on OpRisk losses. We assume therefore that the operational loss is a function of workload. Following the literature on workload and operational performance, we propose to model the workload on personnel in a financial institution as the number of transactions divided by the headcount. One

https://doi.org/10.1017/S0269964817000250 Published online by Cambridge University Press

advantage of the formula we propose is that the headcount is proportional to the amount invested in human resources, and hence can be replaced by the amount of capital invested in human resources. Following Iyer and Rossetti [19], we assume that a similar formula holds for the workload on equipment. Furthermore, we assume that the OpRisk losses increase convexly in both types of workload. In what follows, we propose a general convex function to model OpRisk losses for the entire company.

Consider a financial firm that needs the two types of assets to conduct its operations: infrastructure (e.g., computer equipment) and headcount (personnel). A manager of such a firm has to invest in these two types of assets in order to run its operations and has to determine the amounts to invest that minimize the firm's total investment cost plus its total OpRisk losses. Let the variable N denote the total number of transactions that have to be executed in a given time period. The decision variables of the firm's manager are I_{e} and I_{h}, which denote the amounts to be invested in equipment and in human resources. Based on our previous discussion, we define the workload on equipment as N/I_{e} and the workload on personnel as N/I_{h}. Let $C(N, I_{\mathrm{e}}, I_{\mathrm{h}})$ denote the total cost function, and let $h_{\mathrm{e}}(N/I_{\mathrm{e}})$, $h_{\mathrm{h}}(N/I_{\mathrm{h}})$, $h_r(N/I_{\mathrm{e}}, N/I_{\mathrm{h}})$ denote the loss functions corresponding to equipment workload, personnel workload and interaction effects between equipment and personnel. As proposed by the Basel Committee on Banking Supervision [2], we specifically identify the interaction effects between personnel and equipment here. We consider here the interaction term as a function of the product of the two types of workload, and in its general form can be written as

$$h_r(N/I_{\mathrm{e}}, N/I_{\mathrm{h}}) = h_r\left(\left(\frac{N}{I_{\mathrm{e}}}\right)^{\alpha}\left(\frac{N}{I_{\mathrm{e}}}\right)^{\beta}\right), \tag{3.1}$$

where $\alpha, \beta \geq 1$. Note that one may also consider the interaction term as a function of N^a, where $a \leq 1$, but we can see in the later analysis that under this condition, most of the results will actually be even more straightforward. Moreover, a convex function of N seems to be more consistent with empirical evidence. We can formulate our cost minimization problem as

$$\begin{aligned}&\min_{I_{\mathrm{e}}>0, I_{\mathrm{h}}>0} \; C(N, I_{\mathrm{e}}, I_{\mathrm{h}})\\ &= \min_{I_{\mathrm{e}}>0, I_{\mathrm{h}}>0} \left\{h_{\mathrm{e}}\left(\frac{N}{I_{\mathrm{e}}}\right) + h_{\mathrm{h}}\left(\frac{N}{I_{\mathrm{h}}}\right) + h_r\left(\left(\frac{N}{I_{\mathrm{e}}}\right)^{\alpha}\left(\frac{N}{I_{\mathrm{e}}}\right)^{\beta}\right) + (I_{\mathrm{e}} + I_{\mathrm{h}})\right\}.\end{aligned} \tag{3.2}$$

In this section, we consider the deterministic version of the problem; stochastic settings are discussed in subsequent sections. Before analyzing the structure of optimal results, we specify the assumptions with regard to the proposed functions.

ASSUMPTION 3.1: *The real variable functions $h_e(x)$, $h_h(x)$, and $h_r(x)$ are*

(i) first and second order differentiable on $\mathbb{R}^n$;
(ii) increasing and strictly convex in x.

Under Assumption 3.1, we can show that an optimal solution $(I_{\mathrm{e}}^*, I_{\mathrm{h}}^*)$ that minimizes the total cost $C(N, I_{\mathrm{e}}, I_{\mathrm{h}})$ exists and is unique.

LEMMA 3.2: *The objective function $C(N, I_e, I_h)$ of problem (3.2) is strictly convex in (I_e, I_h) and therefore, the problem has a unique optimal solution.*

https://doi.org/10.1017/S0269964817000250 Published online by Cambridge University Press

We are able to characterize the structure of optimal solution $(I_{\mathrm{e}}^*, I_{\mathrm{h}}^*)$ with respect to the number of transactions N. If we would have considered N as a production quantity in a manufacturing firm, it could be argued, via an "economies of scale" argument, that with an increasing N the required investment would increase concavely. However, this is not obvious in our setting with regard to the operational risk in a financial services firm. The following proposition states how the optimal investments depend on the number of transactions N.

PROPOSITION 3.3: *The optimal investments in equipment (I_e^*) and in personnel (I_h^*) increase monotonically in the total number of transactions N. However, both $I_e^*/N^{\alpha+\beta}$ and $I_h^*/N^{\alpha+\beta}$ decrease in N.*

PROOF: In our objective function (3.2), I_{e} and I_{h} are symmetric, so we only have to show that I_{e}^* increases monotonically in N. Let $y_{\mathrm{h}} \triangleq N^{\alpha+\beta}/I_{\mathrm{h}}$. The objective function now becomes

$$\min_{I_{\mathrm{e}}>0, y_{\mathrm{h}}>0} \; C(N, I_{\mathrm{e}}, y_{\mathrm{h}}) = h_{\mathrm{h}}\left(\frac{y_{\mathrm{h}}}{N^{\alpha+\beta-1}}\right) + h_{\mathrm{e}}\left(\frac{N}{I_{\mathrm{e}}}\right) + h_r\left(\frac{y_{\mathrm{h}}}{I_{\mathrm{e}}}\right) + I_{\mathrm{e}} + \frac{N^{\alpha+\beta}}{y_{\mathrm{h}}}.$$

We first show that $C(N, I_{\mathrm{e}}, y_{\mathrm{h}})$ is submodular in $(N, I_{\mathrm{e}}, y_{\mathrm{h}})$. Taking the second-order partial derivatives of C, we have

$$\frac{\partial^2 C}{\partial N \partial y_{\mathrm{h}}} = -(\alpha+\beta-1)\frac{h_{\mathrm{h}}'}{N^{\alpha+\beta}} - (\alpha+\beta-1)\frac{y_{\mathrm{h}} h_{\mathrm{h}}''}{N^{2(\alpha+\beta)-1}} - \frac{(\alpha+\beta)N^{\alpha+\beta-1}}{y_{\mathrm{h}}^2} < 0,$$

$$\frac{\partial^2 C}{\partial N \partial I_{\mathrm{e}}} = -\frac{h_{\mathrm{e}}'}{I_{\mathrm{e}}^2} - \frac{N h_{\mathrm{e}}''}{I_{\mathrm{e}}^3} < 0,$$

$$\frac{\partial^2 C}{\partial y_{\mathrm{h}} \partial I_{\mathrm{e}}} = -\frac{h_r'}{I_{\mathrm{e}}^2} - \frac{y_{\mathrm{h}} h_r''}{I_{\mathrm{e}}^3} < 0.$$

The above inequalities follow from the monotonicity and convexity properties of the functions $h_{\mathrm{h}}(\cdot)$, $h_{\mathrm{e}}(\cdot)$, and $h_r(\cdot)$. Therefore, $C(N, I_{\mathrm{e}}, y_{\mathrm{h}})$ is submodular in $(N, I_{\mathrm{e}}, y_{\mathrm{h}})$, and the optimal solution $(I_{\mathrm{e}}^*, y_{\mathrm{h}}^*)$ is increasing in N. This monotonicity property follows from the Increasing-Optimal-Selections Theorem (see [39]). From a similar approach, we can prove that I_{h}^* and $y_{\mathrm{e}}^* = N^{\alpha+\beta}/I_{\mathrm{e}}^*$ increase in N as well. ■

From Proposition 3.3, we know that if we define the total investment budget B as $B = I_{\mathrm{e}} + I_{\mathrm{h}}$, then B increases monotonically in the number of transactions N. However, B does not necessarily increase concavely in N. The only structure we can establish is that the curvature of $B^*(N)$ is bounded by the curvature of a general convex function $N^{\alpha+\beta}$ which depends on the interaction term. This result shows that "economies of scale" may not hold in our setting. However, if we remove the interaction term here, we can show through a similar approach that the unit investment I_{e}/N or I_{h}/N actually decreases in N, which is consistent with "economies of scale".

We now show numerically for the following special case that this $B^*(N)$ can increase as a function of N either concavely or in the form of an S-curve under the impact of the interaction process. Let δ denote the ratio of the amount of money invested in equipment, that is, $I_{\mathrm{e}} = \delta B$ and $I_{\mathrm{h}} = (1-\delta)B$. We can further show that for fixed δ, the curvature of $B^*(N)$ is bounded by a linear function. We define such function here as a *sublinear function.*

Example 3.4 (A Quadratic Loss Function): Consider the following quadratic cost minimization problem. Let δ be the ratio of the amount invested in equipment, that is, $I_{\mathrm{e}} = \delta B$

https://doi.org/10.1017/S0269964817000250 Published online by Cambridge University Press

and $I_{\mathrm{h}} = (1-\delta)B$. Rewriting (3.2) yields

$$\min_{0<\delta<1, B>0} \kappa_{e,1}\left(\frac{N}{\delta B}\right) + \kappa_{e,2}\left(\frac{N^2}{\delta^2 B^2}\right) + \kappa_{h,1}\left(\frac{N}{(1-\delta)B}\right) + \kappa_{h,2}\left(\frac{N^2}{(1-\delta)^2 B^2}\right) + \gamma\left(\frac{N^2}{\delta(1-\delta)B^2}\right) + B, \tag{3.3}$$

where we refer to $\kappa_{e,1}(\kappa_{h,1})$, and $\kappa_{e,2}(\kappa_{h,2})$ as the linear and quadratic effects parameters for equipment (personnel) and to γ as the interaction parameter.

In Figure 1, we choose two sets of parameters for $(\kappa_{e,1}, \kappa_{e,2}, \kappa_{h,1}, \kappa_{h,2}, \gamma)$: (a) (10, 20, 10, 30, 10); and (b) (0.00001, 0.000002, 0.00001, 0.000003, 0.000001). The two figures show different shapes. Figure 1a shows that the optimal investment B^* increases concavely in the number of transactions N. In this scenario, each additional transaction requires an additional investment, but less so than the previous transaction (a behavior that seems to be consistent with "economies of scale"). On the other hand, in Figure 1b the optimal investment takes the shape of an S-curve, with very small increases in investments when the number of transactions is either very small or very large. Therefore, in Figure 1b we see first a convex pattern followed by a concave pattern. The S-curve appears in an environment when a small number of transactions simply do not cause operational losses of any significance; for

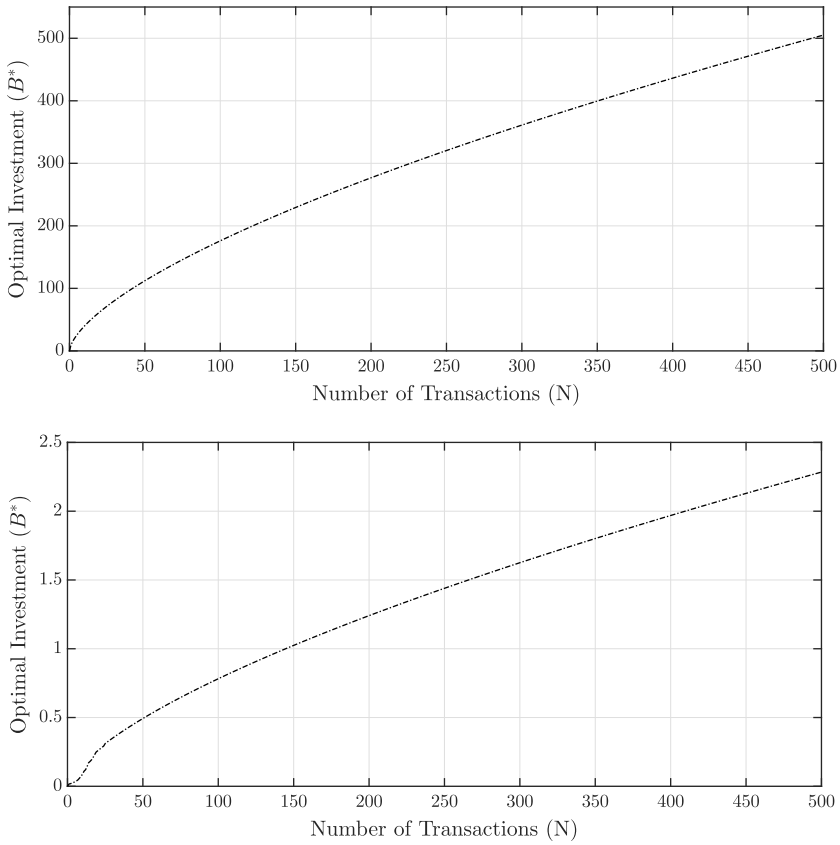


FIGURE 1. Optimal investment B^* changes with the number of transactions N.

https://doi.org/10.1017/S0269964817000250 Published online by Cambridge University Press

example, the operational losses caused by the workload are even less than the investments that are being made. In this type of an environment, the investments increase very slowly when the number of transactions is low; it remains like that until a point is reached when the increment in investment is less than the loss reduction. Note that in both Figures 1a and 1b the optimal investment increases concavely in the asymptotic regime of large N. We further quantify this growth rate in the asymptotic regime in the next section.

The results in this numerical example are consistent with Proposition 3.3, which shows that both I_e/N^2 and I_h/N^2 decrease in N. That is, neither I_e nor I_h increases as fast as N^2. Therefore, when the S shape holds, even though the optimal investments may initially increase convexly in N, they do not increase as fast as the number of transactions when N gets larger. In the next section, we study the impact of the convexity level in more detail.

Now, we show in the following corollary that if we fix the investment ratio δ, we can actually prove the sublinear relationship between the optimal investment B^* and N. In other words, B^*/N decreases in N. The proof can be found in the Appendix.

COROLLARY 3.5 (Sublinearity): *With δ fixed, the optimal investment B^* in problem (3.3) increases sublinearly in the total number of transactions N.*

This special example in Corollary 3.5 shows that if we fix the investment ratio, then the "economies of scale" property is satisfied again. This is still due to the impact of the interaction term. Without the interaction term, we can see that our investment problem is actually separable with respect to I_e and I_h, and hence the optimal investment is independent of the ratio δ.

4. A STOCHASTIC MODEL WITH POLYNOMIAL OBJECTIVE FUNCTIONS

In the previous section, we analyzed a general function (3.2) in a deterministic setting. In this section, we discuss the problem with the number of transactions N being random. We use a general polynomial function of degree i to characterize convex functions $h_e(N/I_e)$ and $h_h(N/I_h)$, specifically

$$h_e\left(\frac{N}{I_e}\right) = \left(\frac{N}{I_e}\right)^i, \quad \text{and} \quad h_h\left(\frac{N}{I_h}\right) = \left(\frac{N}{I_h}\right)^i.$$

It is easy to see here that the larger i is the more convex the loss function would be. Therefore, an advantage of using such a polynomial function is that we can easily assess the curvature level by the index i. We consider this i as the *severity level* at which the workload affects operational risk losses.

As was stated in the beginning, different business units within a financial firm are highly correlated. So operational risk losses due to an excessive workload in one business unit may affect other business units as well, resulting in operational risk losses at the other units. However, the losses in another unit may be a different function (with a different severity level) of the workload. Therefore, the total operational risk losses caused by manual and automated processes in a specific unit of a financial services firm should be represented by a combination of the operational risk losses of the unit itself and also losses caused by contagion effects at other units, that is,

$$\sum_{i=1}^{n}\left(\kappa_{e,i}\left(\frac{N}{I_e}\right)^i + \kappa_{h,i}\left(\frac{N}{I_h}\right)^i\right).$$

https://doi.org/10.1017/S0269964817000250 Published online by Cambridge University Press

The values $\kappa_{\mathrm{e},i} \geq 0$ ($\kappa_{\mathrm{h},i} \geq 0$) may depend then also on the number of business units in the firm that are affected with severity level i by an excessive equipment (personnel) overload. The more business units the firm has and the more diversified the firm is, the more terms appear in the loss function; the higher the n is and the more vulnerable or sensitive the firm is to operational risk. In what follows, in describing the impact of a larger n we may use terms such as "more vulnerable" and "more risk sensitive" interchangeably. Next, we use a general product function to characterize the interaction term of the two types of workload:

$$h_\gamma\left(\frac{N}{I_{\mathrm{e}}}, \frac{N}{I_{\mathrm{h}}}\right) = \gamma\left(\frac{N}{I_{\mathrm{e}}}\right)^{\alpha}\left(\frac{N}{I_{\mathrm{h}}}\right)^{\beta}, \tag{4.1}$$

where $\gamma > 0$ is the interaction parameter. Here $\alpha, \beta \geq 1$ capture the relative weights of the two types of workload in the interaction process, that is, a higher α implies that the equipment workload is more important in the interaction process in its contribution to the losses. Moreover, since the two types of workload are non-separable in the interaction process, we can see that if $\alpha > \beta$, it also means one unit of investment in equipment can be more efficiently converted to the loss reduction. The impact of the number of transactions should be viewed as the effect of the interaction process in its entirety on the losses.

Now, let δ denote the proportion invested in equipment, and B be the total investment. Rewriting our objective function yields the following expectation minimization problem:

$$\min_{0<\delta<1, B>0}\left\{E\left[\sum_{i=1}^{n}\left(\kappa_{\mathrm{e},i}\left(\frac{N}{\delta B}\right)^{i} + \kappa_{\mathrm{h},i}\left(\frac{N}{(1-\delta)B}\right)^{i}\right) + \gamma\left(\frac{N}{\delta B}\right)^{\alpha}\left(\frac{N}{(1-\delta)B}\right)^{\beta}\right] + B\right\}, \tag{4.2}$$

It is easy to see that (because of the positive coefficients) the larger n is, the more convex the function is and the more vulnerable the firm is. We start with an analysis of the optimal value δ^*. Let

$$\Delta_i = \left[\left(\frac{\kappa_{\mathrm{h},i}}{\kappa_{\mathrm{e},i}}\right)^{1/i+1} + 1\right]^{-1}. \tag{4.3}$$

It is a form of a weighted relative ratio of the number of business units of severity level i. We will repeatedly use this term throughout this paper. The following proposition specifies upper and lower bounds for δ^*.

PROPOSITION 4.1: *The optimal investment ratio δ^* is bounded from below and from above by $\mathcal{L}$ and $\mathcal{U}$ respectively, which are given by*

$$\mathcal{L} = \min\left\{\min_{1\leq i\leq n} \Delta_i,\ \frac{\beta}{\alpha+\beta}\right\} > 0; \tag{4.4}$$

$$\mathcal{U} = \max\left\{\max_{1\leq i\leq n} \Delta_i,\ \frac{\beta}{\alpha+\beta}\right\} < 1; \tag{4.5}$$

and we have $\mathcal{L} \leq \delta^* \leq \mathcal{U}$.

PROOF: Rewriting problem (3.3), we have

$$\begin{aligned}&\min_{0<\delta<1, B>0}\left\{E\left[\sum_{i=1}^{n}\left(\kappa_{\mathrm{e},i}\left(\frac{N}{\delta B}\right)^{i} + \kappa_{\mathrm{h},i}\left(\frac{N}{(1-\delta)B}\right)^{i}\right) + \gamma\left(\frac{N}{\delta B}\right)^{\alpha}\left(\frac{N}{(1-\delta)B}\right)^{\beta}\right] + B\right\}\\ &= \min_{0<\delta<1, B>0}\left\{\sum_{i=1}^{n} E[N^i]\left(\frac{\kappa_{e,i}}{\delta^i B^i} + \frac{\kappa_{h,i}}{(1-\delta)^i B^i}\right) + \frac{\gamma}{\delta^\alpha(1-\delta)^\beta B^{\alpha+\beta}} E[N^{\alpha+\beta}] + B\right\}.\end{aligned}$$

https://doi.org/10.1017/S0269964817000250 Published online by Cambridge University Press

Note that the minimum of the interaction term $1/\delta^{\alpha}(1-\delta)^{\beta}$ is achieved when $\delta = \beta/(\alpha+\beta)$. For any $1 \le i \le n$,

$$\frac{\kappa_{\mathrm{e},i}}{\delta^i} + \frac{\kappa_{\mathrm{h},i}}{(1-\delta)^i} \tag{4.6}$$

decreases over $0 < \delta < \Delta_i$ and increases over $\Delta_i < \delta < 1$. Therefore, it achieves its minimum at $\delta = \Delta_i$. We have now

$$\min\left\{\min_{1\le i\le n} \Delta_i, \frac{\beta}{\alpha+\beta}\right\} \le \delta^* \le \max\left\{\max_{1\le i\le n} \Delta_i, \frac{\beta}{\alpha+\beta}\right\}. \tag{4.7}$$

We can show (4.7) by contradiction. If $\delta^* > \max\{\max_{1\le i\le n} \Delta_i, \beta/(\alpha+\beta)\}$, then for each $1 \le i \le n$, function (4.6) at the value δ^* given by

$$\frac{\kappa_{\mathrm{e},i}}{\delta^{*i}} + \frac{\kappa_{\mathrm{h},i}}{(1-\delta^*)^i}$$

is greater than the corresponding local minimum value achieved at Δ_i. Therefore, the objective function at this δ^* cannot be a global minimum, which is a contradiction. A similar contradiction holds for

$$\delta^* < \min\left\{\min_{1\le i\le n} \Delta_i, \frac{\beta}{\alpha+\beta}\right\}.$$

Therefore, we claim that inequality (4.7) holds here.

The upper and lower bounds here are given by (5) and (6). Moreover, since for each $1 \le i \le n$ the functions (5) and (6) lie between 0 and 1, we conclude that δ^* is bounded away from the boundary $(0, 1)$, and is independent of the variables B and N. ■

Proposition 4.1 specifies upper and lower bounds for the optimal investment ratio that are independent of the number of transactions N and are identified by the relative value of the coefficients. This result guarantees the bounded behavior of the optimal δ^* when N goes to infinity. Therefore, it provides us with a basis for analyzing the behavior of the optimal B^* in the asymptotic regime $E[N^n] \to \infty$ (note that the nth moment is of interest when the degree of the polynomial function is n) and $E[N] \to 0$, and in what follows we refer to it as the *heavy trading environment* and *light trading environment*, respectively. We show the asymptotic results in the following two subsections.

4.1. Heavy Trading Asymptotics

We consider in this subsection the asymptotics when $E[N^n] \to \infty$. We first have the following proposition.

Proposition 4.2: *The optimal investment B^* satisfies the following asymptotics:*

$$B^* \sim \begin{cases} O\left(\left(E[N^n]\right)^{1/n+1}\right) & \text{for } n > \alpha+\beta \text{ as } E[N^n] \to \infty \\ O\left(\left(E[N^{\alpha+\beta}]\right)^{1/\alpha+\beta+1}\right) & \text{for } n \le \alpha+\beta \text{ as } E[N^{\alpha+\beta}] \to \infty \end{cases}. \tag{4.8}$$

https://doi.org/10.1017/S0269964817000250 Published online by Cambridge University Press

PROOF: First, we define the function

$$F(\delta, B) = \sum_{i=1}^{n} E[N^i] \left(\frac{\kappa_{e,i}}{\delta^i B^i} + \frac{\kappa_{h,i}}{(1-\delta)^i B^i} \right) + \frac{\gamma}{\delta^\alpha (1-\delta)^\beta B^{\alpha+\beta}} E[N^{\alpha+\beta}] + B. \tag{4.9}$$

At its minimum, we have

$$\left.\frac{\partial F}{\partial \delta}\right|_{(\delta^*, B^*)} = 0, \qquad \left.\frac{\partial F}{\partial B}\right|_{(\delta^*, B^*)} = 0,$$

which implies that

$$\begin{aligned} &\sum_{i=1}^{n} E[N^i] \left(-\frac{i\kappa_{e,i}}{(\delta^*)^{i+1}(B^*)^i} + \frac{i\kappa_{h,i}}{(1-\delta^*)^{i+1}(B^*)^i} \right) \\ &\quad + \left(-\frac{1}{\alpha(\delta^*)^{\alpha+1}(1-\delta^*)^\beta} + \frac{1}{\beta(\delta^*)^\alpha (1-\delta^*)^{\beta+1}} \right) \frac{\gamma}{(B^*)^{\alpha+\beta}} E[N^{\alpha+\beta}] = 0, \end{aligned} \tag{4.10}$$

$$\begin{aligned} &-\sum_{i=1}^{n} E[N^i] \left(\frac{i\kappa_{e,i}}{(\delta^*)^i (B^*)^{i+1}} + \frac{i\kappa_{h,i}}{(1-\delta^*)^i (B^*)^{i+1}} \right) \\ &\quad - \frac{\gamma(\alpha+\beta)}{\delta^*(1-\delta^*)(B^*)^{\alpha+\beta+1}} E[N^{\alpha+\beta}] + 1 = 0. \end{aligned} \tag{4.11}$$

Due to the impact of the interaction term, we consider two cases: $n > \alpha + \beta$ and $n \le \alpha + \beta$. Consider first $n > \alpha + \beta$. From (4.11) and Proposition 4.1, there exist constants C_1, C_2 that depend only on $(\kappa_{e,i})_{i=1}^n, (\kappa_{h,i})_{i=1}^n, \gamma$ such that

$$0 < C_1 \le \frac{E[N^n]}{(B^*)^{n+1}} + \frac{E[N^{n-1}]}{(B^*)^n} + \cdots + \frac{E[N^2]}{(B^*)^3} + \frac{E[N]}{(B^*)^2} \le C_2 < \infty. \tag{4.12}$$

By (4.12), we have $C_2 \ge E[N^n]/(B^*)^{n+1}$ and hence $B^* \ge C_2^{-(1/n+1)} (E[N^n])^{1/n+1}$. It is clear that $B^* \to \infty$ as $E[N^n] \to \infty$. By the monotonicity property of the L^p norm, we have for any $1 \le i \le n-1$,

$$\left(E[N^i]\right)^{1/i} \le \left(E[N^n]\right)^{1/n}, \tag{4.13}$$

and thus from (4.12), we have

$$C_1 \le \frac{E[N^n]}{(B^*)^{n+1}} + \frac{(E[N^n])^{(n-1)/n}}{(B^*)^n} + \cdots + \frac{(E[N^n])^{2/n}}{(B^*)^3} + \frac{(E[N^n])^{1/n}}{(B^*)^2}.$$

Moreover, since it is easy to verify that $B^* \ge C_2^{-(1/n+1)} (E[N^n])^{1/n+1}$ and we are interested in $E[N^n] \to \infty$, we can assume that $B^* \ge 1$. Hence, for any $1 \le i \le n-1$,

$$\frac{(E[N^n])^{i/n}}{(B^*)^{i+1}} \le \left(\frac{E[N^n]}{(B^*)^{n+1}} \right)^{i/n}.$$

Therefore, we get

$$C_1 \le \frac{E[N^n]}{(B^*)^{n+1}} + \sum_{i=1}^{n-1} \left(\frac{E[N^n]}{(B^*)^{n+1}} \right)^{i/n}.$$

This implies that $E[N^n]/(B^*)^{n+1}$ must be bounded from below by some constant, and hence $B^* \le C_3 (E[N^n])^{1/n+1}$ for some constant C_3.

https://doi.org/10.1017/S0269964817000250 Published online by Cambridge University Press

In our second case, when $n \leq \alpha + \beta$, equation (4.12) can be replaced by

$$0 < C_1 \leq \frac{E[N^{\alpha+\beta}]}{(B^*)^{\alpha+\beta+1}} + \cdots + \frac{E[N]}{(B^*)^2} \leq C_2 < \infty,$$

and by a similar analysis, we can show that $B^* \sim O((E[N^{\alpha+\beta}])^{1/\alpha+\beta+1})$ as $E[N^{\alpha+\beta}] \to \infty$. ■

This proposition quantifies the scale of the necessary investment B^* in a heavy trading environment. Moreover, it is easy to see that if in this asymptotic regime $E[N^n]$ is of the order of power n, the larger the n the more investment is needed. This condition actually holds for most distributions as we will see in the next section. As we discussed before, the n reflects the firm's vulnerability towards operational risks. This result actually shows us that in general, the more vulnerable firms need larger investments in a heavy trading environment, but the scale of the necessary budget increases concavely in the asymptotic regime.

In the remaining part of this section we analyze the asymptotic behavior of the optimal investment ratio δ^* that determines the allocation rule. From Proposition 4.1, we know that δ^* is bounded from below and above. The next proposition shows the asymptotic allocation rule for a heavy trading environment.

PROPOSITION 4.3: *If $E[N^n] \to \infty$, then the optimal investment ratio δ^* converges to the limits*

$$\delta^* \to \begin{cases} \Delta_n & \text{if } n > \alpha + \beta \\ \eta & \text{if } n = \alpha + \beta, \\ \dfrac{\beta}{\alpha + \beta} & \text{if } n < \alpha + \beta \end{cases} \tag{4.14}$$

where η is the unique solution in the interval $(0, 1)$ to the equation

$$-\frac{2\kappa_{e,2}}{\eta^{\alpha+\beta+1}} + \frac{2\kappa_{h,2}}{(1-\eta)^{\alpha+\beta+1}} - \gamma\left(\frac{1}{\alpha\eta^{\alpha+1}(1-\eta)^{\beta}} - \frac{1}{\beta\eta^{\alpha}(1-\eta)^{\beta+1}}\right) = 0. \tag{4.15}$$

PROOF: We use again the function F as defined in (4.9). At the minimum, we have $\frac{\partial F}{\partial \delta}|_{(\delta^*, B^*)} = 0$, that is,

$$\sum_{i=1}^{n} E[N^i]\left(-\frac{i\kappa_{e,i}}{(\delta^*)^{i+1}(B^*)^i} + \frac{i\kappa_{h,i}}{(1-\delta^*)^{i+1}(B^*)^i}\right)$$
$$+\left(-\frac{1}{\alpha(\delta^*)^{\alpha+1}(1-\delta^*)^{\beta}} + \frac{1}{\beta(\delta^*)^{\alpha}(1-\delta^*)^{\beta+1}}\right)\frac{\gamma}{(B^*)^{\alpha+\beta}}E[N^{\alpha+\beta}] = 0.$$

Rearranging the terms yields

$$\frac{E[N^n]}{(B^*)^n}\left(-\frac{n\kappa_{e,n}}{(\delta^*)^{n+1}} + \frac{n\kappa_{h,n}}{(1-\delta^*)^{n+1}}\right) \tag{4.16}$$
$$= \sum_{i=1}^{n-1}\frac{E[N^i]}{(B^*)^i}\left(\frac{i\kappa_{e,i}}{(\delta^*)^{i+1}} - \frac{i\kappa_{h,i}}{(1-\delta^*)^{i+1}}\right)$$
$$+\left(\frac{1}{\alpha(\delta^*)^{\alpha+1}(1-\delta^*)^{\beta}} - \frac{1}{\beta(\delta^*)^{\alpha}(1-\delta^*)^{\beta+1}}\right)\frac{\gamma E[N^{\alpha+\beta}]}{(B^*)^{\alpha+\beta}}. \tag{4.17}$$

Consider the case $n > \alpha + \beta$. For any i, we have $E[N^i] \leq (E[N^n])^{i/n}$ from the monotonicity of the L^p norm. As $E[N^n] \to \infty$, we have by Proposition 4.2 that $B^* \sim O((E[N^n])^{1/n+1})$.

https://doi.org/10.1017/S0269964817000250 Published online by Cambridge University Press

Therefore, from (4.25) and Proposition 4.1, there exist some constants $K_1, K_2 > 0$, so that

$$\begin{aligned}\left|\frac{-n\kappa_{e,n}}{(\delta^*)^{n+1}} + \frac{n\kappa_{h,n}}{(1-\delta^*)^{n+1}}\right| &\le K_1 \frac{(B^*)^n}{E[N^n]} \sum_{i=1}^{n-1} \frac{E[N^i]}{(B^*)^i} \\ &\le K_1 \frac{(B^*)^n}{E[N^n]} \sum_{i=1}^{n-1} \frac{(E[N^n])^{i/n}}{(B^*)^i} \\ &\le K_2 \frac{(E[N^n])^{n/n+1}}{E[N^n]} \sum_{i=1}^{n-1} (E[N^n])^{(i/n)-(i/n+1)} \\ &\le (n-1)K_2 \frac{(E[N^n])^{n/n+1}}{E[N^n]} (E[N^n])^{(n-1)/n(n+1)} \\ &= (n-1)K_2 (E[N^n])^{(-1)/n(n+1)} \to 0,\end{aligned}$$

as $E[N^n] \to \infty$. Therefore, the LHS converges to zero and

$$\delta^* \to \Delta_n, \quad \text{as} \quad E[N^n] \to \infty.$$

When $n = \alpha + \beta$, following the same arguments as before, we can show that there exists a $\delta = \eta$ such that

$$-\frac{2\kappa_{e,2}}{\eta^{\alpha+\beta+1}} + \frac{2\kappa_{h,2}}{(1-\eta)^{\alpha+\beta+1}} - \gamma\left(\frac{1}{\alpha\eta^{\alpha+1}(1-\eta)^{\beta}} - \frac{1}{\beta\eta^{\alpha}(1-\eta)^{\beta+1}}\right) \to 0,$$

as $E[N^n] \to \infty$. Define

$$G(\eta) = -\frac{2\kappa_{e,2}}{\eta^{\alpha+\beta+1}} + \frac{2\kappa_{h,2}}{(1-\eta)^{\alpha+\beta+1}} - \gamma\left(\frac{1}{\alpha\eta^{\alpha+1}(1-\eta)^{\beta}} - \frac{1}{\beta\eta^{\alpha}(1-\eta)^{\beta+1}}\right).$$

Then $G(0) = -\infty$ and $G(1) = \infty$ and $G'(\delta) > 0$. Therefore, there exists a unique solution η^* on $(0,1)$ to the equation $G(\eta) = 0$ and $\delta^* \to \eta^*$ as $E[N^n] \to \infty$.

Finally, when $n < \alpha + \beta$, following the same arguments as before, we can show that

$$-\gamma\left(\frac{1}{\alpha(\delta^*)^{\alpha+1}(1-\delta^*)^{\beta}} - \frac{1}{\beta(\delta^*)^{\alpha}(1-\delta^*)^{\beta+1}}\right) \to 0,$$

as $E[N^n] \to \infty$ and thus $\delta^* \to \beta/(\alpha+\beta)$. ■

Proposition 4.3 shows that the optimal δ^* in a heavy trading environment, will always converge to a fixed ratio. It is easy to see that if n becomes very large, that is, $n \to \infty$, the δ^* converges to $1/2$. Since a larger n represents a more vulnerable firm, we can see from Proposition 4.3 that when the market is highly volatile it is optimal to have balanced investments in infrastructure and personnel at financial firms that are sensitive to operational risks. A more unbalanced investment, in a highly liquid market, could cause large losses at risk sensitive firms.

More practically speaking, it could bring about larger operational risk losses when risk sensitive firms expand fast in personnel or in technology in a highly volatile market. For example, Knight Capital's glitch happened when they were using their new trading software at exactly the same time when the NYSE Euronext (NYX) was rolling out its new Retail Liquidity Program. The losses would have been smaller if more people would have been hired to monitor the program more closely.

https://doi.org/10.1017/S0269964817000250 Published online by Cambridge University Press

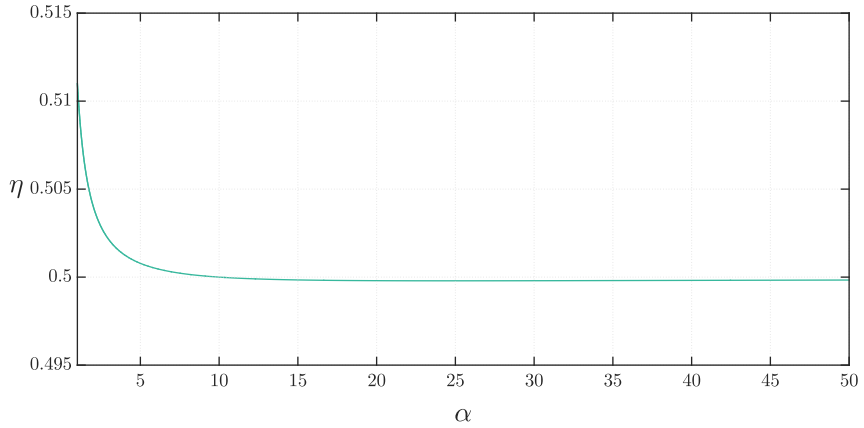


FIGURE **2.** η as a function of α.

We further conduct a numerical analysis to study equation (4.15) to better understand how η depends on α and β. We set $\kappa_{e,2} = 10$, $\kappa_{h,2} = 10$, $\gamma = 10$, $\beta = 10$, and let α vary from 1 to 50. Similarly, in the second experiment, we set $\alpha = 10$, and let β vary from 1 to 50. From Figure 2 we can see that η (the asymptotic δ^*) would decrease as α increases, and then would converge to 0.5 as α becomes very large. The initial decrease is due to the fact that when α increases, we know that the equipment workload plays a more and more important role in the interaction term, and hence the investment in equipment also has a larger impact on the loss reduction. However, when α increases $n = \alpha + \beta$ also increases, and therefore, the losses caused by both equipment and personnel workload increase. In order to balance the risk losses, we also need to invest more in personnel in order to reduce the personnel workload. Therefore, when α becomes very large we would have $\eta \to 0.5$. We can see a similar result in Figure 3. Initially when β increases we would have the investment in equipment increase, and thus the investment in personnel decrease, and then when β becomes very large we have $\eta \to 0.5$.

Finally, we analyze the tail effects in a heavy trading environment. In order to do so, we analyze three of the most commonly used distributions for the transaction volume N, namely the Exponential distribution, the lognormal distribution, and the Pareto distribution. We

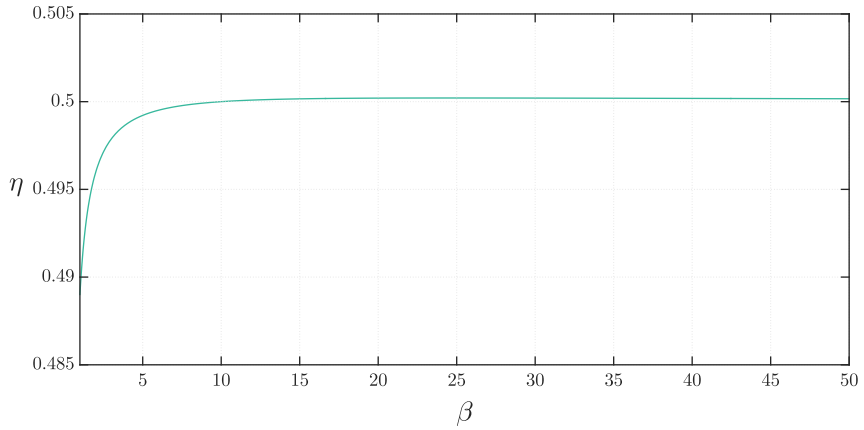


FIGURE **3.** η as a function of β.

https://doi.org/10.1017/S0269964817000250 Published online by Cambridge University Press

analyze the impact of their tails on the optimal investments. To simplify our notation, we consider the special case where $\alpha = \beta = 1$. However, the results can be easily applied to the general case as well. The asymptotic behavior of the B^* are then quantified as functions of the parameters that reflect the heaviness of the tails of the corresponding distributions, that is, the rate parameter λ of the Exponential, the variance σ^2 of the Lognormal, and the shape parameter α of the Pareto. Given the same means and variances, the Pareto distribution is known to have a heavier tail than the Lognormal, and the Lognormal is known to have a heavier tail than the Exponential. Therefore, in the remainder of this section, we control the mean and variance of the three distributions, and conduct numerical experiments to compare the impact of the tails on the optimal investments. We start our discussion with the Exponential distribution.

Example 4.4 (Exponential Distribution): If N, rate $\lambda > 0$, then the nth moment is given by

$$E[N^n] = \frac{n!}{\lambda^n}.$$

As $\lambda \to 0$, $E[N^n] \to \infty$. Hence, from Proposition 4.2, it follows that

$$B^* \sim \begin{cases} O\left(\left(\frac{1}{\lambda}\right)^{n/n+1}\right) & \text{for } n > 2 \text{ as } \lambda \downarrow 0 \\ O\left(\left(\frac{1}{\lambda}\right)^{2/3}\right) & \text{for } n \leq 2 \text{ as } \lambda \downarrow 0 \end{cases}. \tag{4.18}$$

Example 4.5 (Lognormal Distribution): If N follows a Lognormal distribution, that is, $lnN \sim N(\mu, \sigma^2)$, then the nth moment is given by

$$E[N^n] = e^{n\mu + (1/2)n^2\sigma^2}.$$

As $\sigma^2 \to \infty$, $E[N^n] \to \infty$. Hence, from Proposition 4.2 it follows that

$$B^* \sim \begin{cases} O\left(e^{\frac{1}{2}\frac{n^2}{n+1}\sigma^2}\right) & \text{for } n > 2 \text{ as } \sigma \to \infty \\ O\left(e^{\frac{2}{3}\sigma^2}\right) & \text{for } n \leq 2 \text{ as } \sigma \to \infty. \end{cases} \tag{4.19}$$

Example 4.6 (Pareto Distribution): Consider now the Pareto distribution (Type I), which has a polynomial decay in its tail probability. The survival function of the Pareto random variable X can be written as

$$\bar{F}(x) = Pr(X > x) = \begin{cases} \left(\frac{x_m}{x}\right)^{\alpha} & \text{for } x \geq x_m \\ 1 & \text{for } x < x_m \end{cases}.$$

Here $x_m > 0$ is the scale parameter and $\alpha > 0$ is the shape parameter, which is known as the *tail index*. We have $P(X \geq x) \sim O(x^{-\alpha})$ as $x \to \infty$. It is easy to see that when $\alpha \leq n$,

https://doi.org/10.1017/S0269964817000250 Published online by Cambridge University Press

$E[N^n] = \infty$. Therefore, as $\alpha \downarrow n$, $E[N^n] \to \infty$, and one can show that

$$E[N^i] \sim O(1), \qquad \text{for any } 1 \le i \le n-1 \text{ as } \alpha \downarrow n,$$

and that

$$E[N^n] \sim O\left(\frac{1}{\alpha - n}\right), \quad \text{as} \quad \alpha \downarrow n.$$

Hence, by Proposition 4.2 we have,

$$B^* \sim \begin{cases} O\left(\left(\frac{1}{\alpha - n}\right)^{1/n+1}\right) & \text{for } n > 2 \text{ as } \alpha \downarrow n \\ O\left(\left(\frac{1}{\alpha - 2}\right)^{1/3}\right) & \text{for } n \le 2 \text{ as } \alpha \downarrow 2 \end{cases}. \tag{4.20}$$

For the three types of distributions discussed above, a heavier tail would correspond to a smaller rate parameter λ for the Exponential distribution, a larger variance σ^2 for the Lognormal distribution, and a smaller shape parameter α for the Pareto distribution; and would require thus, in all these cases, a larger investment. Also, a more vulnerable financial firm (i.e., a larger n) requires a larger investment. Moreover, the asymptotic functions with a larger n always dominate asymptotic functions with small n, which implies that the tail effects increase with n. In other words, the tail events have a more significant impact on risk sensitive firms. Therefore, risk sensitivity or vulnerability exacerbates the effects of the right tail.

To illustrate, we run numerical experiments with N having a Pareto distribution, and get the optimal B^* in our objective function

$$\min_{0<\delta<1, B>0} \left\{ E\left[\sum_{i=1}^{n} \left(\kappa_{e,i}\left(\frac{N}{\delta B}\right)^i + \kappa_{h,i}\left(\frac{N}{(1-\delta)B}\right)^i\right) + \gamma \frac{N}{\delta B}\frac{N}{(1-\delta)B}\right] + B \right\},$$

for $n = 1, 2, 3, 4$, respectively, and parameters $(\kappa_{e,1}, \kappa_{h,1}, \kappa_{e,2}, \kappa_{h,2}, \kappa_{e,3}, \kappa_{h,3}, \kappa_{e,4}, \kappa_{h,4}, \gamma)$ chosen as $(30, 50, 60, 50, 70, 60, 80, 60, 50)$.

Figure 4 shows the impact of the tail index α of a Pareto distribution on the optimal investment B^*. In this figure, the heaviness of the tail is characterized by $C - \alpha$, and C is a

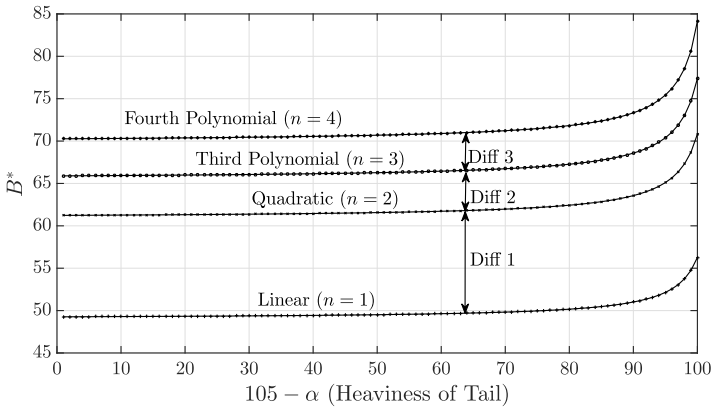


FIGURE 4. The tail effects with Pareto distribution.

https://doi.org/10.1017/S0269964817000250 Published online by Cambridge University Press

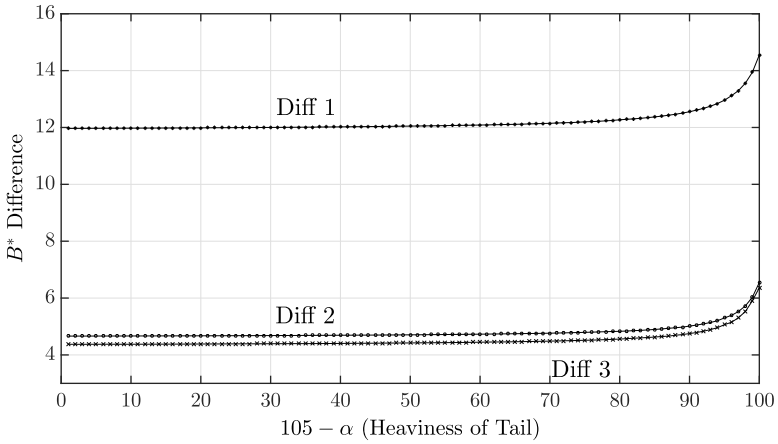


FIGURE **5.** Tail effects and vulnerability interaction.

constant chosen to be 105. By definition, we know that the smaller the α is, the heavier the tail is. Therefore, the larger $C - \alpha$ is, the heavier the tail. Moreover, we choose the scale parameter $x_m = 10$, and change the integer α from 104 to 5. We can see that the larger n is, the more investment is needed. Also, the optimal investment increases rapidly when the tail becomes heavier. However, in this example the interaction effects between the vulnerability n and the heaviness of the tail are not that clear. Therefore, we compute the differences between two subsequent polynomial function (as marked in Figure 4) and show in Figure 5 that the differences are increasing when the heaviness of tail increases. We also show the numerical values of the differences in Table B.1 of Appendix B. This indeed shows that tail effects and vulnerability reinforce one another. We furthermore present another numerical experiment to show the interaction effects in Appendix C.

In general, we find that if the total operational losses increase convexly in the number of transactions, it will in general be advantageous for firms to invest more when the distribution of the number of transactions has a heavier tail. This logic is quite clear, as the convexity implies that a more volatile number of transactions (i.e., a heavier tail) results in larger expected losses than a less volatile number of transactions. Convexity in the loss function is quite plausible for some firms, particularly those who are specializing in high technology trading strategies and are prone to system crashes in heavy trading.

4.2. Light Trading Asymptotics

In this subsection, we analyze the *light trading environment*, which is defined as $E[N^n] \to 0$ and reflects the impact of the left tail. The budget allocation decisions in this environment are of particular importance when the transaction level is low, which occurs often when a firm is newly established. We compare the asymptotics of the investments in this environment with those in the heavy trading environment, and find that the behavior of the investment decisions are more diverse in this environment.

In Section 4.1, in the heavy trading environment, we consider the asymptotic behaviors of B^* and δ^* when $E[N^n] \to \infty$. In this section, we analyze the asymptotics of B^* and δ^* when $E[N^n] \to 0$. Before going into the analysis of the asymptotics, note that in the previous sections we had to focus on $E[N^n] \to \infty$ rather than on $E[N] \to \infty$. Because of

https://doi.org/10.1017/S0269964817000250 Published online by Cambridge University Press

the monotonicity of the L^p norm, we know that

$$E[N] \leq E[N^n]^{1/n}$$

and if $E[N] \to \infty$, then $E[N^n] \to \infty$. So in a heavy trading environment, we only need to consider the latter case. However, in a light trading regime, we have the opposite relationship. If $E[N^n] \to 0$, then $E[N] \to 0$. Hence we consider the situation $E[N] \to 0$ instead.

Moreover, in a light trading environment, some additional assumptions are needed. The reason is that we need to compare $E[N^n]$ and $E[N]$ as $E[N] \to 0$, because in terms of their relative scales, any scenario is possible. For instance, consider $N = \epsilon X$, where X is a positive random variable with all moments being finite. Then, it is easy to see that $E[N^n] \sim O(\epsilon^n)$ as $\epsilon \to 0$. On the other hand, consider $N = \epsilon X$, where X is a positive random variable with finite mean but $E[X^n] = \infty$ for any $n \geq 2$. Then $E[N] \sim O(\epsilon)$ as $\epsilon \to 0$ and $E[N^n] = \infty$ for any $n \geq 2$. Therefore, we can see that the behavior of optimal investments would be more diverse in a light trading regime.

Assume now that all the moments of N are finite. We consider two scenarios: (i) all the moments of N are of the same order; or (ii) the L^n norms of N are of the order of power n.

ASSUMPTION 4.7: *Assume that for any $n \in \mathbb{N}$, there are positive constants c_n and C_n so that for any $E[N]$ sufficiently small, we have*

$$c_n E[N] \leq E[N^n] \leq C_n E[N].$$

This condition appears often when the distribution has a thin left tail, for example, the Normal distribution. In other words, this is the case when the probability of a low market demand is not too high. We now characterize the asymptotics in this setting.

PROPOSITION 4.8: *Under Assumption 4.7, the optimal B^* and δ^* satisfy the following asymptotics:*

$$B^* \sim \begin{cases} O\left((E[N])^{1/n+1}\right) & \text{for } n > \alpha + \beta \\ O\left((E[N])^{1/\alpha+\beta+1}\right) & \text{for } n \leq \alpha + \beta \end{cases} \quad \text{as} \quad E[N] \to 0, \tag{4.21}$$

and

$$\delta^* \to \begin{cases} \Delta_n & \text{if } n > 2 \\ \eta & \text{if } n = \alpha + \beta \\ \dfrac{\beta}{\alpha + \beta} & \text{if } n < \alpha + \beta, \end{cases} \tag{4.22}$$

where η is the unique solution in the interval $(0, 1)$ to the equation

$$-\frac{2\kappa_{e,2}}{\eta^{\alpha+\beta+1}} + \frac{2\kappa_{h,2}}{(1-\eta)^{\alpha+\beta+1}} - \gamma\left(\frac{1}{\alpha\eta^{\alpha+1}(1-\eta)^{\beta}} - \frac{1}{\beta\eta^{\alpha}(1-\eta)^{\beta+1}}\right) = 0. \tag{4.23}$$

https://doi.org/10.1017/S0269964817000250 Published online by Cambridge University Press

PROOF: Note that the bounds in Proposition 4.1 for δ^* still hold. Again, let

$$F(\delta, B) := \sum_{i=1}^{n} E[N^i] \left(\frac{\kappa_{e,i}}{\delta^i B^i} + \frac{\kappa_{h,i}}{(1-\delta)^i B^i} \right) + \frac{\gamma}{\delta^\alpha (1-\delta)^\beta B^{\alpha+\beta}} E[N^{\alpha+\beta}] + B.$$

At the minimum, $\partial F/\partial B = 0$, which implies

$$- \sum_{i=1}^{n} E[N^i] \left(\frac{i\kappa_{e,i}}{(\delta^*)^i (B^*)^{i+1}} + \frac{i\kappa_{h,i}}{(1-\delta^*)^i (B^*)^{i+1}} \right) - \frac{\gamma(\alpha+\beta)}{\delta^*(1-\delta^*)(B^*)^{\alpha+\beta+1}} E[N^{\alpha+\beta}] + 1 = 0. \tag{4.24}$$

Consider first the case $n > \alpha + \beta$. By Assumption 4.7, Proposition 4.1 and (4.24), there exist constants m and M so that

$$0 < m \le \frac{E[N]}{(B^*)^{n+1}} + \frac{E[N]}{(B^*)^n} + \cdots + \frac{E[N]}{(B^*)^2} \le M < \infty.$$

When $E[N] \to 0$, we have $B^* \to 0$ and for any $B^* \le 1$,

$$\frac{E[N]}{(B^*)^{n+1}} \le \frac{E[N]}{(B^*)^{n+1}} + \frac{E[N]}{(B^*)^n} + \cdots + \frac{E[N]}{(B^*)^2} \le n \frac{E[N]}{(B^*)^{n+1}},$$

which implies that $B^* \sim O((E[N])^{1/n+1})$ as $E[N] \to 0$.

Next, consider the case $n \le \alpha + \beta$. Again, by Assumption 4.7, Proposition 4.1 and (4.24), there exist constants m and M so that

$$0 < m \le \frac{E[N]}{(B^*)^2} + \cdots + \frac{E[N]}{(B^*)^{\alpha+\beta+1}} \le M < \infty,$$

which implies that $B^* \sim O((E[N])^{1/\alpha+\beta+1})$ as $E[N] \to 0$.

At the minimum, $\partial F/\partial \delta|_{(\delta^*, B^*)} = 0$, that is,

$$\sum_{i=1}^{n} E[N^i] \left(-\frac{i\kappa_{\mathrm{e},i}}{(\delta^*)^{i+1}(B^*)^i} + \frac{i\kappa_{\mathrm{h},i}}{(1-\delta^*)^{i+1}(B^*)^i} \right) + \left(-\frac{1}{\alpha(\delta^*)^{\alpha+1}(1-\delta^*)^\beta} + \frac{1}{\beta(\delta^*)^\alpha (1-\delta^*)^{\beta+1}} \right) \frac{\gamma}{(B^*)^{\alpha+\beta}} E[N^{\alpha+\beta}] = 0.$$

Rearranging the terms yields

$$\frac{E[N^n]}{(B^*)^n} \left(-\frac{n\kappa_{\mathrm{e},n}}{(\delta^*)^{n+1}} + \frac{n\kappa_{\mathrm{h},n}}{(1-\delta^*)^{n+1}} \right) \tag{4.25}$$

$$= \sum_{i=1}^{n-1} \frac{E[N^i]}{(B^*)^i} \left(\frac{i\kappa_{\mathrm{e},i}}{(\delta^*)^{i+1}} - \frac{i\kappa_{\mathrm{h},i}}{(1-\delta^*)^{i+1}} \right) + \left(\frac{1}{\alpha(\delta^*)^{\alpha+1}(1-\delta^*)^\beta} - \frac{1}{\beta(\delta^*)^\alpha (1-\delta^*)^{\beta+1}} \right) \frac{\gamma E[N^{\alpha+\beta}]}{(B^*)^{\alpha+\beta}}. \tag{4.26}$$

When $n > \alpha + \beta$, by Assumption 4.7, as $E[N] \to 0$, we have $B^* \to 0$ and therefore

$$-\frac{n\kappa_{\mathrm{e},n}}{(\delta^*)^{n+1}} + \frac{n\kappa_{\mathrm{h},n}}{(1-\delta^*)^{n+1}} \to 0,$$

which yields $\delta^* \to \Delta_n$.

https://doi.org/10.1017/S0269964817000250 Published online by Cambridge University Press

When $n = \alpha + \beta$, following similar arguments, we have

$$-\frac{2\kappa_{\mathrm{e},2}}{\eta^{\alpha+\beta+1}} + \frac{2\kappa_{\mathrm{h},2}}{(1-\eta)^{\alpha+\beta+1}} - \gamma\left(\frac{1}{\alpha\eta^{\alpha+1}(1-\eta)^{\beta}} - \frac{1}{\beta\eta^{\alpha}(1-\eta)^{\beta+1}}\right) \to 0.$$

By the same arguments as in the proof of Proposition 4.3, we have $\delta^* \to \eta$, where η is the unique solution in $(0,1)$ to the equation

$$-\frac{2\kappa_{\mathrm{e},2}}{\eta^{\alpha+\beta+1}} + \frac{2\kappa_{\mathrm{h},2}}{(1-\eta)^{\alpha+\beta+1}} - \gamma\left(\frac{1}{\alpha\eta^{\alpha+1}(1-\eta)^{\beta}} - \frac{1}{\beta\eta^{\alpha}(1-\eta)^{\beta+1}}\right) = 0.$$

Finally, following a similar argument, it is easy to prove that if $n < \alpha + \beta$, then $\delta^* \to \beta/\alpha + \beta$. ■

In this case, when all the moments of N are of the same order, we find that the asymptotic allocation rules in the heavy and light trading environments are the same. Moreover, both the asymptotic allocation rule and the investment asymptotics depend on the risk sensitivity of the firm, that is, the n still plays an important role. We can see that when n increases, we would need larger investments in the asymptotic regime of a light trading environment.

Next, we study the scenario when the L^n norms of N are of the order of power n.

Assumption 4.9: *Assume that for any $n \in \mathbb{N}$, there are positive constants c_n and C_n so that for any $E[N]$ sufficiently small, we have*

$$c_n(E[N])^n \le E[N^n] \le C_n(E[N])^n.$$

This condition appears often when the distribution has a heavy left tail, that is, the exponential distribution. In other words, this is the case when the probability of a low market demand is quite high.

Proposition 4.10: *Under Assumption 4.9, the optimal B^* and δ^* satisfy the following asymptotics:*

$$B^* \sim O\left(\sqrt{E[N]}\right), \quad \text{and} \quad \delta^* \to \Delta_1, \quad \text{as} \quad E[N] \to 0. \tag{4.27}$$

Proof: First, by Assumption 4.9, Proposition 4.1 and (4.24), there exist some constants m and M so that

$$0 < m \le \frac{(E[N])^n}{(B^*)^{n+1}} + \frac{(E[N])^{n-1}}{(B^*)^n} + \cdots + \frac{E[N]}{(B^*)^2} \le M < \infty,$$

which implies that $B^* \sim O(\sqrt{E[N]})$ as $E[N] \to 0$. It is straightforward to see that this result is independent of the relationship between n and $\alpha + \beta$. Next, recall that at the

https://doi.org/10.1017/S0269964817000250 Published online by Cambridge University Press

minimum, we have

$$\sum_{i=1}^{n} E[N^i]\left(-\frac{i\kappa_{\mathrm{e},i}}{(\delta^*)^{i+1}(B^*)^i}+\frac{i\kappa_{\mathrm{h},i}}{(1-\delta^*)^{i+1}(B^*)^i}\right)$$
$$+\left(-\frac{1}{\alpha(\delta^*)^{\alpha+1}(1-\delta^*)^{\beta}}+\frac{1}{\beta(\delta^*)^{\alpha}(1-\delta^*)^{\beta+1}}\right)\frac{\gamma}{(B^*)^{\alpha+\beta}}E[N^{\alpha+\beta}]=0.$$

Note that $E[N^i]\sim O((E[N])^i)$ and $B^*\sim O(\sqrt{E[N]})$ as $E[N]\to 0$. Hence, we conclude that

$$-\frac{\kappa_{\mathrm{e},1}}{(\delta^*)^2}+\frac{\kappa_{\mathrm{h},1}}{(1-\delta^*)^2}\to 0,$$

which implies that, as $E[N]\to 0$, $\delta^*\to\Delta_1$. ■

Under Assumption 4.9 we find that in a light trading environment the asymptotic behavior of the optimal investment is independent of n. Firms with different levels of vulnerability will invest therefore with the same ratio, which depends only on the relative number of business units that have the lowest severity level. Intuitively, a higher workload in a specific business unit will cause larger losses at the unit itself, and smaller losses at units that have limited business interactions with the given unit. Proposition 4.10 shows us that the investment ratio may actually only depend on the number of more distant business units, and the investment decisions rely more on the contagion effects among business units in a light trading regime.

5. CONCLUDING REMARKS

We consider a resource allocation model for analyzing investment strategies to minimize expected operational risk losses. We consider four characteristic properties of financial services firms: the randomness of the transaction volumes, the positive correlations between the different business units of a firm, the relatively high probabilities of extreme events, and the severe potential damages that can be caused by automation of processes. Therefore, we study the losses caused by operational risk events that are either due to excessive stress on equipment and systems or due to an excessive workload on personnel, and introduce a functional form of the operational risk losses of the firm that is characterized by the number of transactions and the amounts being invested. Taking into account that financial markets are highly volatile, our analysis focuses on the impact of risk sensitivity and heavy tails on the optimal investments in asymptotic regimes. In general, we expect our model to provide some guidelines with regard to budget allocations in financial institutions, especially taking into consideration market uncertainties and trading activity levels.

Our initial analysis is based on a general convex loss function in a deterministic setting. We find that in this general case, the optimal investment exhibits a certain monotonicity property with regard to the number of transactions. Also, because of the effects of the interaction process, in contrast to manufacturing settings, "economies of scale" effects do not play a role in our setting. Moreover, we show for a special quadratic loss function that the optimal investment increases in the number of transactions either concavely or according to an S-curve.

In a stochastic setting, the number of transactions is a random variable from an arbitrary distribution of which the moments are known. We propose a general polynomial loss function of degree n and quantify the losses of different business units according to different severity

https://doi.org/10.1017/S0269964817000250 Published online by Cambridge University Press

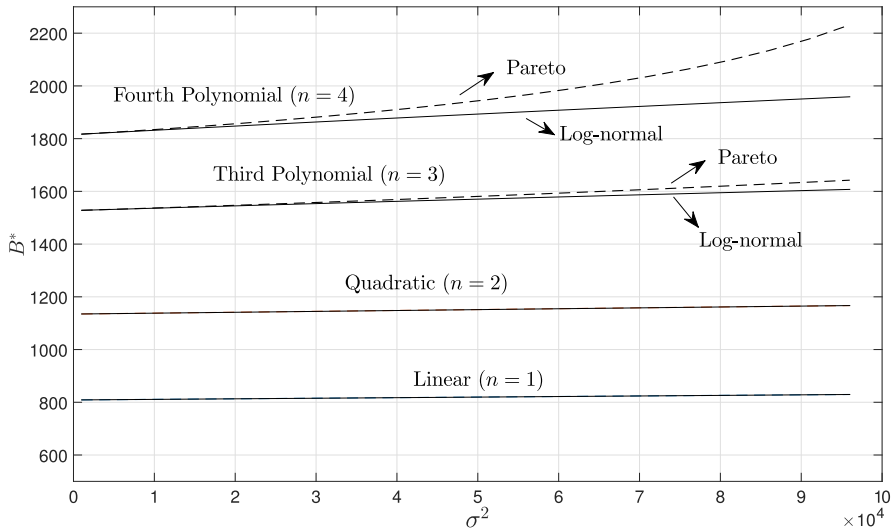


FIGURE **6.** Lognormal versus Pareto.

levels up to n. The asymptotic behaviors of the optimal investments in heavy and light trading environments are then characterized based on the moments of the distribution of the number of transactions. We find that when the market is highly liquid, that is, has a large expected transaction volume, it is optimal for a risk sensitive firm to use a balanced investment strategy. Moreover, we analyze the tail effects of various different types of distributions for transaction volumes on the optimal investments. We find that both a heavier right tail of the distribution of the number of transactions and a greater vulnerability will increase the optimal investments, and furthermore, that the two effects reinforce one another in a heavy trading environment. Finally, compared to a heavy trading environment, the behaviors of the optimal investments in a light trading environment are more diverse. With a heavier left tail, the investment will not depend on the risk sensitivity. So firms will invest according to the same ratio in a light trading environment.

Several research directions may be of interest in the future. First, we only considered allocation rules with regard to two general types of assets. One can actually consider more detailed categories of assets within a financial services firm. This problem may lead then to more complicated interaction processes. Second, one can consider a more detailed flow of transactions, and come up with a network model. Finally, models motivated by other categories of operational loss types may also be of interest. For instance, one can consider the moral hazard problem of rogue traders or the principal agent problem between managers and traders.

Acknowledgments

The authors are grateful to the two anonymous referees and the associate editor for their constructive reviews. The authors are also grateful for the many comments made by Marcelo Cruz (e-trade), Remco Bloemkolk (ING), and Pavel Shevchenko (CSIRO, Australia). Their comments have helped improve the manuscript considerably.

References

1. Basel Committee on Banking Supervision (2006). The first pillar-minimum capital requirements. BIS consultative document.

https://doi.org/10.1017/S0269964817000250 Published online by Cambridge University Press

2. Basel Committee on Banking Supervision (2011). Principles for the sound management of operational risk. BIS consultative document.
3. Bish, E.K. & Wang, Q. (2004). Optimal investment strategies for flexible resources, considering pricing and correlated demands. *Operations Research* 52(6): 954–964.
4. Bish, E.K., Zeng, X., Liu, J., & Bish, D.R. (2012). Comparative statics analysis of multiproduct newsvendor networks under responsive pricing. *Operations Research* 60(5): 1111–1124.
5. Black, F. & Litterman, R. (1992). Global portfolio optimization. *Financial Analysts Journal* 48(5): 28–43.
6. Bocker, K. & Kluppelberg, C. (2005). Operational VaR: a closed-form approximation. *Risk Magazine* 18(12): 90–93.
7. Boudreau, J.W. (2004). Organizational behavior, strategy, performance, and design in management science. *Management Science* 50(11): 1463–1476.
8. Boudreau, J.W., Hopp, W., McClain, J.O., & Thomas, L.J. (2003). On the interface between operations and human resources management. *Manufacturing and Services Operations Management* 5(3): 179–202.
9. Cassenti, D.N., Kelley, T.D., & Carlson, R.A. (2010). Modeling the workload-performance relationship. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 54(19): 1684–1688.
10. Chapelle, A., Crama, Y., Hübner, G., & Peters, J.P. (2008). Practical methods for measuring and managing operational risk in the financial sector: a clinical study. *Journal of Banking and Finance* 32(6): 1049–1061.
11. Cheng, F., Gamarnik, D., Jengte, N., Min, W., & Ramachandran, B. (2007). Modeling operational risks in business processes. *Journal of Operational Risk* 2(2): 73–98.
12. Chernobai, A.S., Rachev, S.T., & Fabozzi, F.J. (2007). *Operational risk: A guide to Basel II capital requirements, models, and analysis*. New York: John Wiley and Sons.
13. Chod, J. & Rudi, N. (2005). Resource flexibility with responsive pricing. *Operations Research* 53(3): 532–548.
14. Cooper, W.L. (2002). Asymptotic behavior of an allocation policy for revenue management. *Operations Research* 50(4): 720–727.
15. Cruz, M.G. (2002). *Modeling, measuring and hedging operational risk*. New York: John Wiley and Sons.
16. Deloitte (2013). Directors, CFOs Share Views on Risk, but Diverge on Time Spent. http://deloitte.wsj.com/riskandcompliance/2013/04/16/.
17. Gregoriou, G.N. (2009). *Operational risk toward basel III: Best practices and issues in modeling, management, and regulation*. New York: John Wiley and Sons.
18. Hora, M. & Klassen, R.D. (2013). Learning from others' misfortune: factors influencing knowledge acquisition to reduce operational risk. *Journal of Operations Management* 31(1): 52–61.
19. Iyer, R.K. & Rossetti, D.J. (1986). A measurement-based model for workload dependence of cpu errors. *IEEE Transactions on Computers* 100(6): 511–519.
20. Jarrow, R.A. (2008). Operational risk. *Journal of Banking and Finance* 32(5): 870–879.
21. Jarrow, R.A., Oxman, J., & Yildirim, Y. (2010). The cost of operational risk loss insurance. *Review of Derivatives Research* 13(3): 273–295.
22. Karger, D.R., Oh, S., & Shah, D. (2014). Budget-optimal task allocation for reliable crowdsourcing systems. *Operations Research* 62(1): 1–24.
23. Karsten, F., Slikker, M., & Van Houtum, G.J. (2015). Resource pooling and cost allocation among independent service providers. *Operations Research* 62(2): 476–488.
24. Kc, D.S. & Terwiesch, C. (2009). Impact of workload on service time and patient safety: An econometric analysis of hospital operations. *Management Science* 55(9): 1486–1498.
25. Kuntz, L., Mennicken, R., & Scholtes, S. (2014). Stress on the ward: Evidence of safety tipping points in hospitals. *Management Science* 64(1): 754–771.
26. Leippold, M. & Vanini, P. (2005). The quantification of operational risk. *Journal of Risk* 8(1): 59–85.
27. Luss, H. (2012). *Equitable resource allocation: Models, algorithms, and applications*. New York: John Wiley and Sons.
28. Markowitz, H. (1952). Portfolio selection. *The Journal of Finance* 7(1): 77–91.
29. Meier, H., Christofides, N., & Salkin, G. (2001). Capital budgeting under uncertainty-an integrated approach using contingent claims analysis and integer programming. *Operations Research* 49(2): 196–206.
30. Nash, P. & Gittins, J.C. (1977). A hamiltonian approach to optimal stochastic resource allocation. *Advances in Applied Probability* 9(1): 55–68.
31. Neil, M., Fenton, N., & Tailor, M. (2005). Using Bayesian networks to model expected and unexpected operational losses. *Risk Analysis* 25(4): 963–972.

https://doi.org/10.1017/S0269964817000250 Published online by Cambridge University Press

32. Powell, A., Savin, S., & Savva, N. (2012). Physician workload and hospital reimbursement: Overworked physicians generate less revenue per patient. *Manufacturing and Services Operations Management* 14(4): 512–528.
33. Ross, S.M. & Wu, D.T. (2013). A generalized coupon collecting model as a parsimonious optimal stochastic assignment model. *Annals of Operations Research* 208(1): 133–146.
34. Royset, J.O. & Szechtman, R. (2013). Optimal budget allocation for sample average approximation. *Operations Research* 61(3): 762–776.
35. Scharfman, J.A. (2008). *Hedge fund operational due diligence: Understanding the risks*. New York: John Wiley and Sons.
36. Sparrow, A. (2000). A theoretical framework for operational risk management and opportunity realisation. No. 00/10 New Zealand Treasury.
37. Tan, T.F. & Netessine, S. (2014). When does the devil make work? An empirical study of the impact of workload on worker productivity. *Management Science* 60(6): 1574–1593.
38. Thomas, P., Teneketzis, D., & Mackie-Mason, J.K. (2002). A market-based approach to optimal resource allocation in integrated-services connection-oriented networks. *Operations Research* 50(4): 603–616.
39. Topkis, D.M. (2011). *Supermodularity and complementarity*. Princeton, NJ: Princeton University Press.
40. Van Mieghem, J.A. (2003). Capacity management, investment, and hedging: Review and recent developments. *Manufacturing and Services Operations Management* 5(4): 269–302.
41. Van Mieghem, J.A. & Dada, M. (1999). Price versus production postponement: capacity and competition. *Management Science* 45(12): 1639–1649.
42. Wu, D.T. & Ross, S.M. (2015). A stochastic assignment problem. *Naval Research Logistics (NRL)* 62(1): 23–31.
43. Xu, Y., Pinedo, M., & Xue, M. (2017). Operational risk in financial services: A review and new research opportunities. *Production and Operations Management* 26(3): 426–445.

APPENDIX

Appendix A: Proof

Proof of Corollary 3.5: Let $C(N, B)$ denote the objective function of B and N in (3.3). Following a similar logic as in Proposition 3.3, we first show submodularity. Taking the second order partial derivative of $C(N, B)$ with respect to N and B, we obtain:

$$\frac{\partial^2 C(N,B)}{\partial B \partial N} = -\frac{\kappa_{\mathrm{e},1}}{\delta B^2} - \frac{4\kappa_{\mathrm{e},2} N}{\delta B^3} - \frac{\kappa_{\mathrm{h},1}}{(1-\delta)B^2} - \frac{4\kappa_{\mathrm{h},2} N}{(1-\delta)B^3} - \frac{4\gamma N}{\delta(1-\delta)B^3} < 0.$$

This proves that B^* increases in N. Next, we show sublinearity. Define $R = B/N$ as a new decision variable. Rewriting the total cost function as $C(N, R)$, we have:

$$C(N,R) = \frac{\kappa_{\mathrm{e},1}}{\delta R} + \frac{\kappa_{\mathrm{e},2}}{\delta^2 R^2} + \frac{\kappa_{\mathrm{h},1}}{(1-\delta)R} + \frac{\kappa_{\mathrm{h},2}}{(1-\delta)^2 R^2} + \frac{\gamma}{\delta(1-\delta)R^2} + RN.$$

If we can prove $C(N, R)$ to be supermodular in N and R, then R^* decreases in N and B^* increases sublinearly in N. The reason is that if $C(N, R)$ is supermodular in (N, R), then it is submodular in $(-N, R)$. Therefore, $R^*(-N) = arg\min C(-N, R)$ is increasing in $(-N)$ and decreasing in N. Taking the second order partial derivative of $C(N, R)$ with respect to N and R, we obtain:

$$\frac{\partial^2 C(N,R)}{\partial R \partial N} = 1 > 0.$$

This proves that B^* increases sublinearly in N. ■

https://doi.org/10.1017/S0269964817000250 Published online by Cambridge University Press

Appendix B:

TABLE **B.1.** Tail effects and vulnerability interaction simulation (Figure 5). (The column "Num" corresponds to the points in the figure)

Num	Diff 1	Diff 2	Diff 3	Num	Diff 1	Diff 2	Diff 3
2	4.374	4.661	11.966	52	4.426	4.710	12.055
4	4.374	4.663	11.970	54	4.429	4.715	12.061
6	4.375	4.663	11.969	56	4.434	4.720	12.069
8	4.376	4.666	11.975	58	4.440	4.724	12.077
10	4.377	4.667	11.976	60	4.447	4.730	12.085
12	4.377	4.664	11.972	62	4.455	4.736	12.095
14	4.378	4.669	11.979	64	4.458	4.742	12.105
16	4.380	4.670	11.981	66	4.467	4.749	12.117
18	4.381	4.671	11.984	68	4.476	4.756	12.131
20	4.384	4.673	11.986	70	4.485	4.764	12.147
22	4.385	4.674	11.989	72	4.496	4.775	12.166
24	4.387	4.676	11.992	74	4.509	4.787	12.182
26	4.389	4.677	11.995	76	4.520	4.800	12.205
28	4.391	4.679	11.998	78	4.538	4.817	12.232
30	4.394	4.681	12.001	80	4.556	4.830	12.270
32	4.395	4.683	12.005	82	4.587	4.854	12.303
34	4.397	4.684	12.007	84	4.613	4.879	12.345
36	4.398	4.687	12.012	86	4.647	4.913	12.399
38	4.401	4.689	12.016	88	4.692	4.956	12.465
40	4.404	4.691	12.021	90	4.751	5.010	12.553
42	4.408	4.694	12.026	92	4.833	5.092	12.672
44	4.411	4.697	12.031	94	4.959	5.202	12.842
46	4.415	4.700	12.036	96	5.148	5.390	13.115
48	4.419	4.703	12.042	98	5.527	5.725	13.560
50	4.421	4.707	12.048	100	6.364	6.550	14.554

Appendix C: Numerical Experiment

We present here another numerical experiment to show the relationship between the vulnerability and the tail effects. If we fix n and keep the means and variances of the distributions the same, then a heavier tail requires a larger investment. Furthermore, we show in the following experiments that a higher risk sensitivity, that is, a larger n, exacerbates the tail effects, which is also a conclusion from the theoretical results.

In the following numerical example, we use the variance σ^2 to capture the level of market uncertainty. When the mean μ is fixed, a larger σ^2 implies that the number of transactions is more variable around its mean. We would like to study the impact of the heaviness of the tail with different levels of uncertainty. We compare, therefore, in Figure 6 the behavior of the Pareto distribution with that of the Lognormal distribution. We choose a Pareto distribution with mean $\mu = 1000$ and a variance σ^2 in a range from 1000 to 100,000. A first observation is that the investment increases with the variance. A second observation is that when the degree n of the polynomial is either 1 or 2, the two distributions exhibit exactly the same behavior (since they have the same mean and variance). However, when the degree of the polynomial function increases (the n increases), we can see a clear spread between the investment curves of the two distributions. Assuming the same mean and variance, the Pareto distribution has a heavier tail and requires a larger investment. The spread increases when the degree of the polynomial function increases, which shows an increasing impact of the tail effects. So we conclude that the vulnerability of the firm and the tail effects reinforce one another with regard to their impact on the optimal investment level in the asymptotic regime.

https://doi.org/10.1017/S0269964817000250 Published online by Cambridge University Press