The Impact of Workload on Operational Risk: Evidence from a Commercial Bank

Yuqian Xu
Gies College of Business, University of Illinois at Urbana-Champaign
Tom Fangyun Tan
Cox School of Business, Southern Methodist University
Serguei Netessine
The Wharton School, University of Pennsylvania

Abstract

Operational risk is now among the three most significant types of risks in the financial services industry, and its management is mandated by Basel II regulation. This paper studies how banks' operational risk event frequency (or error rate) is affected by the workload in order to inform better labor decisions. To achieve this goal, we use a unique data set from a commercial bank in China that contains 1,441 operational risk events over 16 months. We find that workload has a U-shaped impact on operational risk frequency. More specifically, the error rate of operational risk events would decrease first as workload increases and then increase. Based on the causal relationships between workload and operational risk events and profit, respectively, we discuss the bank capital allocation impact of changing the staffing level among branches to reduce operational risk losses. We find that employing a flexible staffing rule can significantly reduce the number of operational risk events by more than 3% under different scenarios. In addition, this significant performance improvement can be achieved by adding even a little bit of flexibility to the process by allowing employees to either switch their business lines in the same branch or switch branches within the same business lines on a quarterly basis.

Key Words: operational risk, workload, frequency, severity, U-shape, optimal staffing.

1 Introduction

Operational risk (OpRisk) in financial services is defined as *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, 2006). For example, it includes execution and process management-related risk events, such as data entry errors, accounting errors, failed mandatory reporting, and negligent loss of client assets. OpRisk is one of the three major risks (together with credit risk and market risk) that banks face and have a dire need to minimize, especially after the 2007-2009 global financial crisis, which cost the banking sector trillions of dollars due to poor risk management (Ashby, 2010). For example, according to Barclays 2014 Annual Report[1], OpRisk accounts for 9% of its total risk exposure (around 3,285 million USD), tied with market and liquidity risk (9%), and second only to credit risk (72.4%), while the remainder is due

[1] http://www.home.barclays/annual-report-2014.html.

to various other risks (e.g., funding, conduct risks). Because of its significance, regulators throughout the world generally require their banks to set aside capital in reserve in order to protect themselves from OpRisk events. The Basel Committee on Banking Supervision (BCBS) in Basel, Switzerland, whose members include 27 major economies (e.g., the U.S., the U.K., Japan and the BRIC countries), makes recommendations and sets guidelines with regard to minimum capital requirements for risk management (Marrison, 2005). After making recommendations for the credit and the market risks in 1988 (i.e., Basel I), the Basel Committee issued a new set of regulations and guidelines with regard to capital reserves for the OpRisk in 2004, which is referred to as Basel II. Recent regulations (Sarbanes-Oxley Act, 2002 and Basel, 2012) have made operational risk management even more important in the financial industry. Hence, mitigating the OpRisk to reduce the necessary capital reserves is a practically significant goal that global banks earnestly strive for.

Furthermore, unlike credit and market risks, OpRisk is often perceived by management as more controllable (Deloitte, 2013) because proper monitoring processes can prevent such risk events from happening (Kaplan and Mikes, 2012). Yet in practice, the compliance system alone does not guarantee that OpRisk events can be averted. Companies are often advised to promote a corporate culture that enables prudentdecision-making processes that reduce the opportunity for OpRisk (Healy and Serafeim, 2019). Therefore, there are many opportunities for operations management researchers to improve our understanding about the causes of the OpRisk and make better operational decisions to mitigate such risks in the future. Nevertheless, despite such opportunities, little empirical Operations Management research has been done on it,probably because the OpRisk data is rarely available for academic research.

In this paper, we study how one specific risk–workload (defined as the total number of transactions handled per employee), an important work environment factor (see Bendoly et al., 2006, and more details will be provided in Section 2) affects the operational risk events frequency (or error rate) and severity (potential loss scale). In particular, we use a unique longitudinal operational risk data set from a Chinese commercial bank that contains 1,441 operational risk events from January 1, 2014 to April 30, 2015 to study how workload affects operational risk. Adopting an instrumental variable (IV) approach to account for potential endogeneity issues, we find that workload has a U-shaped relationship with operational risk error rate. More specifically, the error rate of operational risk events would decrease first and then increase as workload increases. Using the detailed categories of our risk events data, we further find that, under a low workload scenario, employees tend to make performance-seeking risks, while under a high workload scenario, employees tend to make errors or have quality degradation due to cognitive multitasking.

Based on the causal relationships between workload and operational risk events and profit, respectively, we discuss the impact of bank capital allocation by changing staffing levels among branches to reduce operational risk losses. Keeping the risk severity level the same, we find that employing a flexible staffing rule can significantly reduce operational risk losses by more than 3% under different scenarios. In addition, a significant performance improvement can be achieved by adding even a little bit of flexibility to the process by allowing employees to either switch their business lines in the same branch or switch branches within thesame business lines on a quarterly basis.

The contributions of this paper are three-fold. First, to the best of our knowledge, our paper provides the first empirical analysis of one of the causes of operational risks in the banking industry, while the previous studies tended to model OpRisk as an event with exogenous probability distribution. In particular, we establish a causal link between workload, an important work environment factor that is determined by management, and the error rate of operational risk events. Second, we revisit the growing area in operations management on the impact of workload on operational performance, and we broaden our understanding about workload and staffing decisions in the financial industry. This area is understudied in the empirical OM literature. Third, our empirical study enables us to explain the variation in OpRisk events, so we can build a capital allocation model to re-optimize bank staffing levels among branches to improve OpRisk management. Our paper therefore provides a novel consideration for the labor optimization literature: staffing to avert the OpRisk events.

2 Related Literature

Our research contributes to three streams of literature: i) retail banking operations, ii) labor management and iii) OpRisk in financial services, respectively.

The first stream of literature relevant to our work concerns the retail banking operations. Prior work in this stream has looked into consumer channel adoption decisions. Hitt and Frei (2002) explore the difference between electronic and physical distribution channels by studying the case of personal-computer-based (PC) banking. Campbell and Frei (2004) use a unique data set from a financial services firm to study the persistence in customer profitability. Xue et al. (2007) study the effect of customers' banking channel usage on retail banking performance. Campbell and Frei (2010) research the impact of consumers' adoption of online banking channels on their interactions with a major U.S. retail bank. More recent work in retail banking operations has studied service quality issues. Buell et al. (2016) probe customers' reaction to the increasing service quality competition, and find that for firms to maintain high quality levels they need to attract and retain profitable consumers over time. In general, existing literature on retail banking operations has focused on channel decisions and customer management; however, no research has been done to study OpRisk and bank operations.

The second stream of relevant work studies optimal labor decisions in service industries. One large stream is labor management in retail (please see Ton and Huckman, 2008; Kesavan and Mani, 2015, for a comprehensive literature review). Staffing is a key managerial decision in this setting because it affects operational performance. For example, Perdikaki et al. (2012) find that increasing staffing level by one standard deviation from the sample mean improves the marginal returns to traffic from \$10.00 to \$11.32 per person. In addition, Chuang et al. (2016) suggest that sufficient staffing critically enables retailers to fully seize the sales opportunities of increasing store traffic. If understaffed, however, stores are estimated to suffer from lost sales by 8.56%, and lower profitability by 7.02%, a result found in an apparel retailer setting (Mani et al., 2015). Beyond estimating the counter-factual effect of staffing decisions, Fisher et al. (2017) implement a new staffing rule and validate the counter-factual estimation in practice. They find that their staffing rule increases revenues by 4.5%, and annual profits by \$8.9 million, adjusting for the additional labor costs. In this study, we additionally show the importance of staffing decisions by analyzing a new pathway of how staffing affects operational performance through the workload.

Studying workload and operational risks adds to an increasing number of Operations Management papers that have answered the call for researching how external factors, such as workload, affect workers' performance (Boudreau et al., 2003). These papers have mostly examined the impact of workload on operational performance in a healthcare setting, but none of these have studied the financial sector, one of the most significant sectors of the economy[2]. For instance, KC and Terwiesch (2009) conduct an empirical analysis of the impact of workload on service time using operational data from patient transport services in cardiothoracic surgery. In a later paper, KC and Terwiesch (2012) find that the occupancy level of a cardiac intensive care unit is negatively correlated with patients' length of stay. In addition, Powell et al. (2012) find that overworked physicians generate less revenue per patient because of a workload-induced reduction in due diligence with regard to paperwork. Kuntz et al. (2014) discover a nonlinear relationship between hospital workload and mortality rates. Berry and Tucker (2016) analyze two years of inpatient data from 203 hospitals in California and find an N-shaped relationship between occupancy and length of stay. Aral et al. (2012) find that multitasking (a proxy for workload) level exhibits diminishing returns to project output in a mid-size executive recruiting firm. Moving away from the healthcare industry, Tan and Netessine (2014) analyze a large, detailed operational data set from a restaurant chain and show an inverted-U-shaped relationship between the workload and the waiter's service speed and sales output. In this paper, we broaden the studies on workload to the financial services industry, an economically significant industry.

Finally, compared to the extensive literature on market risk and credit risk (e.g., French et al., 1987; Jarrow and Turnbull, 1995; Altman and Saunders, 1997; Dowd, 2007), OpRisk in financial services has generally received little attention in the past. Nevertheless, given its practical significance and new regulations, OpRisk has been receiving growing interest in the academic literature in recent years. We contribute to this burgeoning literature in three ways. First, much of the academic research has tended to discuss the modeling, the measurements, and the regulations of OpRisk at the strategic level rather than at the operational level, which we study in this paper. For example, Cruz (2002) discusses the background and the definition of operational risk, explains measurement methods, and discusses operational risk management strategies. Chernobai et al. (2007) provide the framework and guidelines regarding operational risk background and measurement models based on Basel regulations. Scharfman (2008) examines the operational risk management framework and measurements with a focus on hedge fund operational risk. None of these studies, however, discuss the operational-level decisions that cause operational risks, which is the focus of this paper.

Second, most of the recent papers tend to focus on the statistical modeling of aggregate operational risk loss distributions to estimate the Operational Value at Risk (Ops-VaR) rather than explaining what factors may explain the variation of OpRisk events. For example, Neil et al. (2005) propose a Bayesian Network approach to model both expected and unexpected OpRisk losses. Bocker and Kluppelberg (2005) find a closed-form approximation for Ops-VaR exists when the distribution of OpRisk loss data are heavy-tailed. Unlike these papers, we examine how an important work environmental factor (i.e., workload) affects the frequency and severity of OpRisk events, so that managers can make better decisions to distribute an optimal workload.

[2]https://www.selectusa.gov/financial-services-industry-united-states

Third, recent papers on OpRisk have started to focus on the financial applications, but only a handful of papers have considered the actual operations management decisions. For example, Leippold and Vanini (2003) propose a theoretical model together with numerical experiments to quantify risk losses for banks through their value chain. Jarrow (2008) suggests a modeling framework for firm asset pricing with operational risk losses. Within this asset-pricing framework, Jarrow et al. (2010) further studied the OpRisk insurance contract. It is noteworthy that these papers tended to neglect the granular-level operations decisions (e.g., workforce management), which should significantly affect the OpRisk events in an applied context. Unlike these papers, we examine how an important endogenously determined by management work environmental factor (i.e., workload) affects the frequency and severity of OpRisk events, and we point out that managers can make better decisions to distribute workload optimally. The limited OM literature has recently started to study the causes of OpRisk in non-finance settings. For example, Shah et al. (2017) study the causes of product recalls in the automotive industry. Hora and Klassen (2013) conduct a field experiment to analyze factors that affect firm manager's knowledge acquisition from operational risk losses of other firms. One goal of our paper is to take an initial step in filling the gap between operational risk in financial services and operations management by showing the implications of a fundamental OM decision (i.e., staffing decision) for mitigating OpRisk in financial services.

3 Context and Hypothesis Development

3.1 Retail Banking Branch

Before developing our theoretical hypothesis, we explain the specific context of our bank branches. We study the lowest-level branch in the Chinese commercial bank system, comparable to a credit union branch in the United States. This level is technically called a "savings branch" in Chinese. It handles only basic personal and business banking, which mainly includes deposit and savings services, processing loans, and selling wealth management products, such as mutual funds and bonds. It generally employs two types of workers: tellers and managers.

The tellers are the front line employees. They sit behind a glass window and directly serve the customers. They are the largest in quantity, and the busiest. During the shift, they press a button to call the next customer. They first verify the customer's identity, and then process the service request. Depositing and withdrawing money is the most common personal banking service in China because of the late adoption of the ATM. Other services include issuing cashier's checks, transferring funds, selling wealth management products, and processing personal loans. At the end of their shift, the tellers need to balance the accounts and secure all the stamps and blank checks. The second position is the branch manager. He/she is responsible for the branch performance.

All the branch employees have three parts of income—base salary, performance-based bonus, and commissions. The base salary and the bonus are the largest parts of the income, and are typically the same across the same level of employees. A bonus will be reduced if an employee makes mistakes, such as the OpRisk violations. The commission is related to the number and the amount of transactions and the wealth management products that the employee processes. The store-level transaction amounts and wealth management product sales also contribute to the individual commission.

3.2 OpRisk Auditing Process

The internal audit function is required by the Basel Committee on Banking Supervision (June 2012) and it works according to the following important principles—"*Principle 2: The bank's internal audit function must be independent of the audited activities, which requires the internal audit function to have sufficient standing and authority within the bank, thereby enabling internal auditors to carry out their assignments with objectivity,*" "*Principle 4: Internal auditors must act with integrity*," and "*Principle 7: The scope of the internal audit function's activities should ensure adequate coverage of matters of regulatory interest within the audit plan.*" Therefore, the auditors should be able to confidently say that they dutifully and precisely capture almost all such errors. However, they may capture those errors at different time points; some errors might be caught before losses are realized, while some might be caught after losses are realized. Auditors face significant repercussions if a loss happens that was not detected in time. In addition, all the auditors must be well trained to conduct the auditing process—"*Principle 3: Professional competence, including the knowledge and experience of each internal auditor and of internal auditors collectively, is essential to the effectiveness of the bank's internal audit function.*"

3.3 Hypothesis Development

Building on the extensive literature about the effect of workload on performance, we propose four main mechanisms through which workload could affect OpRisk occurrence. The first two mechanisms will suggest a negative relationship between workload and the frequency of OpRisk events, while the second two will predict a positive relationship. We will argue that the negative effect dominates under low workload, and that the positive effect dominates under high workload. This argument reconciles these seemingly conflicting mechanisms, and leads to our main hypothesis.

Workload Reduces OpRisk (Negative Effect) The first negative effect mechanism is motivation, which can be strengthened by workload to increase workers' human capacity and thence performance (Deci et al., 1989). Indeed, additional workload can increase arousal regarding the work, which helps workers stay "in the zone" (Bendoly and Prietula, 2008; Bendoly, 2011). Increased workload may also be perceived as exciting and setting challenging goals. Such goals may improve workers' motivation according to goal-setting theory (Locke, 1968; Latham and Locke, 1979). In addition, research in cognitive psychology has found that extra workload triggers the cortex to release hormones that enhance cognitive performance (Lupien et al., 2007). On the other hand, a very light workload may trigger workers to fill the idle time with irrelevant and counter-productive activities, aka "Parkinson's Law"(Parkinson, 1958).

When workload is low at banks (e.g., branch traffic is low), bank employees are more likely to be inclined to engage in counter-productive activities, such as chit-chatting with coworkers, checking their phones, playing with games, and attending to their own personal affairs, all of which can distract workers' attention and make OpRisk-related mistakes. For example, absentminded workers may fail to carefully verify all the required documents/information, such as credit history, for a loan service. On the opposite side, increasing workload can reduce such idle time and stimulate the workers to expend more effort to ensure their work quality and accurately follow the OpRisk protocol.

The second negative effect mechanism is economic multitasking (e.g., Holmstrom and Milgrom, 1991), which suggests that employees may rationalize their effort provision towards different tasks to maximize their rewards. This theory implies that varying workload may change the risk and return of the tasks, thus prompting workers to reallocate their attention priority. For example, Tan and Netessine (2014) find that restaurant waiters have a strong incentive to increase attention to generating sales from each table at the expense of slower service speed under a light workload (measured in terms of the number of tables that a waiter simultaneously handles) because waiters want to maximize their earnings (i.e., tips, which are directly related to sales) from the limited number of tables that they are assigned to, and because the waiting cost of customers is low. Under heavy workload, however, waiters have a different incentive to shift the focus onto faster service speed because the waiting cost becomes high, and because turning tables faster will seat new customers sooner, and these customers generally spend more money per unit of time than incumbent customers.

Bank employees are faced with two possible income-generating "tasks"—regular day-to-day duties and alternative reward-enhancing activities. The former task typically includes accurately and efficiently processing various customer service requests, which are tied to bonus, commission and promotion. The latter is exemplified by various malicious activities, such as coercing customers into making a deposit before releasing a loan payment in order to boost deposit performance, aggressively selling financial products, or even committing fraud. Employees engage in these activities because they may be pressured by aggressive performance targets linked to bonuses and promotions. All of these activities may boost income in the short run, but they violate the OpRisk protocol. When the workload of regular business increases, the bank employees may start to curb these alternative reward-seeking activities and more diligently follow the protocol because 1) they are more likely to reach their performance goals of regular duties to earn the rewards (i.e., less pressured by the aggressive performance target), 2) they have little latitude to undertake such reward-seeking activities, and 3) violation of OpRisk protocol may be caught and penalized[3].

We argue that the negative effect of workload dominates when the overall workload is low. First, idle time is more likely to happen under low workload, triggering the motivation mechanism. Second, workers have extra incentives to neglect OpRisk protocol to pursue alternative award-enhancing activities when regular business workload is low (i.e., economic multitasking). Hence, we expect to observe that increasing workload reduces the number of OpRisk events, especially those award-enhancing ones, when the overall workload is low.

However, when the overall workload is high, a further increase in workload may increase OpRisk events as we discuss below.

Workload Increases OpRisk (Positive Effect) The first positive effect mechanism is cognitive multitasking, which suggests that workers will become less capable of focusing on an individual task when they have

[3]We assume that the employees will prioritize regular legal duties because illegal activities bear two additional costs - a chance of being caught and punished and moral/ethical costs.

to pay attention to an increasing number of tasks because of limited cognitive capacity (Charron and Koechlin, 2010). In other words, when the cognitive load is high, any additional task will consume a portion of cognitive bandwidth at the cost of other tasks (Schmidt and DeShon, 2007), causing more errors and service quality degradation. A significant amount of empirical literature supports this theory. For example, Powell et al. (2012) find that overloaded physicians become careless about insurance paperwork, which reduces revenue per patient. KC (2013) discovers that when doctors become extremely busy in the emergency room, they take longer to discharge patients while providing lower quality care. In addition, in a Japanese bank's home loan application-processing line, which is relevant to the setting of this study, Staats and Gino (2012) report that having the workers specialize in one task improves their productivity in a single day because alternating focus among multiple tasks may distract workers' attention in the short term.

Bank employees in our empirical setting also perform multiple cognitive tasks. For example, bankers usually handle multiple loan requests simultaneously because processing loans takes time while awaiting approvals. When workload increases, they are more and more likely to lose focus on any particular task, causing them to make errors and increase OpRisk of various severity, which may range from forgetting to make photo copies of required documents to failing to verify the validity of clients' information, to losing important documents/seals or even forgetting to lock the safe.

The second mechanism concerns various workload-induced anti-productive emotions. Excessively high workload may exhaust workers and reduce their physical and cognitive capacities, making them prone to errors (Cakir et al., 1980; Setyawati, 1995). In addition, heavy workload can stress and frustrate workers, who may consequently cut corners and produce low-quality work (Peters and O'Connor, 1980; Oliva and Sterman, 2001; Bendoly, 2011). Moreover, extra workload can confuse and intimidate workers because various tasks may create conflicting goals and exacerbate the difficulty of accomplishing these tasks, which can lead to a lack of commitment and motivation to fulfill them (Donahue et al., 1993; Dalton and Spiller, 2012). Empirically, Kuntz et al. (2014) examine the discharge records and discover that high hospital occupancy increases front-line clinical workers' stress hormones and forces them to ration resources and become more error-prone, thus increasing patient's mortality rate.

As front-line workers in banks, tellers have to perform multiple duties, such as depositing, transfer, withdrawals, and issuing cashier's checks and money order. In addition, they need to promote the bank's products, resolve various customer issues, batch and process proof of work, while following all the OpRisk standards. When their workload expands, these tellers may encounter all the aforementioned anti-productive emotions and consequently violate the OpRisk protocols. For example, a tired teller may type the wrong deposit amount into the system or mishandle counterfeit money. In addition, a frustrated teller may become impatient with clients and even violently quarrel with them, damaging the bank's reputation and future business. A confused or intimidated worker may even commit financial fraud.

We posit that when the overall workload is high, further increasing workload may increase OpRisk, especially those execution-related or corner-cutting errors. Under a heavy workload, employees are more likely to reach or exceed their cognitive and physical capacities, which triggers both cognitive multitasking mechanism and anti-productive emotions to take effect.

To sum, we formally hypothesize that:

There is a down-up relationship between workload and OpRisk occurrence. That is, as workload increases, the frequency of OpRisk events will first decrease and then increase, controlling for everything else.

4 Data

Our empirical setting is based on 49 branches that belong to one major Chinese retail bank in Jiangsu Province. Jiangsu is one of the largest provinces in China, with the second-highest GDP nationwide[4]. In 2013, the bank implemented a new system to record operational risk events, from which we collected our data. Our data covers the time period from January 1, 2014 to April 30, 2015, when 1,441 operational risk events in total were observed.

The data consist of four parts—operational risk events, transaction-related information, branch characteristics, and employee demographics. In particular, the operational risk event data includes bank branch id, event description (in text format), date of occurrence, and severity level. In addition, the transaction-related information contains transaction date, type, and size. There are five general types of transactions in our data set, namely, deposit/withdrawal for individual client, credit/loan for individual client, sales of wealth management products for individual client, deposit/withdrawal for business client (firm), and credit/loan for business client (firm). The branch characteristics capture the total number of employees, branch address, and its distance to the headquarters. The employee demographic information includes education, industry working experience, and working position (associated with the type of transaction) in the branch.

4.1 Operational Risk Events

According to the Basel II definition[5], there are seven categories of operational risk events, i.e.,
(1) Internal Fraud: misappropriation of assets, tax evasion, intentional mismarking of positions, bribery.
(2) External Fraud: theft of information, hacking damage, third-party theft and forgery.
(3) Employment Practices & Workplace Safety: discrimination, workers compensation, employee health and safety.
(4) Clients, Products, & Business Practices: market manipulation, antitrust, improper trade (aggressive sales), product defects, fiduciary breaches, account churning.
(5) Damage to Physical Assets: natural disasters, terrorism, vandalism.
(6) Business Disruption & System Failures: utility disruptions, software failures, hardware failures.
(7) Execution, Delivery & Process Management: data entry errors, accounting errors, failed mandatory reporting, negligent loss of client assets.

In general, our focal bank has labeled the 1,441 risk events into four categories according to the Basel II definition: (i) data-entry errors, (ii) failed mandatory reporting, (iii) operation failure, and (iv) aggressive sales. Here categories (i), (ii), and (iii) belong to Execution, Delivery & Process Management (ED&PM),

[4]http://www.guancha.cn/economy/2017_01_27_391590.shtml
[5]https://en.wikipedia.org/wiki/Operational_risk

which corresponds to losses arising from an unintentional or negligent failure to meet a professional obligation. On the other hand, category (iv) belong to Clients, Products, & Business Practices (CP&BP), which corresponds to intentional improper trade due to performance-seeking behaviors. In addition to the four labels provided by the bank, we also conduct text-mining analysis of the risk event, description in our Appendix B. We then further categorize our risk events into ten types through text mining, and we conduct robustness check with alternative specification of risk types in our fixed-effects model.

Some operational risk events in our dataset cause immediate losses to the bank. For example, "Branch X issued loan contract with an interest rate of 1.23 percent instead of 7.23 percent" (translated from Chinese). However, some events may cause losses only in the long run. One such risk event is recorded as follows, "On August 25, 2014, Branch X issued Company X RMB 2 million (Note: Chinese currency) loan, and wrote the wrong maturity date as September 1, 2014 (Error in the date)." Another such event is documented as follows, "On September 10, 2014, Branch X issued Company X business loans without checking the collateral." In sum, these OpRisk events are related to process conformance (Ton and Huckman, 2008), and will considerably cost the bank either in the short or in the long term. All of the events in our dataset were caught and recorded by the audit department of our focal bank, which checks the operational processes of each branch every week. Once a risk event is caught, the audit department assigned an operational risk severity score (following the internal risk severity score standard) to each event, which is determined by the potential losses of each event. Note that if these errors were caught by the audit department, the bank must to correct them, and hence it wouldn't negatively affect the performance. However, since our research question is how workload affects the risk events/errors' severity and frequency when these events happened, the correction approach after detection does not affect our results. On the other hand, it is possible that some errors are not caught/reported during our observation time period. If such measurement error is more likely to happen when the workload is low, our finding of a negative coefficient of workload is a conservative estimate. By contrast, if the underreporting is more likely to happen when the workload is high, our result may underestimate the effect of workload on error rates. Nevertheless, since the auditing department is operating completely independently from the retail locations, we believe that the measurement error of the operational risk events is equally likely to happen regardless of the workload level, so that our estimation on the workload effects should be unbiased.

4.2 Risk Measures

In this subsection, we define our dependent variables that are related to operational risk losses. We examine two performance measures: the error rate and the average risk severity level per event because they are the two most important performance measures for operational risk losses in practice (Cruz, 2002). In particular, error rate is computed as the total number of risk events divided by the total number of transactions handled by employee i at branch j in month t, namely,

$$ErrorRate_{ijt} = \frac{\text{total number of risk events}_{ijt}}{\text{total number of transactions}_{ijt}}. \tag{1}$$

We call this variable "error rate" because the operational risk events in our dataset are mainly human errors or mistakes. One advantage of using error rate as compared to frequency is that the error rate is scale free relative to the number of transactions. Later we conduct Poisson regression on frequency controlling for the number of transactions to further validate our results.

The second dependent variable is $Severity_{ijt}$, which is calculated as the average severity level of all the errors made by employee i at branch j in month t. The severity level is defined by the central bank of China to reflect the potential losses of each event. It is measured on a scale from 0 to 100, with the higher value indicating a more severe risk. Although it is almost impossible to objectively quantify all the losses, the severity level serves as an approximation of actual loss severity because it is measured based on the rules from both Basel regulation and the People's Bank of China (the regulator). To be more specific, the severity score is estimated following the Basel standard (see Basel Committee on Banking Supervision, 2011), which requires each bank to design its own "scorecard" system to record potential operational risk loss of each event with a severity score. The "scorecard" system adopted by our focal bank follows the standard from the People's Bank of China.

4.3 Independent Variables

The main independent variable that we study is the workload denoted by $Load_{ijt}$, which is the number of transactions handled by employee i at branch j during month t. We also use an alternative definition of workload to reflect the utilization of the employee capacity in our robustness check section. Following previous literature (see Tan and Netessine, 2014; Macchiavello and Morjaria, 2015) on linear and quadratic terms to study inverted-U or U-shaped relationships, we further take the square of $Load_{ijt}$ and call it $Load_{ijt}^2$ to test our hypothesis about the non-linear effect of workload. We standardize these variables by first subtracting their means and then dividing them by the standard deviations, so that the variables are between zero and one. Note that this mean-centering standardization of both workload variables can also reduce their correlation. To reduce the concern of multicollinearity, we show the correlations before and after the mean-centering. Before the mean-centering, the correlation coefficient between the linear and non-linear term is 0.156; and after the mean-centering, the correlation coefficient is 0.102. Therefore, we do not think our findings will suffer from multicollinearity.

There is not much guidance in the literature on which controls to use in a study like ours since we are the first to study empirically causes of operational risks in banks. Nevertheless, in addition to the workload measure, we propose the following three types of control variables. First, we account for the learning effects. Following Lapré and Nembhard (2011) and Clark et al. (2013), the cumulative workload is computed as the total workload from the beginning of our observation period till the date under study. Second, we control the risk event types by introducing the percentage measure of the four risk types ((i) data entry errors, (ii) failed mandatory reporting, (iii) operation failure, and (iv) aggressive sales). We later change these four labeled risk types into the ten categories derived from text mining and conduct our analysis again as a robustness check. The percentage measure is computed as the number of focal risk type events divided by the total number of risk events. Third, we introduce dummy variables of business types. To be more specific, dummy variables for deposit/withdrawal for individual client, credit/loan for individual client, sales of wealth management products for individual client, deposit/withdrawal for business client (firm), and credit/loan for business client (firm). To account for the effects of time-invariant variables, we also add employee, branch, and month-year fixed effects in our main estimation. In our robustness check, we also consider variables of employee education level and employee experience level. The employee education level is a categorical variable[6] from one to four. The employee experience level measures employees' years of related financial industry working experience.

Table 1 presents the summary statistics of our key variables based on 9,954 observations (N=9,954) of 675 employees in 49 branches at a monthly level. On average, each employee at each bank branch handles 1,678 transactions every month, which is equivalent to 56 transactions per day and 5.6 transactions per hour (assuming 10 working hours per day). The risk frequency per month is 1.692, which is consistent with the generally low frequency property of operational risk (Cruz, 2002). The risk severity level on average is 3.565 based on the standard "scorecard" system. Moreover, the variation of risk severity level is quite large, with the standard deviation being 11.2 and the maximum being 94. Notably, certain risk events can cause a significant amount of losses (94 severity score).

Table 1: Summary statistics (monthly).

Variable	Definition	Mean	Sd.	Min	Max
Tran.num	Number of transactions in each branch in 1,000	15.85	1.850	14.00	43.50
Num.empl	Number of employees in each branch	13.77	2.790	6.000	19.00
Rural/Urban	Dummy variable (one for rural area)	0.341	0.305	0.000	1.000
Load	Number of transactions conducted by each employee	1,678	1,145	777.8	71,600
Err.num	Number of risk events	1.692	4.780	0.000	26.00
Severity	Average event severity level	3.565	11.20	0.000	94.00
Edu_lvl	1: high school, 2: bachelor, 3: master, 4: doctor	2.300	1.100	1.000	4.000
Exp_lvl	Years of related financial industry working experience	3.500	3.700	0.000	36.00

5 Estimation and Results

In this section, we discuss our fixed-effects model and the estimation results. Section 5.1 presents the details of the fixed-effects model and the identification strategy; Section 5.2 describes our main empirical results; Section 5.3 discusses the impact of workload under different scenarios and studies subset analyses; Section 5.4 summarizes robustness checks of our main results with alternative workload measure and model specification.

5.1 Model and Identification

We consider the fixed-effects model of the monthly error rate as our main analysis. We here consider our estimation at a monthly level due to the low frequency of risk events (on average 1.6 events happened per month). In particular, we have the following specification for the error rate of employee i at branch j in

[6] 1: high school degree, 2: bachelor degree, 3: master degree, 4: doctorate degree

month t:

$$ErrorRate_{ijt} = \alpha_0 + \alpha_1 Load_{ijt} + \alpha_2 Load^2_{ijt} + \boldsymbol{\alpha} \boldsymbol{X}_{ijt} + \eta_i + \xi_j + \tau_{t_m} + u_{ijt}. \tag{2}$$

The vector $\boldsymbol{X}_{ijt}$ contains a list of covariates including *Cum_load*, the percentage measure of the four risk types, and dummy variables of business types. We also control for the employee, branch, and month-year fixed effects (η_i, ξ_j, and τ_{t_m}). In our analysis, we standardize all variables except binary. We clustered the standard errors at the employee and branch levels. To account for the simultaneity issue that the workload may be a function of the error rate, we adopt an instrumental variables (IV) approach to estimate the fixed-effects 2SLS model.

Although our fixed-effects models control for both observed and unobserved heterogeneity at the individual level, the models may still be prone to some endogeneity issues. For instance, one potential omitted variable could be the branch manager's effort level of engaging in risk management, which should be negatively correlated with the frequency and the severity of OpRisk. In addition, the risk management effort level should be negatively correlated with workload because the manager may be too busy handling transactions to manage OpRisk. Hence, we may potentially underestimate the true impact of workload on risk frequency and severity.

To alleviate the endogeneity issue mentioned above, we use the IV approach, which is widely used to address such endogeneity issues (Kennedy, 2003). The choice of a good IV should meet two conditions, namely relevance and exclusion (Wooldridge, 2010). The relevance condition requires the IV to be correlated with the endogenous variable, while the exclusion condition requires the IV to be uncorrelated with the error term. In essence, the IV should only be correlated with the dependent variable through the endogenous variable. In our estimation, we use two types of IVs. The first IV is the number of transactions online and the number of e-visits online. Our dependent variables are constructed by risk events that happened in the physical branches, therefore they should be independent of the online banking system. However, the load (transaction, e-visits) of online banking should be correlated with the offline workload, see Table 4. The second IV is the Hausman type (following Sheldon, 2016; Xu et al., 2016; Allon et al., 2018), which is the load of all other employees in the same month and the square of the load of all other employees in the same month. The dependent variable of error rate and this IV are positively correlated and the estimate of the instrument in the first stage is statistically significant, see Table 4. In addition, we believe that the average workload of all other branches should not directly affect the cumulative losses of the focal branch.

5.2 Estimation Results

We used a two-stage least square procedure to estimate our fixed-effects models with the four IVs and show our estimation results in Table 2. Column (1) of Table 2 shows the IV estimation results without the nonlinear term, and we find a negative effect of workload with coefficient -0.121 (at a significance level of 0.1%). Next, column (2) of Table 2 shows the IV estimation results with the nonlinear term. We find that the coefficient of $Load^2$ is positive (coefficient = 0.471 significant at 0.1% level). In addition, the linear term is significant and the coefficient is equal to -0.503, which implies that the critical point is approximately 0.534

$(-\alpha_1/2\alpha_2)$. Comparing the values of R^2 in the linear and nonlinear models, we find the nonlinear model has a better goodness-of-fit than the linear model (R^2 =0.562 v.s. 0.554), which further supports our hypothesis about the U-shaped relationship between workload and the error rate. In addition to the workload impact, we can see that employees with a higher cumulative workload are negatively associated with the error rate, which indicates the learning effects of employees. The higher cumulative workload an employee has, the more experience he/she has with handling transactions, and hence fewer errors.

We next consider an alternative measure of workload, which is operationalized as

$$Load_{2ijt} = Tran.num_{ijt}/Tran.cap_{ijt}, \tag{3}$$

where the transaction capacity $Tran.cap_{ijt}$ is defined as the 95% of the maximum monthly number of transactions (a similar measure is used in Jaeker and Tucker, 2016). We then define the quadratic term $Load_2^2$ and use it and $Load_2$ to replace $Load^2$ and $Load$ in our main model. The results are shown in columns (3) and (4) of Table 2. As can be seen, the coefficient of the linear term is significant and negative and the coefficient of the quadratic term is significant and positive, supporting the U-shaped relationship between workload and error rate. Interpreting the two coefficients, the critical point is equal to 0.535 $(-0.491/(2\times 0.459))$. In addition, the estimation results for all other variables remain consistent as our main estimation.

Table 2: Impact of Workload

	Main Model		Alternative Load	
	(1)	(2)	(3)	(4)
	Linear	Non-linear	Linear	Non-linear
Load	-0.121***	-0.503***	–	–
	(0.021)	(0.082)	–	–
$Load^2$	–	0.471***	–	–
	–	(0.078)	–	–
$Load_2$	–	–	-0.119***	-0.491***
	–	–	(0.020)	(0.077)
$Load_2^2$	–	–	–	0.459***
	–	–	–	(0.072)
Cum_load	-0.188***	-0.212***	-0.177***	-0.133***
	(0.030)	(0.041)	(0.023)	(0.021)
Risk_Type (%)	Yes	Yes	Yes	Yes
Business_Type_Dummy	Yes	Yes	Yes	Yes
Employee_FE	Yes	Yes	Yes	Yes
Branch_FE	Yes	Yes	Yes	Yes
Month_Year_FE	Yes	Yes	Yes	Yes
N	9,954	9,954	9,954	9,954
R^2	0.554	0.562	0.563	0.569
adj. R^2	0.551	0.559	0.560	0.568

Standard errors in parentheses

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

5.3 Workload Impact Discussion and Subset Analysis

Based on the discussion above, so far we find the major result that workload has a U-shaped impact on error rate. To further understand the estimation results with our proposed workload mechanism, we first check the detailed text description of operational risk events in our dataset under two extreme scenarios: extremely high workload environment and extremely low workload environment. We define the extremely high workload environment as top 10% highest workload observations. In this scenario, we find that each employee handles between 21 to 48 transactions per day, and in total there are 29 risk events. Similarly, in the 10% lowest workload observations, each employee handles around 0.5 to 1.5 transactions per day, and there are 11 risk events in total. Table 12 in our Appendix shows the detailed description of the 29 risk events under high workload scenario, while Table 13 presents the 11 risk events under low workload scenario. In Table 12, we can see that under high workload, employees tend to make errors or have quality degradation due to cognitive multitasking. For example, the risk event "Issued 3 million RMB business loan with maturity date as the next day" or "Issues 1 million business loan without the loan usage/interest rate" are such errors. In general, there is a variety of risk events in this table, but most of them seem to be simple mistakes due to negligence (more relevant to Execution, Delivery & Process Management type), which neither enhances the performance of employees nor do they appear malicious. By contrast, under low workload (Table 13), employees tend to make performance-seeking risks. For instance, the risk event "Client manager issued 5 million loans without the branch manager's signature" shows that the client manager used discretion without proper approval to issue the loan to the customer, probably in an attempt to reach certain performance target. In fact, all risk events here are related to issuing loans inappropriately, and issuing more loans is both a major component of employees' incentives and, at the same time, loans can be issued maliciously (more relevant to Clients, Products, & Business Practices type).

We next proceed to analyze statistically our preliminary results above through subset analyses. We want to see the impact of workload on the two risk types (Execution, Delivery & Process Management and Clients, Products, & Business Practices) separately. We therefore conduct our first subset analysis based on the two risk types, and present our results in Table 5. From Table 5, we can see that the impact of workload increases convexly for ED& PM risk events, but decreases convexly for CP&BP risk events. This finding indicates that, when workload is low, employees tend to make CP&BP errors (aggressive sales in our data), however the CP&BP error rate decreases as workload increases. On the other hand, when workload is high, employees tend to make ED&PM errors (data entry errors, failed mandatory reporting, and operation failure), and the ED&PM error rate increases as workload increases. Furthermore, we conduct subset analyses on the three categories: (i) data entry errors, (ii) failed mandatory reporting, and (iii) operation failure under ED&PM errors, and present our results in Table 6. From Table 6, we can see that all the results for data entry errors, failed mandatory reporting, and operation failure events hold consistently as in column (2) of Table 5.

In addition, we also conduct subset analysis on the five business types, and we present our results in Table 7. From Table 7, we can see that our main results hold for all the five business lines, and moreover the business lines of credit/loan and sales of wealth management have in general higher risk losses under the same workload (comparing both linear and non-linear effects).

5.4 Robustness Checks

In this subsection, we discuss several robustness checks for the impact of workload. We first consider several alternative specifications of our main model. To start with, we categorize our risk events into ten types through text mining, and we conduct robustness checks with alternative specification of risk types in our fixed-effects model. Details of the text mining analysis can be found in Appendix B. We show our estimation results in columns (1) and (2) of Table 8. We find that both the linear and non-linear effects hold consistently as in our main estimation.

Next, we consider the joint estimation of risk frequency and severity through the seemingly unrelated regression (SUR):

$$ErrorRate_{ijt} = \alpha_0 + \alpha_1 Load_{ijt} + \alpha_2 Load_{ijt}^2 + \boldsymbol{\alpha}\boldsymbol{X}_{kijt} + \eta_i + \xi_j + \tau_{t_m} + u_{ijt}, \tag{4}$$

$$Severity_{ijt} = \gamma_0 + \gamma_1 Load_{ijt} + \gamma_2 Load_{ijt}^2 + \boldsymbol{\gamma}\boldsymbol{X}_{ijt} + \eta_i^s + \xi_j^s + \tau_{t_m}^s + u_{ijt}^s, \tag{5}$$

where $Severity_{ijt}$ is the average severity level of operational risk events for employee i in branch j at month t. We present our estimation results in columns (3) and (4) of Table 8. We find the impact of workload on risk frequency (error rate) to be consistent with our main model, however workload has no statistically significant impact on risk severity. To understand the intuition why workload has no statistically significant impact on risk severity, we consider the following two examples. From the text description of the risk events, we find the severity level for most risk events depends more on the financial aspect of risk than on the operational aspect. For example, one such risk, "Issued collateral-based loan but did not collect collateral" has a severity score of 6. This severity score is mainly determined by the credit rating of the borrower, a given financial characteristics uncontrollable by the bank employees or by operational characteristics. Another example, "Issued 500,000 RMB loan with 0 interest rate" has a severity score of 5. Again, the severity score 5 here is mainly determined by the current interest rate, which workload of employees would have limited impact on.

We further conduct robustness checks with the alternative model specification—zero-inflated Poisson regression. Poisson regression is appropriate here for two reasons: i) our dependent variables (both risk frequency and severity) are count data; and ii) most OpRisk measurement models use a compound Poisson process to capture the convolution of risk frequency and severity (Cruz, 2002). However, our data set contains a lot of zeros, and hence, instead of a regular Poisson regression, we adopt the zero-inflated Poisson regression to deal with the excess of zeros. We also include the employee-related characteristics (time-invariant variables) and rural/urban dummy in the zero-inflated Poisson regression. We present the results in columns (5) and (6) of Table 8. Column (5) of Table 8 shows the estimation results of the non-zero part of the Poisson model, and the results further support our U-shape impact of workload on error rate.

Next, we consider employee-related characteristics (time-invariant variables) and rural/urban dummy in a random effects model as a robustness check. Columns (7) and (8) of Table 8 show consistent results as in our main model. Moreover, the dummy variable rural/urban is positively correlated with the error rate, which indicates the rural areas are subject to more severe operational risk. Finally, education level is negatively correlated with the error rate, which indicates that employees with higher levels of education are less likely to make errors. The coefficient of the experience level is not statistically significant.

In addition to SUR and zero-inflated Poisson, we also consider a spline regression of workload on error rate. Spline regression can be viewed as an extension of the linear models that are used to characterize the specific nonlinear relationship. It has an advantage of being a non-parametric approach, as it does not impose a specific (e.g., quadratic) functional form on the data (Friedman, 1991). Table 9 shows the spline regression results. As can be seen, the coefficient of workload for the piece-wise linear function when workload is low is negative starting from -10.112 at a 5% significance level, and it increases to positive with a value of 9.005 at a 5% significance level. This finding is consistent with a U-shape hypothesis. Furthermore, we conduct a two-line test and Lind test. The two-line test is used to test the U-shape relationship between x (independent variable) and y (dependent variable) with two separate lines that characterize the low and high value of x separately. The two-line test follows Simonsohn (2016). We have the p-value of 0.015 for the left line and 0.023 for the right line, which suggests the statistical significance of a U-shape. The Lind-test is another method to test the U-shape relationship, which characterizes both necessary and sufficient conditions for such a relationship. Following Lind and Mehlum (2010), we conduct the Lind test to validate the quadratic specification, and we have a p-value of 0.026 of the test, which rejects the null hypothesis of monotone or inverse U-shape, see Table 10.

Next, we discuss robustness checks with additional control variables and alternative IVs. First, we consider the interaction effects of the cumulative workload and workload. We find no statistically significant effects of the interaction term and the effects of all the other variables remain the same, see column (1) of Table 11. Second, we consider the serial correlation effects of the error rate with the lagged error variable. We find no statistically significant effects of the lagged error term and the effects of all the other variables remain the same, see column (2) of Table 11. Third, we consider the impact of cubic workload. We find the coefficient of the cubic workload variable is not statistically significant and the effects of all the other variables remain the same, see column (3) of Table 11. Fourth, we delete the two branches with high error frequency, and run our main estimation model with the remaining 47 branches and show consistent results in column (4) of Table 11. Finally, we consider alternative IVs—weather[7], lagged workload, and lagged quadratic workload. In particular, we consider both the linear and quadratic terms of temperature (the monthly average temperature near the branch location (Cachon et al., 2013)) and the "rain" variable (the amount of precipitation) as weather IVs. We show consistent results in column (5) of Table 11.

6 Discussion on the Optimal Staffing Level

Banks are profit centers and profit maximizers. However, the profit maximization problem is normally conducted in a more qualitative way based on managers' personal experience, rather than using a quantitative math model, and within some limits that are set by the corporate office. Therefore, one goal of our paper is to show to the bank's decision makers the importance of including operational risk in staffing decisions through a quantitative model. Our empirical results suggest a U-shape relationship between workload and

[7] https://www.worldweatheronline.com/lang/en-us/v2/historical-weather.aspx?locid=455878&root_id=382895&wc=local_weather&map=~/rugao-weather-history/jiangsu/cn.aspx

error rate. This finding can help banks make better capital allocation decisions on the optimal staffing level among retail branches so as to reduce operational risk losses. In this section, we propose a capital allocation model to help make the optimal staffing decision to reduce operational risk losses.

For the banks to decide on the capital allocation to offset the operational risk losses, they need to first estimate the potential losses. As mandated by regulation, a bank typically uses a frequency distribution to project the total number of loss events in a given time period and severity distribution to represent the potential loss amount of each risk event (Frachot et al., 2001; Guegan and Hassani, 2013a,b). The industry practice is to further assume that the frequency and severity distributions are independent. The total loss is then computed by the convolution of these two distributions using a compound model. In our case, the operational risk losses L_{ijt} of employee i at branch j in month t could be written as:

$$L_{ijt} = \sum_{n=1}^{N_{ijt}} X_n, \quad \text{for} \quad t = 1, \ldots, 16, \;\; i = 1, \ldots 49, \tag{6}$$

where X_n is the severity of each risk event at each of the 49 branches in each of the 16 months, and N_{ijt} is the monthly frequency of the risk events. The monthly risk frequency is found to be affected by the workload (a division of the number of transaction by the staffing level) through a U-shaped relationship based on our empirical investigation. In other words, keeping the transaction number constant, optimally changing the staffing levels for each business type (k) in each branch (j) at each month (t), denoted as $\boldsymbol{S} = \{S_{111}, \ldots, S_{kjt}, \ldots\}$, affects the risk frequency as follows:

$$ErrorRate_{ijt}(\boldsymbol{S}) = \alpha_0 + \alpha_1 Load_{ijt}(\boldsymbol{S}) + \alpha_2 Load^2_{ijt}(\boldsymbol{S}) + \boldsymbol{\alpha X}_{ijt} + \eta_i + \xi_j + \tau_{t_m} + u_{ijt}.$$

Given that

$$N_{ijt}(\boldsymbol{S}) = Tran_{ijt} \cdot ErrorRate_{ijt}(\boldsymbol{S}),$$

we can then obtain the frequency of operational risk events. With both X_n and $N_{ijt}(\boldsymbol{S})$ defined, the total loss for all the branches over time is

$$L(\boldsymbol{S}) = \sum_{t=1}^{T} \sum_{j=1}^{J} \sum_{i=1}^{I} \sum_{s=1}^{S_{kjt}} \sum_{n=1}^{N_{ijt}(\boldsymbol{S})} X_n.$$

Since X_n are assumed to be i.i.d, the expected total loss should be

$$E[L(\boldsymbol{S})] = \sum_{t=1}^{T} \sum_{j=1}^{J} \sum_{i=1}^{I} \sum_{s=1}^{S_{kjt}} N_{ijt}(\boldsymbol{S}) E[X_n]. \tag{7}$$

In our study, we focus on the optimal staffing rules with the same service capacity (675 employees). Therefore, the optimization problem is equivalent to

$$\min_{\boldsymbol{S}} \sum_{t=1}^{T} \sum_{j=1}^{J} \sum_{i=1}^{I} \sum_{s=1}^{S_{kjt}} N_{ijt}(\boldsymbol{S}). \tag{8}$$

Under the current practice of our focal bank, there is a fixed staffing policy, meaning employees are dedicated to the particular business type and branch, which may lead to an issue that the workload of some business lines and branches might be either too high or too low from time to time. On the other hand, however, based on our empirical investigation, operational risk events are minimized when the workload is neither too high nor too low. Therefore, we propose new staffing rules that add more flexibility to the banking processes. In particular, we consider three types of staffing rules: (i) full flexibility (employees can switch across any business types and branches), (ii) flexibility across business types (employees can only switch across business types in the same bank branch), and (iii) flexibility among branches (employees can switch branches but have to stay within the same business type). We consider two scenarios, one is that employee switch happens every month, and the other one is every quarter.

We first construct the benchmark case, where the staffing rule follows the current bank practice, but the workload of the same business type and branch is equally split among employees. Under this benchmark case, workload is more balanced among employees within the same business type in the same branch, and the total number of risk events is 1,397. Now we compare this benchmark case with the three types of new staffing rules. We first consider the full flexibility staffing rule with monthly switch. We find that under the full flexibility rule, the total number of risk events drops to 1,253, a 10.3% decrease. Next, we consider the monthly switch across business lines, and we find that the total number of risk events drops to 1,342, a 3.9% decrease. Finally, we consider the monthly switch among branches, and we find that the total number of risk events drops to 1,325, a 5.2% decrease.

Next, we consider the flexible staffing rule with a quarterly switch. Again, we first consider the full flexibility staffing rule with quarterly switch. We find that under the full flexibility rule, the total number of risk events drops to 1,306, a 6.5% decrease. Next, we consider the quarterly switch across business lines, and we find that the total number of risk events drops to 1,350, a 3.4% decrease. Finally, we consider the quarterly switch among branches, and we find that the total number of risk events drops to 1,336, a 4.4% decrease.

In general, keeping the risk severity level the same, we find that employing a flexible staffing rule can significantly reduce operational risk losses for more than 3% under different scenarios, which is practically significant given the 10-year annualized return in money market between 2004 and 2013 was about 1.7 % [8]. Moreover, we can achieve a more than 3% performance improvement by adding even a little bit of flexibility to the process by allowing employees to either switch their business lines in the same branch or switch branches within the same business lines on a quarterly basis.

[8] https://www.forbes.com/sites/advisor/2014/04/24/why-the-average-investors-investment-return-is-so-low/#71d11a2e111a

7 Conclusion

In this paper, we use a detailed operational risk data set gathered from a commercial bank to study the effects of workload on operational risks in terms of error rate and severity. We adopt an IV estimation strategy to address potential endogeneity issues. We find a U-shaped relationship between workload and error rate. That is, when the overall workload is low, increasing workload will reduce the error rate; however, when the overall workload is high, increasing workload increases the error rate. To explain the mechanism of such an empirical finding, we discover that under the low workload scenario employees tend to take performance-enhancing risks as workload increases. By contrast, under the high workload scenario, as workload further increases, employees tend to make more errors or have quality degradation due to cognitive multitasking. Although we find a U-shaped relationship between workload and error rate, we do not observe a statistically significant impact of workload on risk severity because the importance of exogenous external factors of the risk severity (losses), such as market conditions (e.g., interest rate, stock prices) and the value of the transaction, seem to outweigh the impact of endogenous factors, such as worker performance.

We find that banks' current staffing decisions create avoidable operational risk events because of imbalanced workloads among branches. We therefore propose a capital allocation model to help make the optimal staffing decision to reduce operational risk losses under the same service capacity (number of employees). In general, we find that keeping the risk severity level the same, employing a flexible staffing rule can significantly reduce operational risk losses by more than 3% under different scenarios. Moreover, we can achieve a more than 3% performance improvement by adding even a little bit of flexibility to the process by allowing employees to either switch their business lines in the same branch or switch branches with the same business lines on a quarterly basis.

Our study makes the following contributions to the literature. To the best of our knowledge, our paper is the first empirical work that analyzes the causes of operational risks from an OM perspective in the banking industry, while the previous studies tended to model OpRisk as an exogenous probability distribution. In particular, we establish a causal link between workload, an important work environment factor, and the error rate of operational risk events. Second, we revisit a growing area in operations management—the impact of workload on operational performance—and broaden our understanding about the importance of the workload and staffing decisions in the financial industry, which is understudied in the empirical OM literature. Third, our empirical study enables us to explain the variation in OpRisk events, so that we can build a capital allocation model to re-optimize bank staffing levels among branches and improve OpRisk management.

Finally, it is important to understand the limitations of our work and establish future research directions. First, although our dataset is unique with respect to the operational risk events collection, we only cover the category of data entry errors, accounting errors, failed mandatory reporting and negligent loss of client assets, which are mainly caused by bank workers. Clearly, there are still other types of operational risk events that we did not study in this paper. For example, "External Fraud" is committed by a third-party. "Damage to Physical Assets" entails those losses arising from natural disasters or terrorism/vandalism. "Business Disruption and System Failures" incur losses arising from the disruption of business/system failure. These events are mainly exogenous and beyond the scope of this paper. An interesting future research direction could be in conducting field experiments to study the internal fraud or external fraud events and explore incentive issues. Second, our data does not contain customer demographic information, because this type of information is too sensitive for most banks to share. Future research with more granular-level data on customer demographic information can focus on how the customer side heterogeneity would affect workload and thus operational risks. Third, our work focuses on the workload and operational risk events in the physical branches. Given the growing adoption of online and mobile banking, it is worth examining how online and mobile banking channels would affect OpRisk. Finally, our optimal staffing policy ignores other implications of changing staffing levels, such as profit and customer satisfaction in our optimization model, because the focus of our model is to illustrate the importance of considering operational risk losses when banks make staffing-level decisions. However, these features could serve as potential future research directions.

Appendix A: Tables

Table 3: Gross Losses by Event Type–Reported to ORX over the Period 2008-2012.

Risk Event Type	Total (Million €)	% of Total
Internal Fraud	€3,142	2.57%
External Fraud	€12,322	10.06%
Employment Practices & Workplace Safety	€2,844	2.32%
Clients, Products, & Business Practices	€77,505	63.28%
Damage to Physical Assets	€504	0.41%
Business Disruption & System Failures	€2,236	1.83%
Execution, Delivery & Process Management	€23,921	19.53%

Table 4: First-stage regression of Load and Loadsq.

	$Load_1$	$Load_1^2$
Tran_online	0.130***	0.112***
	(0.022)	(0.025)
E-visits	0.451**	0.447**
	(0.167)	(0.165)
Haus_load	0.127**	0.152**
	(0.047)	(0.056)
Haus_loadsq	0.087**	0.072**
	(0.032)	(0.026)
Cum_load	0.016*	-0.020
	(0.007)	(0.018)
Risk_Type (%)	Yes	Yes
Business_Type_Dummy	Yes	Yes
Employee_FE	Yes	Yes
Branch_FE	Yes	Yes
Month_Year_FE	Yes	Yes
N	9,954	9,954
R^2	0.419	0.416
$adj.R^2$	0.417	0.413
Prob>Chi-sq	<0.001	<0.001

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Subset Analysis on ED&PM and CP&BP

	(1)	(2)
	ED&PM	CP& BP
Load	0.174***	-0.732***
	(0.031)	(0.092)
$Load^2$	0.103***	0.397***
	(0.016)	(0.045)
Cum_load	-0.178***	-0.049***
	(0.019)	(0.037)
Business_Type_Dummy	Yes	Yes
Employee_FE	Yes	Yes
Branch_FE	Yes	Yes
Month_Year_FE	Yes	Yes
N	9,954	9,954
R^2	0.416	0.408
adj. R^2	0.413	0.406

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Subset Analysis on Categories under ED&PM

	(1)	(2)	(3)
	Data Entry	Reporting	Operation Fail
Load	0.196***	0.145***	0.108***
	(0.031)	(0.025)	(0.016)
Load^2	0.148***	0.124***	0.087***
	(0.027)	(0.021)	(0.015)
Cum_load	-0.128***	-0.146***	-0.208***
	(0.019)	(0.022)	(0.031)
Business_Type_Dummy	Yes	Yes	Yes
Employee_FE	Yes	Yes	Yes
Branch_FE	Yes	Yes	Yes
Month_Year_FE	Yes	Yes	Yes
N	9,954	9,954	9,954
R^2	0.389	0.405	0.396
adj. R^2	0.387	0.403	0.394

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: Subset Analysis on Business Lines

	(1)	(2)	(3)	(4)	(5)
	Individual Deposit/Withdrawal	Individual Credit/Loan	Individual Sales of Wealth Mgnt	Business Deposit/Withdrawal	Business Credit/Loan
Load	-0.506***	-0.492***	-0.501***	-0.502***	-0.499***
	(0.078)	(0.069)	(0.077)	(0.081)	(0.074)
Load^2	0.463***	0.475***	0.474***	0.469***	0.478***
	(0.059)	(0.061)	(0.057)	(0.052)	(0.051)
Cum_load	-0.112***	-0.134***	-0.128***	-0.118***	-0.135***
	(0.021)	(0.028)	(0.025)	(0.022)	(0.027)
Risk_Type (%)	Yes	Yes	Yes	Yes	Yes
Employee_FE	Yes	Yes	Yes	Yes	Yes
Branch_FE	Yes	Yes	Yes	Yes	Yes
Month_Year_FE	Yes	Yes	Yes	Yes	Yes
N	2,140	2,018	1,968	1,936	1,892
R^2	0.423	0.446	0.397	0.415	0.428
adj. R^2	0.421	0.444	0.395	0.413	0.425

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: Alternative Specification

	Text Mining		SUR		Zero-inflated Poisson		Random Effects	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Linear	Non-linear	Frequency	Severity	Count	Logit	Linear	Non-linear
Load	0.165***	-0.647***	-0.153***	0.611	-0.277***	0.191*	0.253***	-0.937***
	(0.029)	(0.112)	(0.012)	(0.602)	(0.078)	(0.080)	(0.0702)	(0.224)
$Load^2$	–	0.612***	0.126***	0.223	0.241*	0.130	–	0.869***
	–	(0.095)	(0.014)	(0.204)	(0.101)	(0.249)	–	(0.167)
Cum_load	-0.012***	-0.067***	-0.040***	-0.066***	-0.177***	-0.027	-0.201***	-0.063***
	(0.001)	(0.007)	(0.001)	(0.004)	(0.002)	(0.025)	(0.002)	(0.005)
Rural/Urban	–	–	–	–	0.228**	0.212**	0.229**	0.213**
	–	–	–	–	(0.084)	(0.077)	(0.084)	(0.078)
Edu_lvl	–	–	–	–	-0.012**	-0.033	-0.105**	-0.107**
	–	–	–	–	(0.004)	(0.032)	(0.038)	(0.039)
Exp_lvl	–	–	–	–	0.015	0.029	0.142	0.140
	–	–	–	–	(0.065)	(0.064)	(0.150)	(0.150)
Risk_Type (%)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Business_Type_Dummy	Yes	Yes	Yes	Yes	No	No	No	No
Employee_FE	Yes	Yes	Yes	Yes	No	No	No	No
Branch_FE	Yes	Yes	Yes	Yes	No	No	No	No
Month_Year_FE	Yes	Yes	Yes	Yes	No	No	No	No
N	9,954	9,954	9,954	9,954	294,126		9,954	9,954
R^2	0.254	0.226	0.563	0.312	0.105		0.426	0.433
adj. R^2	0.251	0.224	0.560	0.309	0.102		0.423	0.430

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note that for the Zero-inflated Poisson regression, we compute the pseudo R-squared.

Table 9: Spline Regression with Ten Equal Intervals

	Coefficient
0%-10%	-10.112*
	(4.247)
10%-20%	-6.364*
	(2.673)
20%-30%	-3.003*
	(1.261)
30%-40%	-1.246*
	(0.523)
40%-50%	-0.897*
	(0.376)
50%-60%	-0.112*
	(0.046)
60%-70%	0.236*
	(0.098)
70%-80%	1.879*
	(0.789)
80%-90%	4.339*
	(1.823)
90%-100%	9.005
	(10.022)
Cum_load	-0.044*
	(0.018)
Rural/Urban	0.078*
	(0.032)
Edu_lvl	-0.148*
	(0.062)
Exp_lvl	-0.103
	(0.105)
Risk_Type (%)	Yes
N	9,954
R^2	0.243
$adj.R^2$	0.240
Prob>Chi-sq	<0.0001

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 10: Lind and Mehlum U-shape test on Error Rate and Workload

	Lower bound	Upper bound
Interval	-0.534	15.558
Slope	-0.001	0.003
t-value	-2.025	1.957
$P > \|t\|$	0.023	0.026

Specification $f(x) = x^2$ Extreme point: 0.586

Overall test of presence of a U shape: t-value = 1.97, P > |t|=0.025

Note: H1: U shape vs. H0: Monotone or Inverse U shape

Table 11: Robustness Checks

	(1) Cum_Load	(2) Lagged Error	(3) Cubic Load	(4) 47 Branches	(5) Alternative IV
Load	-0.254***	-0.587***	-0.241***	-0.153***	-0.474***
	(0.028)	(0.081)	(0.002)	(0.003)	(0.081)
Loadsq	0.237***	0.552***	0.223***	0.134***	0.453***
	(0.052)	(0.075)	(0.046)	(0.002)	(0.075)
Cum_load	-0.168***	-0.123***	-0.167***	-0.175***	-0.235***
	(0.023)	(0.020)	(0.024)	(0.021)	(0.037)
Load*Cum_load	-0.0243	–	–	–	–
	(0.0303)	–	–	–	–
Lag_error	–	-0.655	–	–	–
	–	(0.641)	–	–	–
Load_cubic	–	–	0.082	–	–
	–	–	(0.101)	–	–
Risk_Type (%)	Yes	Yes	Yes	Yes	Yes
Business_Type_Dummy	Yes	Yes	Yes	Yes	Yes
Employee_FE	Yes	Yes	Yes	Yes	Yes
Branch_FE	Yes	Yes	Yes	Yes	Yes
Month_Year_FE	Yes	Yes	Yes	Yes	Yes
N	3880	3880	3880	3635	3880
R^2	0.536	0.588	0.543	0.529	0.507
adj. R^2	0.534	0.585	0.539	0.525	0.503

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 12: Operational Risk Events under High Workload

1	Issued 3 million RMB business loan with maturity date as the next day.
2	Issued collateral based loan but did not collect collateral.
3	Issued RMB 500,000 loan with 0 interest rate.
4	Issued RMB 800,000 loan with maturity date as 02/2013.
5	Opened new client account without collecting client information into the system.
6	Did not record the withdrawal client ID number.
7	Client manager issued a business loan of 100,000 RMB to a company with an expired license.
8	Deposit without currency tag.
9	Issued 500,000 RMB business loan on 12/2014 with maturity date as 12/2013.
10	Issued 1 million business loan without an interest rate.
11	Opened new account with wrong client name.
12	Issued 100,000 personal loan to a client with credit card delinquency of 43,817.65.
13	Issued 1 million business loan without borrower's name.
14	Issued 1 million business loan without the loan usage.
15	XXX did not log out computer system after the transaction is done.
16	XXX did not lock the safe of the client.
17	Opened the ATM machine without locking it.
18	Issued 500,000 business loan with different interest rates in the system and on paper.
19	Total withdrawal amount is different between the central system and client manager's system.
20	Typed the wrong deposit date.
21	Large amount cash withdrawal without client signature.
22	Opened asset management account without validating client ID.
23	Did not double check the capital in reserve of the branch at the beginning of the day.
24	Loan issuance with different company license numbers in the system and on paper.
25	Asset management transaction without the specification of commission fee.
26	Did not double check the collateral at the end of the season.
27	XXX did not log off computer system when leaving back home.
28	A deposited account without client ID.
29	Issued 1 million business loan with maturity date as the previous month.

Table 13: Operational Risk Events under Low Workload

1	Client manager issued 5 million loan without the branch manager's signature.
2	Issued business loan 1 million to a firm manager without checking his personal debt status.
3	Issued business loan 2 million to a firm without checking the company operations and profits.
4	Issued personal loan to client with age exceeds the allowance limit.
5	Issued 2 million business loan without checking the owner's credit score.
6	Issued 1 million business loan to client with existing business loans but without explanation.
7	Issued 200,000 personal loan without credit check.
8	Issued 1 million business loan to client with default history but without explanation.
9	Issued 1 million business loan to a client who changed the name of guarantor several times.
10	Issued 2 million business loan without validating the purpose of loan.
11	Issued 1 million business loan with wrong interest rate calculation (lower than actual rate).

Appendix B: Text Mining of Risk Events

In this part, we further conduct text mining analysis of our risk events. Utilizing the text mining analysis, we now classify all the risk events in our data set into 10 types (see Figure 1 in our revised paper), i.e., verification failure, guarantor issue, applicant quality, default issue, information inconsistency, operation failure, interest rate or maturity issue, date error, collateral issue, material incomplete. To be more specific, we apply the Han Language Processing (HanLP) package contained in Python to analyze the Chinese text of event description. Generally speaking, we first use the POS tagger to identify the part of speech of each word in each event description in terms of nouns, noun phrases, verbs, etc. We then use an association mining algorithm following Manning and Schutze (1999) to further cluster the candidate features. The association mining here finds correlations among a set of data. After getting the feature (topic) of each event, we use association mining to group similar features together. In the end, we asked contacts in our focal bank to label and classify the topics of all events into 10 categories. We summarize our algorithm in the following table.

Algorithm 1 Topic Extraction

Input: event description E
for sentence $I \in E$ **do**
    POS tagger for each word in I
    Select nouns, verbs, noun phrases, and verb phrases in I as candidates for features of E
    Association mining to group items in all the sentences of E that appear together frequently to find frequent features (itemsets)
end for
Output: set of features f of event E

Finally, the 10 risk event types can be further classified into two general types of operational risk events according to the definition of Basel Committee[9], namely "clients, products and business practices" and "execution, delivery and process management". Utilizing text mining analysis, we can classify all the risk events in our data set into 10 types, see Figure 1 below. The five types – verification failure, guarantor issue, applicant quality, default issue, information inconsistency (a total of 542 events), could be categorized into "clients, products and business practices" level 1 type, since they could be special examples of Fiduciary breaches / guideline violations, account churning, failure to investigate client per guidelines, etc. (level 3 category in the link). The remaining five types – operation failure, interest rate or maturity issue, date error, collateral issue, material incomplete (a total of 1245 events), could be categorized into "execution, delivery and process management", since they could be special examples of Data entry, maintenance or loading error, delivery failure, etc. (level 3 category in the link). Note that our data set only contains risk events occuring in the physical branches, and it does not include those that happened through online banking system.

The ten event categories of our text mining results can also be grouped into the four categories labeled by our focal bank, as follows

(i) Data entry errors: interest rate or maturity issue, information inconsistency, date error;

[9]Referece: https://www.bis.org/bcbs/qisoprisknote.pdf (page 12-13).

(ii) Failed mandatory reporting: material incomplete, verification failure;

(iii) Operation failure;

(iv) Aggressive sales: guarantor issue, default issue, applicant quality, collateral issue.

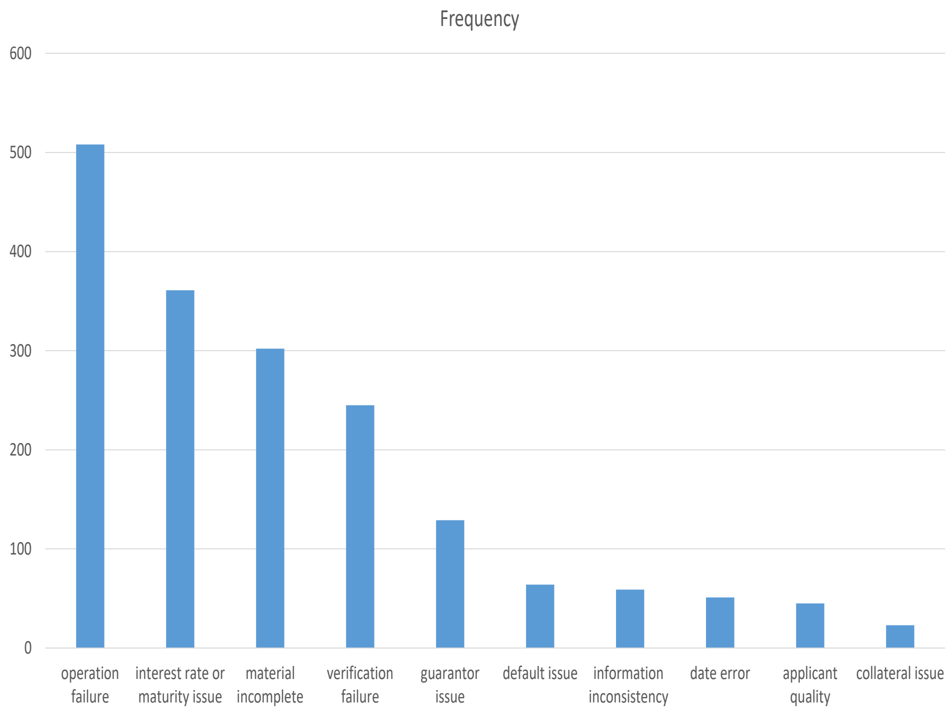

Figure 1: Risk Types Via Text Analysis

References

Allon, Gad, Maxime Cohen, Park Sinchaisri. 2018. The impact of behavioral and economic drivers on gig economy workers. *Available at SSRN 3274628* .

Altman, E. I., A. Saunders. 1997. Credit risk measurement: Developments over the last 20 years. *Journal of Banking & Finance* **21**(11) 1721–1742.

Aral, S., E. Brynjolfsson, M. Van Alstyne. 2012. Information, technology, and information worker productivity. *Information Systems Research* **23**(3-part-2) 849–867.

Ashby, S. 2010. The 2007-2009 financial crisis: learning the risk management lessons. *Financial Services Research Forum, Nottingham*.

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

Basel Committee on Banking Supervision. 2011. Principles for the sound management of operational risk. BIS consultative document.

Basel Committee on Banking Supervision. June 2012. The internal audit function in banks. BIS consultative document.

Bendoly, E. 2011. Linking task conditions to physiology and judgment errors in RM systems. *Production and Operations Management* **20**(6) 860–876.

Bendoly, E., K. Donohue, K.L. Schultz. 2006. Behavior in operations management: Assessing recent findings and revisiting old assumptions. *Journal of Operations Management* **24**(6) 737–752.

Bendoly, E., M. Prietula. 2008. In the 'Zone': The role of evolving skill and transitional workload on motivation and realized performance in operational tasks. *International Journal of Operations & Production Management* **28**(12) 1130–1152.

Berry, J. J., A. Tucker. 2016. Past the point of speeding up: The negative effects of workload saturation on efficiency and patient severity. *Management Science* **63**(4) 1042–1062.

Bocker, K., C. Kluppelberg. 2005. Operational VaR: a closed-form approximation. *RISK-LONDON-RISK MAGAZINE LIMITED* **18**(12) 90.

Boudreau, J. W., W. Hopp, J. O. McClain, L. J. Thomas. 2003. On the interface between operations and human resources management. *Manufacturing and Services Operations Management* **5**(3) 179–202.

Buell, R. W., D. Campbell, F. X. Frei. 2016. How do customers respond to increased service quality competition? *Manufacturing & Service Operations Management* **18**(4) 585–607.

Cachon, G., S. Gallino, M. Olivares. 2013. Does adding inventory increase sales? evidence of a scarcity effect in us automobile dealerships. *Working Paper* .

Cakir, A., DJ Hart, TFM Stewart. 1980. *Visual display terminals: A manual covering ergonomics, workplace design, health and safety, task organization*. Wiley.

Campbell, D., F. Frei. 2004. The persistence of customer profitability: Empirical evidence and implications from a financial services firm. *Journal of Service Research* **7**(2) 107–123.

Campbell, D., F. Frei. 2010. Cost structure, customer profitability, and retention implications of self-service distribution channels: Evidence from customer behavior in an online banking channel. *Management Science* **56**(1) 4–24.

Charron, S., E. Koechlin. 2010. Divided representation of concurrent goals in the human frontal lobes. *Science* **328**(5976) 360–363.

Chernobai, A. S., S. T. Rachev, F. J. Fabozzi. 2007. *Operational Risk: A Guide to Basel II Capital Requirements, Models, and Analysis*. John Wiley and Sons, New York.

Chuang, H. H., R. Oliva, O. Perdikaki. 2016. Traffic-based labor planning in retail stores. *Production and Operations Management* **25**(1) 96–113.

Clark, J. R., R. S. Huckman, B. R. Staats. 2013. Learning from customers: Individual and organizational effects in outsourced radiological services. *Organization Science* **24**(5) 1539–1557.

Cruz, M. G. 2002. *Modeling, Measuring and Hedging Operational Risk*. John Wiley and Sons, New York.

Dalton, A. N., S. A. Spiller. 2012. Too much of a good thing: The benefits of implementation intentions depend on the number of goals. *Journal of Consumer Research* **39**(3) 600–614.

Deci, E.L., J.P. Connell, R.M. Ryan. 1989. Self-determination in a work organization. *Journal of Applied Psychology* **74**(4) 580.

Deloitte. 2013. Directors, CFOs Share Views on Risk, but Diverge on Time Spent. URL `http://deloitte.wsj.com/riskandcompliance/2013/04/16/`.

Donahue, E.M., R.W. Robins, B.W. Roberts, O.P. John. 1993. The divided self: Concurrent and longitudinal effects of psychological adjustment and social roles on self-concept differentiation. *Journal of Personality and Social Psychology* **64**(5) 834–846.

Dowd, K. 2007. *Measuring market risk*. John Wiley & Sons.

Fisher, M., S. Gallino, S. Netessine. 2017. Setting retail staffing levels: A methodology validated with implementation. *Working paper* .

Frachot, A., P. Georges, T. Roncalli. 2001. Loss distribution approach for operational risk .

French, K. R., W. G. Schwert, R. F. Stambaugh. 1987. Expected stock returns and volatility. *Journal of financial Economics* **19**(1) 3–29.

Friedman, J. H. 1991. Multivariate adaptive regression splines. *The Annals of Statistics* 1–67.

Guegan, D., B. K. Hassani. 2013a. Operational risk: A basel ii++ step before basel iii. *Journal of Risk Management in Financial Institutions* **6**(1) 37–53.

Guegan, D., B. K. Hassani. 2013b. Using a time series approach to correct serial correlation in operational risk capital calculation. *The Journal of Operational Risk* **8**(3) 31.

Healy, P., G. Serafeim. 2019. How to scandal-proof your company a rigorous compliance system is not enough. *HARVARD BUSINESS REVIEW* **97**(4) 42–+.

Hitt, L. M., F. X. Frei. 2002. Do better customers utilize electronic distribution channels? the case of pc banking. *Management Science* **48**(6) 732–748.

Holmstrom, B., P. Milgrom. 1991. Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, & Organization* **7** 24–52.

Hora, M., R. D. Klassen. 2013. Learning from others' misfortune: factors influencing knowledge acquisition to reduce operational risk. *Journal of Operations Management* **31**(1) 52–61.

Jaeker, J. B., A. L. Tucker. 2016. Past the point of speeding up: The negative effects of workload saturation on efficiency and patient severity. *Management Science* **63**(4) 1042–1062.

Jarrow, R. A. 2008. Operational risk. *Journal of Banking and Finance* **32**(5) 870–879.

Jarrow, R. A., J. Oxman, Y. Yildirim. 2010. The cost of operational risk loss insurance. *Review of Derivatives Research* **13**(3) 273–295.

Jarrow, R. A., S. M. Turnbull. 1995. Pricing derivatives on financial securities subject to credit risk. *The journal of finance* **50**(1) 53–85.

Kaplan, R. S., A. Mikes. 2012. Managing risks: a new framework. *Harvard Business Review* **90**(6).

KC, D. S. 2013. Does multitasking improve performance? Evidence from the emergency department. *Manufacturing & Service Operations Management* **16**(2) 168–183.

KC, D. S., C. Terwiesch. 2009. Impact of workload on service time and patient safety: An econometric analysis of hospital operations. *Management Science* **55**(9) 1486–1498.

KC, D. S., C. Terwiesch. 2012. An econometric analysis of patient flows in the cardiac intensive care unit. *Manufacturing and Services Operations Management* **14**(1) 50–65.

Kennedy, P. 2003. *A Guide to Econometrics*. MIT Press, Cambridge, MA.

Kesavan, S., V. Mani. 2015. An overview of industry practice and empirical research in retail workforce management. *Retail Supply Chain Management*. Springer, 113–145.

Kuntz, L., R. Mennicken, S. Scholtes. 2014. Stress on the ward: Evidence of safety tipping points in hospitals. *Management Science* **64**(1) 754–771.

Lapré, M. A., I. M. Nembhard. 2011. Inside the organizational learning curve: Understanding the organizational learning process. *Foundations and Trends® in Technology, Information and Operations Management* **4**(1) 1–103.

Latham, G.P., E.A. Locke. 1979. Goal setting: A motivational technique that works. *Organizational Dynamics* **8**(2) 68–80.

Leippold, M., P. Vanini. 2003. The quantification of operational risk. Available at SSRN 481742.

Lind, T., H. Mehlum. 2010. With or without U? the Appropriate Test for a U-Shaped Relationship. *Oxford Bulletin of Economics and Statistics* **72**(1) 109–118.

Locke, E.A. 1968. Toward a theory of task motivation and incentives. *Organizational Behavior and Human Performance* **3**(2) 157–189.

Lupien, S.J., F. Maheu, M. Tu, A. Fiocco, T.E. Schramek. 2007. The effects of stress and stress hormones on human cognition: Implications for the field of brain and cognition. *Brain and Cognition* **65**(3) 209–237.

Macchiavello, R., A. Morjaria. 2015. The value of relationships: evidence from a supply shock to Kenyan rose exports. *American Economic Review* **105**(9) 2911–45.

Mani, V., S. Kesavan, J. M. Swaminathan. 2015. Estimating the impact of understaffing on sales and profitability in retail stores. *Production and Operations Management* **24**(2) 201–218.

Manning, C., H. Schutze. 1999. Foundations of statistical natural language processing. MIT Press, Cambridge, MA.

Marrison, C. 2005. The fundamentals of risk measurement. *The Mathematical Intelligencer* **27**(2) 83–83.

Neil, M., N. Fenton, M. Tailor. 2005. Using bayesian networks to model expected and unexpected operational losses. *Risk Analysis* **25**(4) 963–972.

Oliva, R., J.D. Sterman. 2001. Cutting corners and working overtime: Quality erosion in the service industry. *Management Science* **47**(7) 894–914.

Parkinson, C. 1958. Parkinsons Law: The pursuit of progress.

Perdikaki, O., S. Kesavan, J. M. Swaminathan. 2012. Effect of traffic on sales and conversion rates of retail stores. *Manufacturing & Service Operations Management* **14**(1) 145–162.

Peters, L.H., E.J. O'Connor. 1980. Situational constraints and work outcomes: The influences of a frequently overlooked construct. *Academy of Management Review* **5**(3) 391–397.

Powell, A., S. Savin, N. Savva. 2012. Physician workload and hospital reimbursement: Overworked physicians generate less revenue per patient. *Manufacturing and Services Operations Management* **14**(4) 512–528.

Scharfman, J. A. 2008. *Hedge Fund Operational Due Diligence: Understanding the Risks*. John Wiley and Sons, New York.

Schmidt, A. M., R. P. DeShon. 2007. What to do? the effects of discrepancies, incentives, and time on dynamic goal prioritization. *Journal of Applied Psychology* **92**(4) 928.

Setyawati, L. 1995. Relation between feelings of fatigue, reaction time and work productivity. *Journal of Human Ergology* **24**(1) 129–135.

Shah, R., G. P. Ball, S. Netessine. 2017. Plant operations and product recalls in the automotive industry: An empirical investigation. *Management Science* **63**(8) 2439–2459.

Sheldon, Michael. 2016. Income targeting and the ridesharing market. *working paper* .

Simonsohn, U. 2016. Two-lines: The first valid test of ushaped relationships. *Working Paper* .

Staats, B. R., F. Gino. 2012. Specialization and variety in repetitive tasks: Evidence from a Japanese bank. *Management Science* **58**(6) 1141–1159.

Tan, T. F., S. Netessine. 2014. When does the devil make work? An empirical study of the impact of workload on worker productivity. *Management Science* **60**(6) 1574–1593.

Ton, Z., R. S. Huckman. 2008. Managing the impact of employee turnover on performance: The role of process conformance. *Organization Science* **19**(1) 56–68.

Wooldridge, J. M. 2010. *Econometric analysis of cross section and panel data*. MIT press.

Xu, Yuqian, Mor Armony, Anindya Ghose. 2016. The effect of online reviews on physician demand: A structural model of patient choice. *Available at SSRN* .

Xue, M., L. M. Hitt, P. T. Harker. 2007. Customer efficiency, channel usage and firm performance in retail banking. *Manufacturing and Services Operations Management* **9**(4) 535–558.