

CITY UNIVERSITY OF HONG KONG
香港城市大學

**The Impact of Intelligent Auditing on
Employee Behavior in Commercial Banks:
From the Perspective of Algorithm
Transparency**
商業銀行智能審計對員工行為的影響分析-基於算
法透明度視角

Submitted to
College of Business
商學院
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Business Administration
工商管理博士學位

by

He Hongyu
何紅宇

December 2025
二零二五年十二月

摘要

在金融科技快速發展的背景下，商業銀行正在加速推進數智化轉型，廣泛引入人工智慧、大數據、區塊鏈等新興技術，以期在內部治理中實現效率與風險控制的雙重提升。特別是在內部審計領域，傳統依賴人工抽樣和事後檢查的模式逐漸被智能審計所取代。智能審計系統能夠借助算法模型實現數據的即時採集、業務全面覆蓋與智能預警，從而在審計效率、風險識別精度以及組織治理能力方面展現出前所未有的優勢。這一變革不僅契合了監管部門對審計數位化的政策導向，也為金融機構促進合規經營與穩健發展提供了有力支撐。

隨著智能審計系統在商業銀行的廣泛應用，內部審計實現抽樣審計向全量審計、識別單業務條線風險向全面識別關聯風險、從事後查處向事前和事中預警轉變，從而具有即時性、及時性、持續性、全面性等特點。一方面，智能審計幫助員工及時發現潛在風險、強化合規意識，從而有效減少違規操作並提前規避風險；另一方面，智能審計系統的全面應用與強大功能，可能打破員工的心理安全邊界，使其產生不安與壓力，還會激發員工的防禦傾向，使其在工作中趨於謹慎與被動，甚至抑制主動性行為。由此，企業在引入和不斷發展智能審計系統的同時，也可能引發員工層面複雜而矛盾的心理與行為反應，成為亟需深入研究的重要議題。同時，我國內部審計發展趨勢體現在審計理念由監督制約向管理服務轉變、審計職能由查錯糾弊向增加價值轉變，在內部審計數智化轉型背景下，內部審計如何利用智能審計系統提供管理服務、賦能業務部門，增加組織價值，也值得深入研究和探討。

基於上述背景，本研究引入心理安全理論與資源保存理論作為核心分析框架，試圖揭示智能審計在組織應用中對員工心理與行為的影響機制。心理安全理論強調，員工在組織中的安全感是其積極投入和主動行為的前提，而智能審計系統的引入和發展可能打破這種平衡，誘發不安全感 and 防禦反應。資源保存理論則指出，個體在面對外部壓力時會傾向於保護和保存有限資源，這使得員工在智能審計情境下可能表現出更加謹慎和防禦性的行為。結合這兩大理論，本研究提出以下研究問題：第一，智能審計覆蓋度如何影響員工的不安全感和防禦傾向？進而，這些心理反應如何分別作用於合規行為與主

動性行為？第二，不安全感與防禦傾向是否在智能審計覆蓋度與員工行為結果之間發揮仲介機制？第三，算法透明度在其中起到怎樣的調節作用？其是否呈現出既能緩解不安全感、又能強化防禦傾向的雙刃劍效應？

為深入探討上述問題，本研究以商業銀行 717 名一線及管理崗位員工為調查對象，採用結構化問卷的形式系統收集數據。問卷涵蓋了智能審計覆蓋度、算法透明度、不安全感、防禦傾向、合規行為與主動性行為等關鍵變數，旨在通過量化方法檢驗變數之間的關係。結果表明，智能審計覆蓋度顯著增加了員工的不安全感與防禦傾向，並通過不同的心理路徑最終影響其行為結果：一方面，增強了員工的合規行為，使其在面對風險時更加謹慎守規；另一方面，卻抑制了員工的主動性行為，削弱了其自發改進的動力。進一步的調節效應檢驗顯示，算法透明度在這一過程中發揮了積極作用：較高水準的透明度能夠有效緩解員工的不安全感和防禦傾向，從而減輕智能審計帶來的心理壓力，並在一定程度上增加員工的行為積極性。

通過揭示智能審計覆蓋度與算法透明度對員工心理與行為的作用機制，本研究在理論上：第一，拓展了內部審計領域的研究視角。本研究突破了以往聚焦審計效率與治理績效的局限，將研究延伸至員工心理與行為層面，揭示了智能審計不僅塑造組織治理模式，還深刻影響員工的心理狀態與行為表現，從而為理解 AI 技術在組織中的深層作用提供了新的解釋框架。第二，豐富了心理安全理論與資源保存理論的應用。本研究將兩大理論引入智能審計情境，發現智能審計覆蓋度會通過增加不安全感與防禦傾向而影響員工的合規意願與主動性行為。這一發現不僅拓展了心理安全理論在智能審計情境下的適用性，也說明了資源保存理論在解釋員工防禦性反應方面的價值。第三，補充了算法透明度相關研究。本研究發現，算法透明度在智能審計與員工心理反應之間發揮積極調節作用，能夠有效緩解員工的不安全感與防禦傾向，降低因不確定性帶來的焦慮，從而改善員工的行為反應。該結論為算法透明度的實際應用提供了理論支持。第四，揭示了不安全感與防禦傾向的仲介機制。本研究驗證了二者在智能審計覆蓋度與員工合規行為和主動性行為之間的傳導作用路徑，揭示了智能審計如何通過影響員工態度，進而作用於其行為表現，深化了對員工態度-行為鏈條的理解。

在實踐層面，本研究為商業銀行及其他行業提供了多方面啟示：第一，

組織在引入智能審計系統時，應在效率提升與員工心理安全之間保持動態平衡。內部審計部門在利用智能審計系統更好地發揮監督制約作用的同時，也可以利用其提升管理服務能力，如通過開放部分審計模型給被審計單位使用等方式，協助被審計單位提高風險防範能力，並提升員工心理安全感和主動性行為。第二，適度的透明度有助於員工理解審計邏輯，降低因不確定性帶來的不安全感和防禦傾向，從而緩解智能審計的負面心理效應。但過高的透明度則可能導致算法的權威性與預測能力下降，從而導致審計監督和約束能力的下降。第三，組織還應通過培訓與溝通機制，營造支持性氛圍，幫助員工正確看待智能審計系統的作用與價值。通過開放式溝通，員工不僅能理解制度背後的治理目標，也更容易在合規與主動之間尋求平衡。

關鍵字：智能審計；算法透明度；心理安全；不安全感；防禦傾向；合規行為；主動性行為；審計理念；審計職能

Abstract

Against the backdrop of the rapid development of financial technology, commercial banks are accelerating their digital and intelligent transformation. They have extensively adopted emerging technologies such as artificial intelligence, big data, and blockchain, aiming to achieve dual improvements in both efficiency and risk control within their internal governance. This shift is particularly evident in the field of internal auditing, where the traditional model—relying on manual sampling and post-event inspections—has gradually been replaced by intelligent auditing. Leveraging algorithmic models, intelligent auditing systems enable real-time data collection, comprehensive business coverage, and intelligent early warning. In doing so, they demonstrate unprecedented advantages in audit efficiency, accuracy of risk identification, and organizational governance capabilities. This transformation not only aligns with the regulatory authorities' policy orientation toward audit digitalization, but also provides strong support for financial institutions to promote compliant operations and sound development.

With the widespread application of intelligent auditing systems in commercial banks, internal auditing has undergone a fundamental transformation: it has shifted from sample-based auditing to full-scope auditing, from identifying risks in individual business lines to comprehensively detecting interrelated risks, and from post-event investigation to pre-event and in-process early warning. As a result, internal auditing now exhibits characteristics of real-time responsiveness, timeliness, continuity, and comprehensiveness. On the one hand, intelligent auditing helps employees promptly identify potential risks and strengthen compliance awareness, thereby effectively reducing irregular operations and proactively mitigating risks. On the other hand, the extensive application and powerful functions of intelligent auditing systems may breach the boundaries of employees' psychological safety, triggering feelings of unease and pressure. Furthermore, they may induce defensive tendencies among employees, making them more cautious and passive in their work, and even inhibiting their proactive behaviors. Consequently, while enterprises introduce and continuously advance intelligent auditing systems, they may also trigger complex and contradictory psychological and behavioral responses among employees—an issue that urgently

requires in-depth research. Meanwhile, the development trend of internal auditing in China is reflected in two key shifts: the evolution of auditing concepts from "supervision and restriction" to "management and service," and the transformation of auditing functions from "error detection and correction" to "value addition." Against the backdrop of the digital and intelligent transformation of internal auditing, how internal auditing can utilize intelligent auditing systems to provide management services, empower business departments, and enhance organizational value also merits in-depth exploration and discussion.

Based on the above context, this study adopts Psychological Safety Theory and Conservation of Resources Theory as its core analytical frameworks, seeking to unravel the mechanism through which intelligent auditing influences employees' psychology and behaviors in organizational settings. Psychological Safety Theory emphasizes that employees' sense of safety within an organization is a prerequisite for their active engagement and proactive behaviors. However, the introduction and development of intelligent auditing systems may disrupt this balance, triggering insecurity and defensive reactions. Conservation of Resources Theory, by contrast, posits that when facing external pressures, individuals tend to protect and preserve their limited resources. This explains why employees may exhibit more cautious and defensive behaviors in the context of intelligent auditing. Integrating these two theories, this study proposes the following research questions: First, how does the coverage of intelligent auditing affect employees' sense of insecurity and defensive tendencies? Furthermore, how do these psychological responses impact compliance behaviors and proactive behaviors, respectively? Second, do insecurity and defensive tendencies play a mediating role between the coverage of intelligent auditing and employees' behavioral outcomes? Third, what regulatory role does algorithmic transparency play in this process? Does it exhibit a double-edged sword effect that can both alleviate insecurity and strengthen defensive tendencies?

To address these questions in depth, this study conducted a survey among 717 frontline and managerial employees of commercial banks, collecting data systematically through structured questionnaires. The questionnaires covered key variables, including the coverage of intelligent auditing, algorithm transparency, sense of insecurity, defensive tendencies, compliance behaviors, and proactive behaviors. The goal was to test the relationships between these variables using

quantitative methods. The results indicate that the coverage of intelligent auditing significantly increases employees' sense of insecurity and defensive tendencies, which in turn influence their behavioral outcomes through distinct psychological pathways: on one hand, it enhances employees' compliance behaviors, making them more cautious and rule-abiding when facing risks; on the other hand, it inhibits employees' proactive behaviors and weakens their motivation for spontaneous improvement. Further tests on moderating effects reveal that algorithm transparency plays a positive role in this process: a higher level of transparency can effectively alleviate employees' sense of insecurity and defensive tendencies, thereby reducing the psychological pressure caused by intelligent auditing and to some extent boosting employees' behavioral initiative.

By clarifying the mechanism through which the coverage of intelligent auditing and algorithm transparency affect employees' psychology and behaviors, this study makes several theoretical contributions: First, it expands the research perspective in the field of internal auditing. Breaking free from the previous focus on audit efficiency and governance performance, this study extends research to the realm of employees' psychology and behaviors. It reveals that intelligent auditing not only shapes organizational governance models but also exerts a profound impact on employees' psychological states and behavioral performance, thus providing a new explanatory framework for understanding the in-depth role of AI technology in organizations. Second, it enriches the application of Psychological Safety Theory and Conservation of Resources Theory. By introducing these two theories into the context of intelligent auditing, the study finds that the coverage of intelligent auditing affects employees' willingness to comply and their proactive behaviors by increasing their sense of insecurity and defensive tendencies. This finding not only broadens the applicability of Psychological Safety Theory in the context of intelligent auditing but also demonstrates the value of Conservation of Resources Theory in explaining employees' defensive responses. Third, it supplements existing research on algorithm transparency. The study confirms that algorithm transparency plays a positive moderating role between intelligent auditing and employees' psychological responses: it can effectively alleviate employees' sense of insecurity and defensive tendencies, reduce anxiety caused by uncertainty, and thereby improve employees' behavioral responses. This conclusion provides theoretical

support for the practical application of algorithm transparency. Fourth, it uncovers the mediating mechanism of insecurity and defensive tendencies. This study verifies the transmission path of the two factors between intelligent audit coverage and employees' compliance behaviors and proactive behaviors, reveals how intelligent audit affects employees' attitudes and then acts on their behavioral performance, and deepens the understanding of the employee attitude-behavior chain.

At the practical level, this study offers multiple insights for commercial banks and other industries: First, when introducing intelligent auditing systems, organizations should maintain a dynamic balance between efficiency improvement and employees' psychological safety. While using intelligent auditing systems to strengthen their supervision and restriction functions, internal audit departments can also leverage these systems to enhance their management service capabilities. For instance, by opening part of the audit models to audited units for their use, they can help these units improve their risk prevention capabilities while also boosting employees' psychological safety and proactive behaviors. Second, an appropriate level of transparency helps employees understand the logic of auditing, reduces insecurity and defensive tendencies caused by uncertainty, and thus mitigates the negative psychological effects of intelligent auditing. However, excessively high transparency may undermine the authority and predictability of algorithms, thereby weakening the effectiveness of audit supervision and restraint. Third, organizations should also establish training and communication mechanisms to foster a supportive environment, helping employees develop a correct understanding of the role and value of intelligent auditing systems. Through open communication, employees can not only grasp the governance objectives behind the system but also more easily strike a balance between compliance and proactivity.

Keywords: Intelligent Auditing; Algorithm Transparency; Psychological Safety; Sense of Insecurity; Defensive Tendencies; Compliance Behavior; Proactive Behavior; Auditing Concept; Auditing Function

CITY UNIVERSITY OF HONG KONG
Qualifying Panel and Examination Panel

Surname: HE
First Name: Hongyu
Degree: Doctor of Business Administration
College/Department: College of Business

The Qualifying Panel of the above student is composed of:

Supervisor(s)

Prof. ZHU Xindong Department of Accountancy
City University of Hong Kong

Prof. LU Xianghua Department of Information Management and
Information Systems
Fudan University

Qualifying Panel Member(s)

Prof. WU Haibin Department of Accountancy
City University of Hong Kong

Dr. LIU Shengming School of Management
Fudan University

This thesis has been examined and approved by the following examiners:

Prof. YU Yangxin Department of Accountancy
City University of Hong Kong

Prof. ZHU Xindong Department of Accountancy
City University of Hong Kong

Prof. ZHANG Cheng Department of Information Management and
Information Systems
Fudan University

Prof. LU Xianghua Department of Information Management and
Information Systems
Fudan University

Prof. SU Lixin School of Accounting and Finance
The Hong Kong Polytechnic University

致謝

“霜染楓紅鋪野徑，風搖葦白映寒塘”，抬眼望見窗外深秋的景致，階前楓葉凝霜、塘邊葦絮搖風，才驚覺這段城大與復旦聯合培養的求學時光，竟已隨季節流轉至尾聲。猶記初入課堂，面對複雜理論模型的茫然無措；而今伏案定稿，對學術研究已初窺門徑。滬港雙城的學術之旅、課堂討論的熱烈聲響、家人守望的溫暖目光，都藏在這“霜楓葦白”的時光刻度裏。此刻，伴著深秋的沉靜與澄澈，謹以真摯字句，致謝所有照亮我學術征途的師長、助研、同窗與親友，是你們的托舉，讓我得以在求知路上穩步前行，不負這一段珍貴歲月。

首先，我要深深感謝我的兩位導師——復旦大學盧向華教授和香港城市大學朱鑫東教授。盧向華教授以開闊的學術視野、敏銳的學術洞察力、活躍的學術思想、淵博的學術知識和深厚的科研經驗幫助我確立研究方向和課題選擇，帶領我們攻克一個個難關，在研究規劃構思、理論模型設計、論文寫作指導等每一個學術環節，盧老師都給予了悉心指導。盧老師務實的作風、嚴謹的治學精神令我由衷欽佩。朱鑫東教授以其獨特的研究視角，拓寬了我的學術視野。無論何時與朱老師探討研究課題，他都會毫無保留地分享見解，激勵我挑戰更高目標。朱老師嚴以律己、寬以待人、溫文爾雅、平易近人的風範和人格，是我工作生活的楷模。若不是兩位導師的悉心點撥與諄諄教誨，我難有學術路上的每一次突破。兩位導師以言傳身教的力量，既授我具體的研究方法，更教會我治學的邏輯與態度，這份“魚漁雙授”的栽培，足以讓我終身受益。千言萬語難表心中敬意與感激，唯有在此奉上最誠摯的謝忱與崇高的敬意。願未來以所學為舟、以所能為槳，不負恩師栽培之德，始終追隨其治學初心，篤行不怠。

一項好的課題研究，離不開優秀的助研。在此由衷感謝我的助研——復旦大學劉聖明教授、博士生鐘容與和香港理工大學劉一燁老師。劉聖明教授以其扎實的理論功底把握研究方向，確定管理學理論，擬定問卷設計思路，牽頭進行數據分析。鐘容與博士以細緻的態度推進具體執行。劉一燁老師協助尋找和梳理國內外文獻資料。和他們一起梳理文獻、分析數據、討論方案的過程，充滿了溫暖與力量。他們不僅有扎實的專業能力、認真的敬業精神，

更願意毫無保留地分享思路、協作互助。這段並肩奮鬥的時光，讓我收穫了專業成長，更收穫了值得一生珍視的朋友與榜樣。

感謝香港城市大學商學院和復旦大學管理學院，給我們創造了再學習的平臺，兩校老師的授課和工作坊指導更是為這段學習時光注入了核心價值，他們不僅教會我們具體的知識內容，更引導我們建立系統的思維方式，這種深層次的啟發，會持續影響我們未來的工作與探索。感謝班主任陳彩霞老師、顧丹老師及周娜琦老師，他們工作認真負責，百問不厭，感謝你們在求學路上溫暖相伴！

感謝 DBA2021 班的每一位同學。這段求學路，我們的足跡遍佈廈門、青島、廣州、揚州、巴黎等城市，移動課堂上的思想碰撞、企業參訪時的實踐感悟、課堂外的歡聲笑語、遇到難題時的互相鼓勵、課後聚餐時的真情流露，每一個瞬間都讓這段學習時光閃閃發光。和你們同行的日子，是這段旅程裏最寶貴的禮物。我始終堅信“離短暫而存長久”，畢業不是終點，即便今後相隔萬裏，但仍情意長存。

最後，要向我的家人表達最深切的感謝。你們的支持與付出，是我完成學業的重要動力。感謝父母，你們不僅用愛哺育我長大，更用包容與信任支持我在求學路上不斷探索。感謝先生，你不僅是生活裏的伴侶，更是我學術上的“同行者”，願意傾聽我的研究困惑，更能給出切實的幫助。感謝女兒，你用自己的方式參與我的論文工作，查閱文獻、處理數據時的認真態度，讓我看到了責任與擔當。有你們作為依靠，我才能更從容地追求目標。

回望這段征途，有伏案鑽研的艱辛，有突破瓶頸的欣喜，更有滿溢於心的感恩。從恩師的引路、同窗的陪伴，到家人的支撐、助研的助力，正是每一份溫暖與力量的彙聚，才讓我得以走完這段路。我深知眼前的成果只是學術長河中的涓滴，未來仍需以謙卑之心繼續求索。但這段旅程中收穫的情誼與成長，將永遠是我最寶貴的財富，也會成為我未來前行的不竭動力。

目錄

摘要.....	i
ABSTRACT.....	iv
QUALIFYING PANEL AND EXAMINATION PANEL.....	viii
致謝.....	ix
表目錄.....	xiv
圖目錄.....	xvi
一、引言.....	1
1.1 研究背景與研究問題.....	1
1.1.1 研究背景.....	1
1.1.2 研究問題.....	10
1.2 研究目的與研究意義.....	11
1.2.1 研究目的.....	11
1.2.2 研究意義.....	14
1.3 研究內容.....	17
1.4 研究方法與技術路線.....	18
1.4.1 研究方法.....	18
1.4.2 技術路線.....	22
1.5 本章小結.....	25
二、文獻綜述與理論背景.....	27
2.1 商業銀行內部審計.....	27
2.1.1 內部審計概念與發展.....	27
2.1.2 商業銀行內部審計的作用.....	28
2.1.3 商業銀行內部審計發展趨勢.....	29
2.1.4 商業銀行內部審計數智化轉型.....	30
2.1.5 數智化轉型下內部審計職能變化.....	32
2.2 智能審計.....	33
2.2.1 智能審計概念.....	33
2.2.2 智能審計獨特性.....	35
2.2.3 智能審計影響.....	36
2.3 理論基礎.....	39
2.3.1 資源保存理論.....	39
2.3.2 心理安全理論.....	42

2.4 本章小結	45
三、研究假設與研究模型	47
3.1 變數定義及研究模型	47
3.2 研究模型	49
3.3 假設推導	53
3.3.1 智能審計系統引入對員工態度的影響	53
3.3.2 智能審計系統引入對員工行為的影響	56
3.3.3 算法透明度的雙重影響機制	60
3.4 本章小結	63
四、研究方法	64
4.1 問卷描述	65
4.2.1 智能審計覆蓋度	66
4.2.2 不安全感	66
4.2.3 防禦傾向	67
4.2.4 主動性行為	68
4.2.5 合規行為	69
4.2.6 算法透明度	70
4.3 分析方法	72
4.3.1 基本分析	72
4.3.2 信度分析	72
4.3.3 效度分析	73
4.4 本章小結	74
五、研究分析	76
5.1 樣本資訊	76
5.1.1 受訪者所在部門	76
5.1.2 受訪者人口統計	77
5.1.3 受訪者工作情況	79
5.2 測量模型分析	81
5.2.1 信度分析	81
5.2.2 效度分析	82
5.2.3 共同方法偏差檢驗	85
六、模型驗證與結果	87
6.1 描述性統計與相關分析	87

6.2 主效應: 智能審計覆蓋度對不安全感與防禦傾向的影響.....	89
6.3 仲介效應: 不安全感與防禦傾向的仲介作用.....	90
6.4 調節作用: 算法透明度的調節作用.....	93
6.5 穩健性檢驗: 分層線性模型 HLM.....	98
七、研究結論與展望	100
7.1 研究結論.....	100
7.2 研究創新點.....	101
7.2.1 理論創新.....	101
7.2.2 實踐創新.....	102
7.3 研究貢獻點.....	103
7.3.1 理論貢獻.....	103
7.3.2 實踐貢獻.....	105
7.4 研究局限性.....	106
7.5 研究展望.....	108
參考文獻.....	111

表目錄

表 2-1 智能審計的應用領域.....	34
表 2-2 四種資源類型.....	40
表 2-3 心理安全的定義.....	43
表 3-1 變數定義.....	48
表 3-2 研究假設匯總.....	63
表 4-1 智能審計覆蓋度測量題項.....	66
表 4-2 不安全感測量題項.....	67
表 4-3 防禦傾向測量題項.....	68
表 4-4 主動性行為測量題項.....	69
表 4-5 合規意願的測量題項.....	70
表 4-6 算法透明度測量題項.....	71
表 5-1 問卷回收概況表.....	76
表 5-2 受訪者所在機構分佈.....	77
表 5-3 受訪者性別分佈.....	77
表 5-4 受訪者年齡段分佈.....	78
表 5-5 受訪者受教育水準分佈.....	78
表 5-6 受訪者從事業務分佈.....	79
表 5-7 受訪者職級分佈.....	80
表 5-8 受訪者工作年限分佈.....	81
表 5-9 信度分析表.....	81
表 5-10 模型 AVE 和 CR 指標結果.....	83
表 5-11 因數載荷係數表格.....	83
表 5-12 Pearson 相關與 AVE 平方根值.....	84
表 5-13 模型擬合指標.....	85
表 6-1 描述性統計.....	88
表 6-2 相關分析.....	88
表 6-3 主效應回歸檢驗.....	90
表 6-4 仲介效應回歸檢驗.....	92

表 6-5 總效應和仲介效應路徑	93
表 6-6 調節效應回歸檢驗	95
表 6-7 有調節的仲介效應檢驗	97
表 6-8 多元層次回歸	98
表 7-1 假設檢驗結果	100

圖目錄

圖 1-1 研究框架圖.....	24
圖 3-1 研究框架.....	52
圖 6-1 算法透明度調節效應分析圖（不安全感）.....	96
圖 6-2 算法透明度調節效應分析圖（防禦傾向）.....	96