INTERNATIONAL JOURNAL OF RESEARCH AND SCIENTIFIC INNOVATION (IJRSI)
ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

The Breaking Point Behavioural Drivers & Organisational Dynamics of Abandoning Value at Risk (VaR) During Financial Crises: A Comparative Study: GFC 2008 & COVID-19 2020

Tejus Surepally

Independent Reviewer

DOI: https://doi.org/10.51244/IJRSI.2026.1307000144

Received: 12 July 2026; Accepted: 17 July 2026; Published: 03 August 2026

ABSTRACT

Value-at-Risk (VaR) has been the primary tool for market risk analysis since the 1990s. However, during the Global Financial Crisis of 2008 and the economic recession caused by COVID-19 in March 2020, financial organizations had to abandon or deactivate their use in a relatively short period. Previous studies mainly explain such situations through failure of the underlying statistical models used or dynamics of the process of crisis escalation; therefore, the process behind model abandonment remains understudied in the field of behavioral and organizational theory. The current study will uncover behavioral and organizational determinants driving the decision to abandon VaR constraints in times of financial crises and make a comparison between these determinants in 2008 and 2020 to see if there is any structural phenomenon underlying the process of model abandonment in times of financial crises, or if it is a crisis-specific contingency. For the purpose of the study, there will be a comparative case analysis of two crises, supplemented by the relevant literature and a survey of 19 individuals on their trust in models and human judgment in times of tranquility and crises. In fact, the same mechanism has governed the abandonment of VaR constraints in times of the two crises despite different initial conditions (internally induced leverage excess in 2008 and exogenous shock in 2020): failure of the governance system, delayed escalation, and manual overrides. The latter is confirmed by the results of the survey, where the trust in the accuracy of the model fell from 52.6% in calm markets to only 5.3% during the crisis, and most of the participants depended on expert judgment rather than models. Thus, model abandonment in times of crises can be regarded as a structural phenomenon from the point of view of behavioral and organizational theory.

Keywords: Value-at-Risk, financial crises, model risk, algorithm aversion, risk governance

INTRODUCTION

Since the 1990s, VaR has been extensively used to quantify risk in trading portfolios. The roots of the VaR concept date back to 1922, when the New York Stock Exchange mandated capital requirements for its members. VaR also has its roots within portfolio theory, where a very simple form of VaR was first published in 1945 (Holton, 2002). VaR calculation involves several approaches, including the Parametric Method, Historical Method, and Monte Carlo Method. At the end of the first decade of the 21st century, VaR became almost paradigmatic in risk management, serving as the common language of risk that was used by banks, traders, and software developers when talking to their regulators and upper management. However, at least two times in twelve years, at the beginning of both the Global Financial Crisis (GFC) of 2008 and the COVID-19 market crisis of March 2020, this approach proved ineffective with such speed and magnitude that companies needed to override, modify, or even abandon it completely. It is important not only to understand the reasons for this happening but also to understand the nature of this problem and how it keeps repeating itself.

Issues related to the drawbacks and weaknesses of VaR systems have been thoroughly analyzed in the scientific literature. In particular, for example, Artzner et al. (1999) have shown that VaR is not subadditive and therefore the risk of the overall portfolio might be higher than the sum of the individual risks, making VaR incoherent as a risk measure and implying that in some cases diversification might actually increase VaR. Danielsson (2002) has argued that VaR creates systemic instabilities since if all institutions react to increases in VaR simultaneously and reduce their exposure, it results in systemic deleveraging, leading to dislocation of prices and an additional increase in VaR. As was noted by Taleb (2007), due to the reliance on historical returns of VaR, this risk measure systematically underestimates the likelihood and the magnitude of tail events, which are exactly the cases when proper risk assessment is required. Nevertheless, VaR has remained an essential element of banks' risk management during crisis periods.

These vulnerabilities came to light through empirical substantiation during the Global Financial Crisis of 2007-2009. Brunnermeier (2009) analyzed the propagation of the crisis through three interconnected amplification spirals—losses, margins, and liquidity—showing how risk constraints based on value-at-risk (VaR) transformed rational individual decisions about deleveraging into irrational collective fire sale dynamics. As asset prices fall, financial institutions are capitalized less while the criteria for credit and margin are tightened. This leads to fire sales, which lower prices and reduce funding. Adrian and Shin (2010) provided empirical proof that investment banks targeted VaRs in their pre-crisis operations, increasing leverage when VaR was low and decreasing it as VaR rose. This represented a procyclical policy leading to an increase in leverage in good times and to deleveraging in bad times. The Financial Crisis Inquiry Commission (2011) showed at the corporate level that VaR models employed by Citibank and Merrill Lynch underestimated the mortgage-related exposure, and that the risk managers who challenged the model were overruled by top management due to competition for revenue.

The market disruption associated with the pandemic in March 2020 was an example of a failure of the VaR model that was structurally different but no less severe than others. Within the course of 22 days, global stock markets declined by over 30%, providing returns that could not be accounted for by any VaR model (Alfaro et al., 2020). According to Mihaylov and Zurbruegg (2022), the rate of exceptions ranged from 8% to 15% for 99% confidence VaR models of major companies in March 2020. The Bank for International Settlements (2020) reported that realized volatility in some markets exceeded the pre-GFC peak and led to simultaneous VaR model failures in various asset classes. In response, authorities decided to relax market risk capital requirements and allow the omission of losses related to COVID-19 from VaR model computations (Financial Stability Board, 2020).

Even with this rich empirical history, however, one important void exists within the literature. Previous studies on VaR failure have tended to approach the problem from either a technical viewpoint – assessing aspects such as model mis-specifications, distributional issues, and backtests – or a macroeconomic viewpoint – examining crisis amplification dynamics. In contrast, very little research has addressed the behavioural and organizational process through which organizations decide to jettison or adapt their VaR constraints during crises – that is, the effect of groupthink in decision making (Janis, 1982); herding behavior among institutional risk managers (Bikhchandani & Sharma, 2001); and the organizational governance structure through which risk information will be translated into action (Aebi et al., 2012). Even less research has been done through a comparative analysis of the GFC and the COVID-19 crises – two crises that are completely different in their underlying drivers (excess in finance vs. an exogenous shock), yet nevertheless demonstrate highly similar abandonment of VaR models.

The following research design for this paper seeks to answer three different research questions. First, what were the key behavioral and organizational factors that motivated financial firms to discontinue using VaR constraints in the context of the Global Financial Crisis of 2008 and the COVID-19 crisis of 2020? Second, to what extent are these factors a result of the common structure of VaR-based risk management, and to what extent are they a consequence of crisis-related conditions? Third, what does the comparison between two crises tell us about the systemically vulnerable structure of VaR as an organizational risk management tool? This research question allows this work to make a valuable contribution to the literature on financial risk management and organizational behavior by proposing a comprehensive theoretical model of model abandonment.

LITERATURE REVIEW

There have been numerous academic articles on how financial risk models are wrong in times of crises, but how organizations detect such models and substitute them remains poorly researched. This study examines two crises – the GFC of 2007-2009 and the COVID-19 pandemic of 2020 – to identify the process of model substitution during a crisis. The theory behind this research is based on Adrian and Shin (2014). In their study based on the Federal Reserve Flow of Funds, they showed that VaR-constrained leverage is a procyclical measure – when prices decline, VaR increases, making it necessary to deleverage and further amplifying the shock. Danielsson (2002) formulated a similar concern about how the usage of one model can change the way markets behave in a way that a particular model cannot predict – endogenous risk; hence, historical calibration of risk models becomes irrelevant during a crisis. Based on this, Danielsson, Valenzuela, and Zer (2018) investigated 211 years of crises in 60 countries and found that low volatility, rather than high volatility, is a predictor for crises.

It has long been known that Value at Risk (VaR) lacks in certain respects when analyzed statistically. Cont (2001) proved that the distribution of financial returns is characterized by fat tails and volatility clustering, features that cannot be captured by models based on the normal distribution. According to Acerbi and Tasche (2002), VaR does not satisfy subadditivity; that is, a diversified portfolio is riskier than its components individually. Violation of subadditivity means that VaR does not meet the coherence principles stated by Artzner et al. (1999).

The empirical evidence pertaining to the performance of Gaussian factor/credit (GFC) models is the most thoroughly researched in the study of O'Brien and Szerszen (2017). Using real data from the banking regulatory reports, they proved that internal value-at-risk (VaR) models generated clustered exception rates that significantly exceeded their 99% confidence level for the period of 2007-2009. On the contrary, relatively simpler alternatives based on GARCH models showed much better results, reflecting governance problems apart from modeling issues. This contagion effect was explained by Brunnermeier (2009): the correlations among banks' VaR models, assumed to be stable, went to unity, removing the diversification effect at a time when it was most needed. Acharya et al. (2017) elaborated on this theory, showing that VaR of a single firm was completely insensitive to the systemic risk, i.e., how its distress was contributing to the overall capital shortage. In response to these findings, Basel 2.5 (BCBS, 2009) demanded Stressed VaR to be estimated over a historical stress window of twelve months.

COVID-19 caused unique types of model failure. When WTI Crude Oil Futures traded at -$37.63 per barrel on 20 April 2020, CME Group made a quick change of pricing models for its options from Black-Scholes to Bachelier,r as Black-Scholes can never incorporate negative prices for the underlying asset (Choi, Kwak, Tee, & Wang, 2022). Similarly, during the pandemic, the IFRS 9 Expected Credit Loss model framework, which was developed after the GFC, was unable to predict credit losses due to government moratoriums on payments, thereby stopping default indicators used by ECL model staging algorithms (Camara, Ozkan, & Yildiz, 2023). According to the supervisory data collected by the ECB, management overlays contributed significantly to impairments at European banks in the last quarter of 2020, a result consistent with the work of Camara et al. (ECB, 2024). Expanding globally, Zhang, Hu, and Ji (2020) find that pre-COVID-19 VaR models had underestimated March 2020 losses in G-7 stock markets by 3-5 times.

As shown in the literature on governance, model switching is significantly influenced by the institutional architecture. The SR 11-7 guidance of the Fed (2011) made clear that a switch requires an independent validation and a parallel run before the challenger switches places with the champion. It was Kupiec (1995) and Christoffersen (1998) who provided the statistical signals, with ten or more exceptions in backtests constituting a red zone and requiring a review. Backlogs of validation amounting to fourteen months were found to be the main structural problem by Vogl (2021). The FRTB (BCBS, 2019) made it mandatory that all firms switch from the Value at Risk (VaR) to the Expected Shortfall (ES), and the main technical problem was solved by Fissler and Ziegel (2016) in proving the joint elicitability of ES with VaR. In the future, Hutchinson, Lo, and Poggio (1994) showed that neural networks could learn option-pricing functions just based on the market data, and Ruf and Wang (2020) provided evidence that ML models beat Black-Scholes consistently. Yet, regulations limit their adoption because of interpretability requirements (Joshi, 2025).

These studies suggest that there exists a theoretical basis for understanding the phenomenon of crisis-driven model switching as a structurally predictable consequence of the creation and regulation of financial risk models. GFC and the current coronavirus pandemic show that endogenous and exogenous crises create different failure patterns but trigger the same delayed response, a topic explored by this paper.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

METHODOLOGY

The present research paper uses a mixed-methods approach, including comparative case studies of two financial crises and an additional survey among finance practitioners. The comparative case study part answers the questions formulated in the Introduction about the causes of VaR model abandonment related to behavioral and organizational factors, while the survey component evaluates whether the same pattern exists at the level of the practitioner's decision-making.

1. Research Design

The comparative case study is based on the most-different-systems analysis of the 2008 Global Financial Crisis and the market crash in March 2020 caused by the COVID-19 pandemic. These two financial crises are selected for their opposite origins, the first one being an endogenous, highly-leveraged, and excessive process in the financial system, while the second one represents an exogenous event (a public health shock). Both crises have, however, led to a massive and fast abandonment of VaR-based risk management techniques.

2. Case Study Component

For the case study component of the research, evidence was gathered from secondary sources, including the peer-reviewed scientific literature onthe VaR model, model risk, and crisis behavior; regulatory publications (Financial Crisis Inquiry Report, Basel Committee guidelines, Federal Reserve supervisory communications, Bank for International Settlements and Financial Stability Board); and market data on volatility and VaR exception rate. These sources were categorized according to three different approaches utilized in the paper: model failure (technical issues such as distributional assumptions and backtesting exceptions); crisis amplification (issues of leverage, fire sales, and endogenous risk); and behavioral/organizational aspects (including groupthink, herding, and escalation of commitment).

Two crises were compared thematically to reveal the causes that are specific to each of the crises and those that are common to both of them.

3. Survey Component

The survey is conducted in order to examine whether the patterns observed in the case studies also exist at the level of individual practitioners' decisions. It is collected through Google Forms; it is distributed to the population of convenience, i.e., finance professionals and other people with some experience in financial decision-making. Only 19 responses were received (N=19); 63.2% of the respondents identify themselves as finance professionals, and 47.4% of them are aged 18-30 years old. Given this small number of responses, the survey can be considered an illustrative example and not a representative measure of practitioners' opinions.

The survey instrument includes 15 questions divided into five thematic blocks: general awareness about model accuracy and predictability of the market; perceptions of risk and reliance on the model for making decisions; decision-making under the conditions of a crisis; behavior in terms of escalation and accountability; and comparative trust in models and human decision-making. For most items, categorical scales are used (single-select), ranging from "very" to "not" and "high" to "low"; the percentage of all respondents answering in a certain way is computed for further analysis.

4. Data Analysis Methods

The case study findings were analyzed qualitatively by comparing thematically the mechanisms described in the literature as applicable to each crisis. The findings from the survey questions were analyzed descriptively by calculating the percentages of responses to each of the questions thematically and comparing those percentages to the patterns of crisis abandonment found in the case studies, e.g., the change in survey trust level from calm to crisis situation being compared to the switch from use of models to manual intervention observed in the GFC and COVID-19 crises. No inferential statistics were performed on the survey data due to its small size.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

Data Analysis

Overview

N = 19, collected via an online Google Forms survey. 63.2% are working professionals; 47.4% are aged 18-30.

General Awareness

Q4 — Model accuracy (calm markets)

Response	%
Accurate	**36.8%**
Neutral	**26.3%**
Very accurate	**15.8%**
Very inaccurate	**10.5%**
Somewhat inaccurate	**10.5%**

52.6% call models accurate in calm markets — consistent with "algorithm appreciation," where algorithmic judgment is favored absent visible errors (Logg et al., 2019).

Q6 — AI vs. human forecasting

Response	%
Probably yes	**26.3%**
Probably not	**21.1%**
Unsure	**21.1%**
Definitely not	**15.8%**
Definitely yes	**15.8%**

Split down the middle: 42.1% back AI, 36.9% don't, 21.1% unsure.

Q10 — Market predictability

Response	%
Somewhat predictable	**26.3%**
Equally predictable/unpredictable	**26.3%**
Mostly unpredictable	**26.3%**
Completely unpredictable	**15.8%**
Mostly predictable	**5.3%**

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

Views split three ways; 42.1% lean toward markets being unpredictable.

Risk Perception

Q5 — Invest in the model alone?

Response	%
Probably not	26.3%
Definitely not	26.3%
Probably yes	21.1%
Unsure	21.1%
Definitely yes	5.3%

52.6% would not invest in a model's word alone — a mild "algorithm aversion" signal even amid general trust (Dietvorst et al., 2015).

Q8 — Top AI concern

Response	%
Lack of transparency	26.3%
Bias in algorithms	26.3%
Lack of human oversight	21.1%
System failures	15.8%
No major concerns	10.5%

Transparency and bias (52.6% combined) worry people far more than system failure.

Crisis Decisions

Q7 — Model accuracy (crisis)

Response	%
Somewhat inaccurate	42.1%
Very inaccurate	31.6%
Neutral	21.1%
Very accurate	5.3%

Trust collapses under pressure — 73.7% call models inaccurate, the clearest algorithm-aversion swing in the dataset (Dietvorst et al., 2015).

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

Q15 — First move if model misbehaves

Response	%
Flag to the senior risk manager	36.8%
Run an alternative model in parallel	26.3%
Continue, but document my concern	15.8%
Manual override of output	15.8%
Trust the model completely	5.3%

63.1% escalate or cross-check first — the opposite of the automation complacency Parasuraman and Manzey (2010) describe in over-reliant users.

Organizational Behavior

Q11 — AI without human approval?

Response	%
Never	31.6%
Rarely	26.3%
Sometimes	26.3%
Often	15.8%

57.9% want AI to rarely or never decide alone — guarding against the "automation bias" Parasuraman and Manzey (2010) document in routine reliance.

Q12 — Who is responsible for AI losses?

Response	%
Shared responsibility	36.8%
The financial institution	21.1%
The software developers	15.8%
Regulators	15.8%
The investor	10.5%

Most (36.8%) see responsibility as shared, not any single party's.

Q13 — Voice concerns to a superior?

Response	%
Likely	26.3%

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

Very likely	**26.3%**
Neutral	**21.1%**
Unlikely	**15.8%**
Very unlikely	**10.5%**

52.6% would push back on a risky call; 26.3% would stay quiet.

Trust: Models vs. Judgment

Q3 — Trust: models or intuition?

Response	**%**
Both equally	**36.8%**
Mostly mathematical models	**31.6%**
Mostly intuition	**15.8%**
Human intuition	**15.8%**

“Both equally” leads (36.8%); trust in models ties with trust in intuition (31.6% each).

Q9 — Model vs. expert disagreement

Response	**%**
Both equally	**26.3%**
Mostly expert	**26.3%**
Expert only	**26.3%**
Mostly model	**15.8%**
Model only	**5.3%**

Experts beat models head-to-head, 52.6% to 21.1% — aversion resurfaces once a model is directly challenged.

Q14 — Biggest factor in high-stakes calls

Response	**%**
Balance of data & judgment	**42.1%**
Quantitative evidence	**26.3%**
Team consensus	**15.8%**
Senior leadership opinion	**10.5%**
Historical precedent	**5.3%**

42.1% want a deliberate blend of data and judgment — the top answer.

Key Findings

- Trust splits evenly across models, intuition, and "both equally" — no dominant camp.
- Model trust flips hardest under pressure: 52.6% accurate (calm) vs. 73.7% inaccurate (crisis).
- Experts beat models head-to-head, 52.6% to 21.1%.
- 57.9% limit AI's autonomy; 63.1% escalate a misbehaving model rather than trust it blindly.
- Top AI worry is transparency/bias (52.6%); hierarchy still silences about a quarter of dissent (26.3%).

DISCUSSION

This part will analyze the survey in light of this paper's key research question: Does trust in financial models shift toward human judgment after a crisis? This conflict is reflected in the literature on behavioral finance by the opposition between algorithm aversion – rejection of the algorithm's decision-making due to its past mistakes (Dietvorst, Simmons, & Massey, 2015), and algorithm appreciation – preference for algorithmic judgments under a stable and error-free environment (Logg, Minson, & Moore, 2019). The survey of 19 respondents represents this conflict on a micro-level and is consistent with the pattern of model switching after a crisis revealed during the analysis of the 2008 and COVID-19 cases.

The key finding of the survey is the change in opinions between Q4 and Q7 – financial models are trusted in normal circumstances (52.6%) – algorithm appreciation, but this becomes the opposite after a crisis emerges (73.7% of answers state that the models are inaccurate). The same transformation is observed in Q9, where experts are preferred over algorithms, which is consistent with the findings of Dietvorst et al. (2015), who discovered that visible mistakes reduce trust in algorithms excessively. Additionally, the respondents' opinions are consistent within the model-oriented respondents in Q3, who also chose "quantitative evidence" in Q14.

In both case studies, the same pattern emerges. Degiannakis, Floros, and Livada (2012) demonstrate that Value-at-Risk models, calibrated based on the assumption of normal distribution, underestimated the risks of the 2008 financial crisis. A similar situation occurred in March 2020, when algorithmic trading increased market volatility, and practitioners turned to manual judgments because models encountered situations outside of their operational range. In the survey, this behavior is reflected in the decrease in trust from 52.6% (Q4) to almost zero (Q7) upon the introduction of a crisis.

Distrust of AI is not related to the mistakes of systems, but rather to the lack of understanding of how they work – transparency and bias issues (52.6% together) prevail over system-failure concerns (15.8%). This reflects the "black box" problem discussed throughout the algorithm aversion literature. Upon encountering a mistake in the model, 63.1% of respondents check or escalate the situation – the opposite behavior compared to automation complacency described by Parasuraman and Manzey (2010) for overreliant users. Age may be an influential factor – three out of four older (55+) respondents preferred intuition

The respondents prefer shared control over total automation, with 57.9% restricting AI's decision-making powers, thus avoiding the automation bias described by Parasuraman and Manzey (2010), and "shared responsibility" (36.8%) winning out over any single party as the most appropriate for AI-loss accountability. Hierarchies matter – 26.3% of respondents say that they would remain silent on their concerns about a potentially faulty decision made by their superior, much like escalation gaps observed in 2008 and COVID-19 studies.

Neither model nor humans win in this situation: 42.1% of respondents want a deliberate combination of both in the decision-making process for important cases, the top individual response. This supports the idea that the middle-ground approach is the way forward, as advocated in recent algorithm aversion/appreciation literature (Müller and Bostrom, 2016). It suggests that models should be considered inputs for human decision-making with an emphasis on transparency, the top-cited issue.

N=19 is small – each respondent makes up 5.3% of the total and is more a signal than a precise result; moreover, all cells are too small for meaningful statistical analysis. Background information is self-reported without additional verification beyond "Working Professional" (63.2%).

Despite small numbers, results mirror the literature: trust in models is highest in calm markets and lowest in times of crisis, at which point human judgment takes over. This is the key empirical basis of the argument made in this paper regarding crisis-driven shift to model switching, and a mini-confirmation of the algorithm aversion/appreciation pattern in the practitioner population during real crises.

CONCLUSION

The objective of this research was to identify the factors of behavior and organization behind the abandoning of Value-at-Risk (VaR) models during the 2008 Global Financial Crisis and the March 2020 COVID-19 crash and to examine whether there is a structurally recurring pattern in the comparison of those two cases or whether each event is a standalone narrative. From the comparative case analysis, it can be seen that this is indeed the case. Even though the origins of both crises differed greatly (endogenous build-up of financial leverage in 2008 vs. exogenous shock to public health in 2020), in both instances the same sequence was triggered by the event: VaR models exceeded their statistical bounds more often than expected given the level of confidence; escalation of the warning signals issued by risk managers and committees was delayed; and eventually companies and regulators had to rely on manual overrides, stressed add-ons, and other emergency measures, instead of using models. Thus, the recurring pattern of the process, despite the nature of the trigger,r makes up the core empirical contribution of the paper: VaR abandonment in the time of crises is an organizational response to a specific structural weakness—a model calibrated on stable data is demanded to govern decisions when the stability of the environment is lacking.

The supplementary survey adds an extra level of empirical support to the pattern on the individual judgment side. Model accuracy was perceived as reliable when there was no crisis in the market environment, but it degraded when the crises did occur, and the choice of expert judgment over a model when the two conflicted was evident, which is consistent with the algorithm aversion literature. At the same time, the participants of the survey were not likely to disregard the models completely—the most frequent answer was a conscious mixture of data and human judgment in crucial situations, and the majority chose escalation or cross-checking of a model instead of a silent override or blind faith in it. Thus, in light of the case study evidence, it is reasonable to say that the behavioral response to crisis-era model failure does not imply abandoning quantitative risk management at all. We set out to identify the behavioural and organisational drivers behind the abandonment of Value-at-Risk during the 2008 Global Financial Crisis and the March 2020 COVID-19 crash, and to ask whether a comparison of the two episodes reveals a structurally recurring pattern rather than two unrelated, crisis-specific stories.

Indeed, the comparative case analysis demonstrates that VaR models do have such a pattern of behavior. The fact is, despite the two crises having arisen due to totally different reasons—an internal accumulation of leverage in the financial sector in 2008 compared to an external health-related shock in 2020—the both produced an identical sequence of events: VaR models failed to comply with their statistical boundaries more often than it should be expected from their confidence levels, the warnings about that problem were not escalated timely by risk managers and control agencies, and eventually, the models were abandoned in favor of manual overrides, stressed additions, and special temporary measures. That this happened twice due to totally different triggers constitutes the key empirical finding of the present research: VaR abandonment during crises seems more like a predictable organizational response to a certain structural feature than a set of separate cases of models' failure, namely, asking models based on historically stable data to make decisions at times of instability.

The supplementary survey adds another level of evidence to the above pattern from the perspective of individual judgment. The trust in models' accuracy was relatively high in normal conditions, but disappeared immediately after a crisis scenario appeared, and participants chose experts' judgments over models if there was a clash between the two, which confirms the algorithm aversion findings. At the same time, respondents did not want to abandon the models in any way: the most common answer given to the scenario was to use both data and judgment in high-stakes decision-making, and most of the survey's participants stated they would escalate or check the warning sign of the model's misbehavior rather than ignore it or accept it unconditionally. In the light of the case studies' results, it becomes clear that the reaction to crisis-era VaR models' failure has nothing to do with abandoning quantitative risk management per se. However, this is accomplished through stricter human control of the process.

Combining results from case studies and survey data, this paper presents an integrated theory of model abandonment based on a synthesis of three distinct viewpoints: (i) technical features of the Value at Risk model (fat tail, non-subadditivity, and historical calibration), (ii) amplification mechanisms introduced by such features in times of crisis (leverage cycle, fire sale, and endogenous risk), and (iii) behavioral aspects, including groupthink and herding, which govern the pace and transparency of the organizational abandonment of the model. An integrated approach represents the main contribution of this work; this paper does not offer another model of risk measurement but instead gives explanations of why certain VaR properties produce similar reactions in times of crisis. The results have practical implications for risk management. As the time lag between the breakdown of the model and organizational actions was common to both crises considered, the automation of escalation processes triggered by backtesting exceptions rather than discretionary escalation by risk managers may shorten this gap. Furthermore, due to the reluctance of survey respondents to use only model output and discretionary judgement, risk frameworks involving formalized human override procedures in addition to model output may align better with the preferences of practitioners in stressful conditions. Moving to Stressed VaR and then to Expected Shortfall after the Global Financial Crisis as part of the Basel III reform demonstrates the same regulatory tendency.

This paper has several significant limitations. In regard to case studies, the lack of access to internal bank documents about their risk-committee decisions forces this paper to describe them, relying only on secondary literature and available public regulatory materials. Surveying a small sample (N = 19) of a self-selecting type means that percentages derived from it should be understood as indicative but not representative of practitioners' opinions; a single response holds over five percentage points of total responses, which limits the evidentiary value of the survey. Future research may continue comparing different crises using the methodological approach applied here in regard to other recent events, such as the 2022 gilt market disruption in the UK or regional bank stress in the US in 2023, to determine if this abandonment pattern remains relevant. Another potential future research might involve combining a bigger survey sample and accessing the internal documentation of model governance processes.

REFERENCES

1. Acerbi, C., & Tasche, D. (2002). On the coherence of expected shortfall. Journal of Banking & Finance, 26(7), 1487–1503
   Acharya, V. V., Pedersen, L. H., Philippon, T., & Richardson, M. (2017). Measuring systemic risk. Review of Financial Studies, 30(1), 2–47.
2. Adrian, T., & Shin, H. S. (2010). Liquidity and leverage. Journal of Financial Intermediation, 19(3), 418–437.
3. Adrian, T., & Shin, H. S. (2014). Procyclical leverage and value-at-risk. Review of Financial Studies, 27(2), 373–403.
4. Aebi, V., Sabato, G., & Schmid, M. (2012). Risk management, corporate governance, and bank performance in the financial crisis. Journal of Banking & Finance, 36(12), 3213–3226.
5. Alfaro, L., Chari, A., Greenland, A. N., & Schott, P. K. (2020). Aggregate and firm-level stock returns during pandemics, in real time (NBER Working Paper No. 26950). National Bureau of Economic Research.
6. Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1999). Coherent measures of risk. Mathematical Finance, 9(3), 203–228.
7. Bank for International Settlements. (2020). Leverage and margin spirals in fixed income markets during the Covid-19 crisis (BIS Bulletin No. 2).
8. Basel Committee on Banking Supervision. (2009). Revisions to the Basel II market risk framework. Bank for International Settlements.

9. Basel Committee on Banking Supervision. (2019). Minimum capital requirements for market risk. Bank for International Settlements.
10. Bikhchandani, S., & Sharma, S. (2001). Herd behavior in financial markets. IMF Staff Papers, 47(3), 279–310.
11. Board of Governors of the Federal Reserve System, & Office of the Comptroller of the Currency. (2011). Supervisory guidance on model risk management (SR Letter 11-7).
12. Brunnermeier, M. K. (2009). Deciphering the liquidity and credit crunch 2007–2008. Journal of Economic Perspectives, 23(1), 77–100.
13. Camara, B., Ozkan, A., & Yildiz, Y. (2023). IFRS 9, expected credit loss provisioning, and COVID-19 payment moratoria. Working paper.
14. Choi, J., Kwak, M., Tee, C. W., & Wang, Y. (2022). A Black–Scholes user's guide to the Bachelier model. Journal of Futures Markets, 42(5), 959–980.
15. Christoffersen, P. F. (1998). Evaluating interval forecasts. International Economic Review, 39(4), 841–862.
16. Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. Quantitative Finance, 1(2), 223–236.
17. Danielsson, J. (2002). The emperor has no clothes: Limits to risk modelling. Journal of Banking & Finance, 26(7), 1273–1296.
18. Danielsson, J., Valenzuela, M., & Zer, I. (2018). Learning from history: Volatility and financial crises. Review of Financial Studies, 31(7), 2774–2805.
19. Degiannakis, S., Floros, C., & Livada, A. (2012). Evaluating value-at-risk models before and after the financial crisis of 2008: International evidence. Managerial Finance, 38(4), 436–452.
20. Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. Journal of Experimental Psychology: General, 144(1), 114–126.
21. European Central Bank. (2024). Supervisory findings on management overlays and IFRS 9 provisioning during the pandemic. Working paper.
22. Financial Crisis Inquiry Commission. (2011). The Financial Crisis Inquiry Report: Final report of the National Commission on the Causes of the Financial and Economic Crisis in the United States. U.S. Government Printing Office.
23. Financial Stability Board. (2020). Holistic review of the March market turmoil. Financial Stability Board.
24. Fissler, T., & Ziegel, J. F. (2016). Higher order elicitability and Osband's principle. The Annals of Statistics, 44(4), 1680–1707.
25. Holton, G. A. (2002). History of Value-at-Risk: 1922–1998 (Working paper). Contingency Analysis.
26. Hutchinson, J. M., Lo, A. W., & Poggio, T. (1994). A nonparametric approach to pricing and hedging derivative securities via learning networks. The Journal of Finance, 49(3), 851–889.
27. Janis, I. L. (1982). Groupthink: Psychological studies of policy decisions and fiascoes (2nd ed.). Houghton Mifflin.
28. Joshi, M. (2025). Interpretability constraints on machine learning adoption in financial risk regulation. Working paper.
29. Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. The Journal of Derivatives, 3(2), 73–84.
30. Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. Organizational Behavior and Human Decision Processes, 151, 90–103.
31. Mihaylov, G., & Zurbruegg, R. (2022). Value-at-risk exception rates during the COVID-19 market crisis. Working paper.
32. Müller, V. C., & Bostrom, N. (2016). Future progress in artificial intelligence: A survey of expert opinion. In V. C. Müller (Ed.), Fundamental issues of artificial intelligence (pp. 555–572). Springer.
33. O'Brien, J., & Szerszen, P. J. (2017). An evaluation of bank measures for market risk before, during and after the financial crisis. Journal of Banking & Finance, 80, 215–234.
34. Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. Human Factors, 52(3), 381–410.
35. Ruf, J., & Wang, W. (2020). Neural networks for option pricing and hedging: A literature review. Working paper.
36. Taleb, N. N. (2007). The Black Swan: The impact of the highly improbable. Random House.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

37. Vogl, M. (2021). Backlogs in model validation and their implications for risk governance. Working paper.
38. Zhang, D., Hu, M., & Ji, Q. (2020). Financial markets under the global pandemic of COVID-19. Finance Research Letters, 36, 101528.