The Political Economy of AI-Driven Financial Supervision

Gérard Hertig
ETH Zurich and ECGI

© Gérard Hertig 2022. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

This paper can be downloaded without charge from:
http://ssrn.com/abstract_id=4012048

https://ecgi.global/content/working-papers
The Political Economy of AI-Driven Financial Supervision

Working Paper N° 621/2022
January 2022

Gérard Hertig
Abstract

AI-reliance is expected to improve risk management across the financial services industry, reinforcing the dominance of private ordering in 'normal times'.

From a supervisory perspective, the use of AI can be expected to decrease regulatory enforcement costs while providing technology-advanced players with opportunities to game the regulatory system.

More fundamentally, AI-reliance is unlikely to either significantly improve the prompt and effective handling of systemic incidents or to increase systemic risk. However, the use of AI may go hand-in-hand with significant job losses.

Overall, the use of AI can be expected to have an impact on the respective roles of private ordering and state regulation. The former will become (rapidly) dominant in normal times while the latter will (slowly but increasingly) target systemic issues.

Keywords: Artificial intelligence, compliance, courts, financial supervision, investor protection, systemic risk

JEL Classifications: K22, K23, L22, L23
The Political Economy of AI-Driven Financial Supervision
Table of Content

1 Abstract 3
2 The Increasing Role of AI 4
3 Private Ordering vs State Regulation 7
4 The Politics of AI-driven Supervision 10
  4.1 Increasing Supervisory Independence 10
  4.2 “Private” Interests as a Key Driver 11
  4.3 Macro-Economic Impact 13
5 A Long Way to Go for AI-Driven Supervision 17
1 Abstract

AI-reliance is expected to improve risk management across the financial services industry, reinforcing the dominant of private ordering in ‘normal times’.

From a supervisory perspective, the use of AI can be expected to decrease regulatory enforcement costs while providing technology-advanced players with opportunities to game the regulatory system.

More fundamentally, AI-reliance is unlikely to either significantly improve the prompt and effective handling of systemic incidents or to increase systemic risk. However, the use of AI may go hand-in-hand with significant job losses.

Overall, the use of AI can be expected to have an impact on the respective roles of private ordering and state regulation. The former will become (rapidly) dominant in normal times while the latter will (slowly but increasingly) target systemic issues.

Key Words: Artificial intelligence, compliance, courts, financial supervision, investor protection, systemic risk.
The authorities in charge of financial supervision still rely upon human judgement for oversight purposes. At the same time, they are starting to trust artificial intelligence (AI). There is an awareness of AI-driven analysis being risky due to data quality, transparency and understandability issues; however, AI advances provide supervisory authorities with an opportunity to focus upon domains where humans (still) have an advantage over machines.

1. These developments generate social benefits and costs. On the up side, AI-use is expected to increase labor productivity by up to 40%—in particular, by taking over tedious tasks that contribute to friction within large organizations like government agencies. It follows that lawmaking and enforcement costs should decline, especially when it comes to predicting bank distress, detecting fraud and minimizing money laundering.

On the down side, financial authorities are likely to face AI models they cannot fully comprehend. More importantly, large financial intermediaries may benefit from AI-driven supervision at the expense of their smaller brethren.

2. Given this environment, one does not expect financial authorities to adopt a wait-and-see approach. To begin with, it would open them to ‘obsolete technology’ blame should scandals or market failures occur. In addition, AI-driven supervision is likely to prove popular with taxpayers due to its (perceived) cost-savings potential. Finally and most importantly, financial authorities cannot afford to lag behind AI-reliant private players if they want to remain effective and credible.
This does not mean that financial authorities are wholeheartedly embracing AI-driven supervision; they are aware of AI-reliance raising complex issues from an ethical, legal, political and economic perspective.viii Hence, the European Commission (EC) is emphasizing the need for a common AI approach to avoid single market fragmentation and foster legal certainty,ix whereas the European Central Bank (ECB) follows a cautious approach to using AI for supervisory purposes.x Similarly, the US government sees its AI regulation role as a limited one,xi the idea being that uncertainty about AI-related risks is likely to decrease over time.xii

3. AI-reliance also has behavioral implications. Human beliefs generally depend on recalled personal experiences such as stock market crashes, and not merely on statistical information. In addition, there is a mystifying diversity of human biases; in particular, people tend to over-estimate the frequency of events that are unlikelyxiii or can be broken down in constituent parts.xiv By contrast, the use of AI can be expected to result in more complete and unbiased decision-making. Programming or data bank deficiencies may initially offset some of the benefits of reduced human intervention, but they should vanish over time. The downside is that AI decision-making may run counter to the (subjective) preferences of voters; here, the hope is that their inclinations can be addressed via legislative amendments and the resulting coding adjustments.

4. Overall, the benefits of AI-driven financial supervision can be considered highest for financial centers. In line with this hypothesis, Singapore and Swiss financial authorities are among the most advanced AI users. French and German supervisory authorities are not far behind, whereas


xii Sarah Lichtenstein, Paul Slovic, Baruch Fischhoff and Mark Layman, Judged Frequency of Lethal Events, 4 Journal of Experimental Psychology Human Learning and Memory 551 (1978).
Japan significantly relies on AI to detect market manipulation and fraudulent money transfers.\textsuperscript{xv}

On the other hand, supervisory use of AI remains \textit{unimpressive} in most other European jurisdictions and the US. This deficit is a source of concern, especially given China’s attempts to export its state control approach to data governance and AI use.

5. These concerns are addressed by the 2020 European Commission (EC) \textbf{White Paper} on Artificial Intelligence.\textsuperscript{xvi} The basic aim is to propose a regulatory framework that inspires confidence, given that many market actors consider AI as untrustworthy in view of the privacy, data collection and intended use issues it raises.

The EC identified seven (partly overlapping) requirements that should contribute to ‘\textit{trusted}’ AI: 1) ultimate human control; 2) technical robustness and safety; 3) privacy and data governance; 4) transparency; 5) diversity, non-discrimination and fairness; 6) societal and environmental wellbeing; and 7) accountability.

In addition, the EC submitted a draft \textbf{Artificial Intelligence Act} (AIA) to the European Parliament.\textsuperscript{xvii} The aim is to maintain EU leadership while preserving EU values and fundamental rights. In particular, AI systems posing significant health, safety or fundamental rights risks will have to comply with a set of horizontal requirements. These systems will also be subject to conformity assessment procedures.\textsuperscript{xviii}

6. The draft AIA is viewed as a \textbf{watershed}. It also provides a good opportunity for the EU to engage a regulatory dialogue with the US, which has displayed restraint in regulating AI.\textsuperscript{xix} For example, the draft Algorithmic Accountability Act,\textsuperscript{xx} which aimed at striking a balance between promoting AI and safe, responsible and democratic development, died in Congress.

\textsuperscript{xv} See \url{https://www.jpx.co.jp/english/corporate/news/news-releases/0060/20180319-01.html}.

\textsuperscript{xvi} COM(2020) 65 final.

\textsuperscript{xvii} COM/2021/206 final.


\textsuperscript{xix} Meredith Broadbent, \textit{AI Regulation: Europe’s Latest Proposal is a Wake-Up Call for the United States} (May 18, 2021), available at \url{https://www.csis.org/analysis/ai-regulation-europes-latest-proposal-wake-call-united-states}.

\textsuperscript{xx} Bill H.R. 2231, introduced on October 4, 2019.
without receiving a vote. However, change may be in the making, as evidenced by the more aggressive approach pursued by the draft Algorithmic Justice and Online Platform Transparency Act.\textsuperscript{xxi}

The differences in EU and US approaches may be due to US players favoring voluntary standards. For example, the NASDAQ stock exchange has used AI to detect irregular activities since 2017,\textsuperscript{xxii} and for market surveillance since 2019,\textsuperscript{xxiii} nowadays, its systems review more than 750,000 alerts (unusual price movements, trading errors and potential manipulation) a year.\textsuperscript{xxiv}

3 Private Ordering vs State Regulation

In this evolving environment, one fundamental policy question arises: will the use of AI have an impact on the respective importance of private ordering (via personal and collective norms) and state regulation? The easy answer is 'definitely'; it is harder to predict in which circumstances and to what extent.

1. Overall, private ordering will remain dominant in ‘normal times’. Many executives expect the advent of AI to improve financial institutions’ risk management.\textsuperscript{xxv} There is a simple assumption behind this anticipation: AI being easier to buy than (equivalent) talent, its use has to increase the robustness of financial intermediaries.

It follows that financial authorities are likely to spend fewer resources monitoring market conduct. Instead of investigating compliance with the principle of precaution, they will devote significant attention to the overall robustness of AI-driven systems.

In other words, structural review will pre-empt behavioral analysis.

2. In this context, dealing with systemic risk takes a new meaning. The generalized use of AI should facilitate the private uncovering and management of systemic incidents. In theory, market participants will identify

\textsuperscript{xxi} Bill S. 1996, introduced on May 27, 2021.
\textsuperscript{xxiii} See https://www.finextra.com/pressarticle/80533/nasdaq-uses-ai-for-surveillance-patterns.
\textsuperscript{xxv} See https://www.sqlpower.ca/artificial-intelligence-financial-regulation.
and execute the required measures (margin calls, stop loss, firewalls, etc.) in real time. In practice, however, circuit breakers could hamper their timely execution or, as evidenced by past crises, even shut down transaction systems altogether.

More fundamentally, one must be aware that most of the available data is from normal times. It is not easy to link this data to the determinants of large losses that could threaten overall stability. In fact, history shows that financial crises generally get market participants flatfooted.

Overall, technological improvements is likely to make financial systems more resilient. In particular, one can expect AI-driven supervision to complement the private use of AI, with significant contributions to systemic risk management.

3. In terms of state intervention, this by-and-large positive impact of AI may not prove as decisive as its impact on workforce size and composition.

There is evidence of financial services firms increasingly using AI for risk management (56%) and new products and processes generation (52%) purposes. This evolution will generate significant efficiency gains; hence, US banks expect AI-use to generate a $70 billion reduction in middle-office costs by 2025. However, AI-use will also go hand-in-hand with workforce reductions or reallocations; for example, AI-use is projected to cut or transform 23% of banking jobs in China by 2027, while increasing the output of the remaining 77%.

Given these forecasts, policy-makers as well as industry representative can be expected to adopt measures to minimize their social impact. An obvious strategy is to rely on automation primarily for routine tasks while beefing-up jobs that require human judgment and expertise. For example, the UK Serious Fraud Office uses AI to review privileged document, thus

---

xxx David He and Venessa Guo, Boston Consulting Group, 4 Ways AI will Impact the Financial Job Market (September 14, 2018), available at www.weforum.org.
reducing independent counsel review by 80%. Similarly, augmented demand for protecting the integrity of data collection, processing and storing is likely to have positive workforce effects. A good example of the drivers of data integrity concerns is the US Governmental Accounting Office (GAO) pointing out that cybersecurity incidents have increased by more than 1,000% from 2006 to 2015.

4. AI is also changing the way financial institutions interact with financial supervisors.

For example, financial supervisors have started using AI to identify misbehaving financial advisers and suspicious trading activity. Applications range from mis-selling detection in the mortgage loan and consumer credit contracts area to the generalized detection of financial irregularities.

This new approach is contingent upon the parties trusting each other. This is especially true for operational risks, given the autonomous adaptability of AI-based models and the complexity of the techniques employed.

5. More specifically, the increasing use of AI raises various practical issues. One is whether AI-driven supervision is biased in that it focuses on those areas where data is available or risks are easier to measure. Another issue is that AI performance is less impressive when the past is unlike the future, i.e. when standards are superior to rules. A third issue is that the use of AI generates data authenticity, provenance, ownership and privacy issues, as well as data and concept drift problems.

---


In the US, the Securities and Exchange Commission (SEC) is dealing with such issues by using unsupervised algorithms to detect data patterns and anomalies, and supervised learning algorithms to inject SEC knowledge into the process. These successive algorithms can then be applied to new data, for example those generated by recent SEC filings. However, in a possibly self-interested move, SEC representatives have emphasized that it would be premature to think of AI as the next market regulator— a view shared by the Financial Stability Board (FSB), which has highlighted the need to keep a human touch within the supervisory process.

Still another issue is that financial intermediaries have been slow in translating regulatory expectations into effective internal governance and practices. This has led financial supervisors to clarify their position. For example, the SEC has ruled that AI-use must be disclosed to investors; it is also ramping-up its enforcement actions, as evidenced by Charles Schwab having to set aside $200 million following an investigation about fraudulent charges for robot advising.

4 The Politics of AI-driven Supervision

The move from human to AI-driven financial supervision will result in gains and losses for financial supervisors as well as for financial intermediaries and investors.

4.1 Increasing Supervisory Independence

This new technology world is likely to make financial supervisors more independent from lawmakers.

1. To begin with, decision-making relies on codes, which are harder to assess than human-generated decisions. This barrier to understanding is
compounded by the use of AI allowing for more differentiated outcomes, due to the taking into account of a larger number of variables.

At the same time, managerial failures could be harder to prove in an AI-driven world, unless there is some programming or data gathering deficiency. This situation will not necessarily generate supervisory laxity; however, the latter is more likely to occur when leniency beneficiaries outnumber their opponents.

2. The early 2000s house-price “bubble” provides a good laxity example. Its occurrence was facilitated by supervisors allowing lenders to engage in excessive risk-taking; fundamentally, however, the bubble is attributable to its beneficiaries (“subprime” buyers and home owners) outnumbering its opponents (affluent buyers and renters).

The 2010s introduction of lending platforms offers another laxity example. Due to banks being cautious in the wake of the 2008 credit crisis, retail investors engaged in unsupervised peer-to-peer (P2P) lending. At the time, supervisory authorities tolerated the practice even though it spread transaction risks among non-professional players. This approach is now under review. P2P lending proves to be more complex than expected due to the market getting dominated by institutional lenders.

4.2 “Private” Interests as a Key Driver

Investor protection and systemic risk management are the core goals of financial regulation. Hence, AI-driven financial supervision is likely to be
constrained by (the perception of) its usefulness for small investors and its value in terms of crisis prevention.

1. The extent to which AI-use improves the situation of **small investors** is debatable. In theory, it should be easy to develop algorithms designed to protect their interests. In practice, this is difficult due to small investors often investing via pension funds and other asset managers. Nowadays, AI’s potential contribution to small investor interests seems essentially limited to small claims cases,\(^{\text{xlix}}\) a situation where AI allows for affordable litigation avenues and facilitates class action settlement.\(^{\text{I}}\)

2. AI-reliance may prove more effective when it comes to financial crisis prevention. However, it would be naïve to expect AI-use to eradicate **systemic deficiencies**, especially when it comes to time-inconsistencies and opportunistic behavior.

For example, **financial crises** can be prompted by investors preferring long-term returns while expecting banks to remain well capitalized in the short-term.\(^{\text{lii}}\) Clearly, AI-driven financial regulation will not fundamentally constrain investor preferences or thwart their natural outcome. Or, to take another example, AI-driven financial regulation cannot take care of policymakers’ tendency to deem government bonds resilient to sovereign default.\(^{\text{lii}}\) Nevertheless, AI-driven financial supervision can contribute to the early detection or, at least, the **proper management** of macro-prudential events. It follows that AI-driven financial supervision will, at worse, prove neutral in terms of systemic risks and, at best, contribute to their reduction.

3. In practice, the interests of financial intermediaries and supervisory authorities are the **key drivers** of AI-driven financial supervision.

---


Financial intermediaries can be expected to support or, at least, not to object to the use of AI for supervisory purposes. This can be attributed to they themselves increasingly using AI for management, operational and compliance purposes. For example, a 2019 survey by the Bank of England and the Financial Conduct Authority shows that 57% of respondents were relying upon AI applications for risk management and compliance purposes.\(^{iii}\) AI is also more and more used for disclosure and advice purposes\(^{liv}\) as well as for fraud detection.

Financial supervisors are in a similar position. They already rely on AI to identify financial advisers with higher risk of misconduct as well as suspicious trading activity. Other applications range from uncovering misselling in the mortgage loan and consumer credit contracts area\(^{lv}\) to the generalized detection of financial irregularities. Financial supervisors can be expected to expand AI-use across-the-board, if only to match growing AI-reliance by financial intermediaries.

4.3 Macro-Economic Impact

At this stage, AI-related job losses seem modest and limited to clerical workers.\(^{lvi}\) However, even small increases in job losses can result in public pressure driving regulatory policy off track across-the-board. For example, workers may perceive AI-driven financial supervision as a first step towards generalized use of AI in the financial industry.

1. One way to insulate AI use from labour-related constraints is to provide adequate public information.\(^{lvii}\) To be effective, communication officers must cut through the technical jargon and present the issues in simple, accessible and credible terms.\(^{lviii}\) This could prove problematic, for two

---


\(^{liv}\) Sean Hanno Williams, AI Advice: The Irony of Big Data Disclosures and the New Advice Paradigm, 2020 University of Texas Public Law Research Paper No. 718.

\(^{lv}\) Financial Stability Board, supra note 33.


\(^{lvii}\) Roger B. Myerson, Rethinking the Principles of Bank Regulation: A Review of Admati and Hellwig's The Bankers' New Clothes, 52 Journal of Economic Literature197 (2014).

reasons. To begin with, the parties directly benefiting from AI-reliance (financial intermediaries) have limited market credibility. In addition, the civil servants in charge of financial supervision may not have the expertise and, more importantly, the autonomy needed to generate market confidence.

At this point, the best way to deal with voter pressure could be to combine the disclosure of private information about the labor impact of AI-use with credible commitments of state intervention should AI-driven financial supervision have significant employment consequences.

2. AI-driven financial supervision may also result in a de facto reallocation of powers between lawmakers and financial supervisors.

It is relatively easy to spot competence issues in a non-AI environment. By contrast, abuse of supervisory powers can be hard to detect in an AI-driven world—especially when they favor well-established incumbents. Consequently, it may go unnoticed that the financial industry has become less competitive, with a negative impact on capital accumulation and income distribution. Pushed to its limits, AI-driven supervision may increase the severity of economic and social crises by prompting (ultimately inefficient) State interventions.

3. The COVID-19 environment provides good case studies. At the national level, many financial policymakers favoured the deferring of loan loss recognition to shield banks from the impact of the pandemic and give them time to ‘resurrect’. At the global level, the implementation timeline for outstanding Basel III standards got delayed to buttress banks' Covid resilience.

While the circumstances clearly warranted these interventions, efficiency requires them to be limited to countering market dysfunctionalities— e.g.

---

[³][³] To insure for adequate effectiveness, the action was taken within the Bank for International Settlements’ (BIS), by the Group of Central Bank Governors and Heads of Supervision. See https://www.bis.org/press/p201130.htm.
restricted to the period during which asset values fell below historic costs.\textsuperscript{lxii} However, extreme events are also an opportunity for \textit{industry lobbies} to get overly lenient supervisory treatment.\textsuperscript{lxiii} The COVID 19 crisis was no exception; actually, it exemplifies how to achieve such outcome—i.e. by exploiting the regulatory forbearance vs liquidity shortage trade-off.\textsuperscript{lxiv}

4. One must also take into account that AI technology and financial resources remain concentrated in the hands of a few firms and nations.\textsuperscript{lxv} This goes hand-in-hand with an increasing \textit{digital gap} at both the firm and country levels.\textsuperscript{lxvi}

At the \textbf{firm level}, there is evidence of some enterprises being better at continually improving their performance,\textsuperscript{lxvii} an outcome at least partly attributable to variance in technology and knowledge diffusion.\textsuperscript{lxviii}

At the \textbf{country level}, the digital gap is most visible between rich industrialized countries and the global south. However, this gap can also be observed within OECD countries: according to the Digital Government Index (DGI), the average score across OECD member countries was 0,5, with 15 out of 29 countries surpassing this threshold.\textsuperscript{lxix} Interestingly, Korea, the UK and Columbia were among the best performers whereas Greece, Iceland and Sweden scored the lowest.

5. Digital gaps could hamper \textbf{market access} for several reasons. First, home supervisory authorities may deem some financial intermediaries to have such an AI advantage that they cannot properly supervise them. Second, host supervisory authorities may constrain market access for financial intermediaries whose home supervisors they consider as lacking AI-savviness. Third, political considerations may generate barriers to entry

\textsuperscript{lx} Franklin Allen and Elena Carletti, \textit{The Role of Liquidity in Financial Crises}, Jackson Hole Economic Policy Symposium, available at https://repository.upenn.edu/fnce_papers/99.
for foreign financial intermediaries deemed to be AI-savvier than their domestic counterparts.

These barriers to entry could prove short-lived. AI is already facilitating the emergence of new entities that, due to their size or organization, should be less prone to submit to supervisory moral suasion.\textsuperscript{xix} At the other end of the financial spectrum, robot-advisor algorithms allow individual investors to shape their financial portfolios themselves.\textsuperscript{xixi}

6. Finally and most importantly, supervision effectiveness requires financial regulators to build predictive simulations and other technical capabilities, while limiting bias and risk amplification.\textsuperscript{xixii} There is evidence of measures taken in that direction. For example, the US SEC has developed algorithms that are five times better than random testing at detecting whether it should investigate investment adviser filings;\textsuperscript{xixiii} more fundamentally, the ECB is considering using AI for banking supervision at large.\textsuperscript{xixiv}

Clearly, it may take time for these developments to transform market dynamics. Most financial authorities still use descriptive and diagnostic analytics: recent data shows that less than 10\% rely on more advanced predictive and prescriptive analytics.\textsuperscript{xixv}

This ‘interim’ situation is costly. On the one hand, it hampers AI-use by financial intermediaries to the extent they need regulatory guidance given operational and other risks.\textsuperscript{xixvi} On the other hand, technology-advanced
intermediaries may see this state-of-affairs as an opportunity to game the regulatory system, which could be the source of economic and social harms.

7. This state of affairs is a source of supervisory concern. AI and, more generally information technology developments make it increasingly difficult to continue to discriminate among investors, in particular when this results in systemic differentiations.

Going forward, two approaches seem possible.

One is to use a step-by-step approach, under which AI-driven supervision is first introduced in areas where supervision is already or can more easily be digitalized, such as trading or retail banking. AI-driven supervision would thereafter be extended to more complex areas, such as investment banking or private banking.

The other approach is to select a small set of representative banks, for which AI-driven supervision is introduced across the board. This is a bolder approach and it would require the parallel continuation of the existing supervision mechanisms.

Both approaches raise significant conceptual and methodological issues. The latter approach has the advantage to be a voluntary one, but raises adverse selection issues. The former has the advantage of being inclusive, but may prove much harder to implement.

5 A Long Way to Go for AI-Driven Supervision

Major financial services market players and supervisory authorities increasingly rely on Al for internal purposes.

---


Monetary Authority of Singapore, Principles to Promote Fairness, Ethics, Accountability and Transparency in the Use of Artificial Intelligence and Data Analytics at 7.
On the other hand, private and supervisory use of AI seem to be diverging when it comes to financial market operations. Financial intermediaries’ transactions are increasingly AI-driven, whereas financial supervisors use of AI seems confined to detecting illegal market practices.

More importantly, AI-driven supervision is still in its infancy when it comes to systemic (as opposed to transactional) applications. This is an environment where one should not rush to adjust to technological development. Regardless of peer or political pressures, the best course of action is to adopt a step-by-step AI approach. Otherwise, AI-driven systemic supervision may well increase the severity of economic and social crises by prompting inadequate and ultimately inefficient State interventions.

It follows that there still is and should be a long way to go until financial supervision is comprehensively AI-driven.

Acknowledgment

This work is an outcome of the Future Resilient Systems project at the Singapore-ETH Centre (SEC) supported by the National Research Foundation, Prime Minister’s Office, Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) programme.

Contact

Singapore-ETH Centre
Future Resilient Systems
1 Create Way, #06-01 CREATE Tower
138602 Singapore

www.frs.ethz.ch

Publisher: Singapore-ETH Centre

© Singapore-ETH Centre, October 2021

about ECGI

The European Corporate Governance Institute has been established to improve corporate governance through fostering independent scientific research and related activities.

The ECGI will produce and disseminate high quality research while remaining close to the concerns and interests of corporate, financial and public policy makers. It will draw on the expertise of scholars from numerous countries and bring together a critical mass of expertise and interest to bear on this important subject.

The views expressed in this working paper are those of the authors, not those of the ECGI or its members.
Electronic Access to the Working Paper Series

The full set of ECGI working papers can be accessed through the Institute’s Web-site (https://ecgi.global/content/working-papers) or SSRN:

<table>
<thead>
<tr>
<th>Paper Series</th>
<th>URL</th>
</tr>
</thead>
</table>

https://ecgi.global/content/working-papers