## Money and Mental Health's response to the Bank of England/Financial Conduct Authority DP5/22 on artificial intelligence and machine learning

### Introduction

The Money and Mental Health Policy Institute is a research charity established by Martin Lewis to break the vicious cycle of money and mental health problems. We aim to be a world-class centre of expertise developing practical policy solutions, working in partnership with those providing services, those who shape them, and those using them, to find out what really works. Everything we do is rooted in the lived experience of our Research Community, a group of nearly 5,000 people with personal experience of mental health problems.

This written submission has been informed by the experiences of our Research Community, as well as our wider body of research. Unless otherwise specified, all quotes in this response are drawn directly from the Research Community.

### Background

- In any given year, one in four people will experience a mental health problem which can affect their cognitive and psychological functioning.[1] Over a lifetime, this proportion rises to nearly half the population.[2] However, we do not always know when we are unwell, or receive treatment. Over a third (36%) of people with a common mental disorder have never received a diagnosis, and 62% are not currently receiving treatment.[3]
- Common symptoms of mental health problems, like low motivation, unreliable memory, limited concentration and reduced planning and problem-solving abilities, can make managing money significantly harder.[4] As a result, it is estimated that people with mental health problems pay up to £1,550 more per year for essential services than people without mental health problems.[5]
- People with mental health problems are three and a half times more likely to be in problem debt than those without, and half (46%) of adults in problem debt also have a mental health problem.[6]

### Benefits, risks, and harms of AI

### Q1: Would a sectoral regulatory definition of AI, included in the supervisory authorities' rulebooks to underpin specific rules and regulatory requirements, help UK financial services firms adopt AI safely and responsibly? If so, what should the definition be?

---

[1] McManus S et al. Adult psychiatric morbidity in England, 2007. Results of a household survey. NHS Information Centre for Health and Social Care. 2009.
[2] Mental Health Foundation. Fundamental facts about mental health. 2016.
[3] McManus S et al. Mental health and wellbeing in England: Adult Psychiatric Morbidity Survey 2014. NHS Digital. 2016.
[4] Holkar M. Seeing through the fog. Money and Mental Health Policy Institute. 2017.
[5] Rogers C, Poll H and Isaksen M. The mental health premium. Citizens Advice. 2019.
[6] Holkar M. Debt and mental health: a statistical update. Money and Mental Health Policy Institute. 2019.

MONEY AND MENTAL HEALTH POLICY INSTITUTE

Ultimately, the value of a definition of AI is in how it shapes behaviour and outcomes. In such a wide-ranging and rapidly-evolving landscape, tight definitions of AI within financial services seem likely to become outdated quickly, as methods and applications not currently envisioned are deployed. The likely need to regularly revisit and expand a specific definition would be problematic for regulators as well as for the firms affected.

That approach also risks creating a perimeter problem, with firms incentivised to operate (or be deemed to be operating) on one side of an arbitrary line, in an attempt to minimise or simplify regulatory requirements. With applications of AI likely to stretch beyond a single market - for instance, a single tool could cover information about health but with added financial elements - a restricted focus on particular purposes may lead to regulators overlooking issues.

If the aim from a regulatory perspective is to protect consumers without disproportionate requirements of firms, a broad definition backed up by guidance from the relevant bodies seems a more appropriate approach. This would mirror how the FCA has approached complex issues like vulnerability, combining a high-level statement of what the term means, with guidance on what it can cover and the considerations firms should make when taking action in a range of circumstances. This broader approach could allow for more attention to be given to improving outcomes through more effective monitoring and enforcement, rather than interminable debates over precise definitions. This would also have the benefit of acknowledging the *existing* level of harms that many groups face as a result of AI and non-AI processes.

## Q3: Which potential benefits and risks should supervisory authorities prioritise?

**Benefits**

For consumers with and without mental health problems, AI offers the potential for greater empowerment and control over our finances. There are a number of challenges that we at Money and Mental Health regularly hear about from research participants with mental health problems that AI could help to address.

Before detailing these potential benefits, a key caveat should be noted: there is no reason to expect that many of these benefits will be developed and brought to a mass market without intervention. To use the example of Open Banking, there was much optimism surrounding what it could deliver to people in more vulnerable circumstances. It was hoped that joined-up data could help to prevent harm and improve outcomes. To date, those benefits have failed to materialise at scale, with promising products struggling to find a route to sustainability through profitability or sufficient investment.[7] Without specific action to address that gap with AI, the benefits to customers in more vulnerable circumstances seem less likely to manifest than the harms.

[7] See for instance https://www.bristol.ac.uk/media-library/sites/geography/pfrc/OB4G_Making%20a%20difference.pdf

Potential benefits that supervisory authorities should prioritise include:

- Use of transaction data and other information to enable more accurate identification of when a person is at risk of financial harm - for example, struggling to keep up with payments, gambling excessively or seeing a sudden drop in income - in order to deliver proactive support. Delays in customers reaching out for help remain a key challenge for both creditors and debt advice providers. A common symptom of many mental health problems is difficulty seeking support. Breaking down those barriers - and using AI to understand which messages at what moments are most effective - could greatly reduce the impact of financial difficulty on our long-term finances and health.

- Using similar information to inform people that they may be struggling with their mental health. Spending patterns, the time of day or night transactions are made, the number of times an app is opened or a balance checked can all be indicators of how someone's mental health is faring. In research we conducted in 2019, a participant told us: "In an ideal world… [firms] could help me to understand what areas of spend indicate that I am heading for an episode of depression. I think that there may be a pattern that I follow but can't always see it."[8]

- More effective chatbots. Many people with mental health problems struggle with using the phone, while for others face-to-face contact can be daunting. Currently, solving non-standard issues regularly requires one of these two approaches, as online services remain limited to more routine queries. Chatbots employed by many financial services firms are often poorly set up to understand someone who may be struggling to communicate, which is a common symptom of a number of mental health problems. The development of more effective chatbots, allowing people to ask complex questions, find appropriate information and address problems in a way that works for them, could be hugely beneficial. Prompts or auto-completes that more accurately capture what people are trying to say or ask would also help to reduce the burden that many people face when their mental health problem is affecting their cognitive processing.

- Adaptability and flexibility. A major downside of simpler systems is that they rely on hard and fast rules, for instance, which dates payments are taken and how payments are prioritised. This often fails to gel with the messy reality of daily life for many of us, with income from work arriving at irregular intervals for those in flexible or self-employed roles, benefit awards varying from month to month and unexpected expenses upending carefully-planned budgets. We often hear how, particularly when it coincides with a period of poor mental health, one missed payment can lead to a spiral of late fees, charges and distressing contact from creditors, worsening the cycle of money worries and mental health problems. The potential for AI to more intelligently respond to someone's evolving financial picture, to move money around accounts, or simply to notify in advance when a payment appears likely to fail, would help more of us to avoid that downward spiral of financial difficulty.

[8] https://www.moneyandmentalhealth.org/wp-content/uploads/2019/10/Data-Protecting-report.pdf

- Improved matching of customers with products and services. Many people do not shop around for the best deal. This can be for a number of reasons, including limited financial literacy or confidence, a lack of time or, particularly for people experiencing a mental health problem, difficulty finding the energy and willpower to do so. This can leave people stuck on bad deals or choosing familiar providers, even when they may not be satisfied. AI could help to address this issue, at a minimum greatly reducing the burden, or more ambitiously allowing greater automation of switching, taking into account people's preferences about communication channels or level of customer service.

**Risks**

Many of the risks outlined below exist with currently-used models and approaches. The concerns below are priorities because the potential power and scale of AI models could deepen and spread the harm. We have also focused on those risks that seem more immediately present, but recognise that some of the potential risks further in the future could be even more damaging. Regulators should be actively monitoring developments that could pose a risk, commissioning research with potentially affected groups and evaluating the impact of new products on specific groups of consumers like people with mental health problems.

- Badly-managed consent. When people knowingly sign up for services - and can withdraw as easily as they signed up - the risk of people's data being used in ways they were not anticipating is reduced. Without that, the potential for a breach of trust is high, undermining the potential to improve outcomes. Where AI enables greater automation or the identification of sensitive personal information such as health conditions, the risks of acting without informed consent are that much greater. Long and confusing terms and conditions statements, permissions that are excessively broad or last for longer than needed all raise this concern.

- Anonymity. The potential breadth of information gathered and used means the dangers of leaks or hacks from AI could be severe. These are of particular concern when it comes to mental health problems, which is a sensitive and often private topic.

- Incomplete data. One in two of us will experience a mental health problem across our lifetimes. As explored in more detail in response to Q6, a far smaller share of us choose to disclose that we have a mental health problem to financial service providers, or may share that information with some firms but not others. AI models that draw conclusions about people with mental health problems based on the subsection who feel comfortable disclosing may be drawing inaccurate conclusions about people who AI identifies as potentially having a mental health problem.

- Outdated data and historical data biases. Mental health problems fluctuate and can be recovered from. This can mean that someone who experienced a period of poor mental health years ago may never experience it again. Similarly, as societal attitudes towards mental health problems have improved in recent years, it may be that outcomes and patterns observable who are flagged in historic data as having had a mental health problem may not be representative of that population today. This is a consideration we have already raised concerns about within the insurance industry, with the lack of transparency about what data is used to inform risk assessments leading to worries that those decisions are based on a different time, with a qualitatively different group of people and without reflecting the changes in medications and treatment now available.[9]

- Discrimination based on accurate data. AI-driven decisions will be more accurate in some cases but that will still result in bad outcomes for people in certain circumstances. Our response to Q6 expands on this point in more detail but the risk of highly-personalised decision-making leading to financial exclusion seems particularly concerning.

- Exploitation of behaviour driven by symptoms of mental health problems. Symptoms of mental health problems can leave us more susceptible to techniques used by firms intended to lead to a certain action. For instance, as recent work by the CMA has demonstrated, online choice architecture can greatly influence our purchasing decisions, leading people to buy products out of a sense of scarcity or a discount, neither of which may actually be the case. But our past research has shown how some people with mental health problems can be prone to spending compulsively at night or in response to advertising. Without effective regulation, AI opens up the possibility of that manipulation becoming even more targeted, latching onto the signs of poor mental health discussed above and taking that as an opportunity to push certain products or services.

A major concern regarding governance is that the above harms could all be happening without firms' or regulators' awareness. Decent data collection and analysis (as required already through the Consumer Duty) should mean that higher-level trends are visible, for instance how certain groups of consumers are faring. But the complexity of AI models means the exact products and mechanisms through which the harm is being caused may be harder to pin down.

## Q6: How could the use of AI impact groups sharing protected characteristics? Also, how can any such impacts be mitigated by either firms and/or the supervisory authorities?

Our response to Q3 covers how people with mental health problems - mostly included under the protected characteristic of disability - may be affected by the use of AI. On mitigation, there may be opportunities related to disclosure of mental health problems to firms. Financial services

[9] https://www.moneyandmentalhealth.org/wp-content/uploads/2023/02/MMHPI-Written-Off-insurance-report.pdf

firms we speak to often stress that they have support available or can offer reasonable adjustments when they know that customers are experiencing poor mental health.

But the customers who do disclose to firms are only the tip of the iceberg. In a representative poll of 5,000 people with mental health problems we commissioned in the summer of 2021, we found that only 14% of respondents had *ever* told a financial service firm about their condition.[10] Of those 14%, it is likely that they have not always disclosed, and not to every firm they engage with, meaning actual awareness among firms is likely to be much lower. Beyond that group, the gold-standard NHS study into the prevalence of mental health problems found that over a third (36%) of people with a common mental disorder have never received a diagnosis, and sometimes we are not aware that what we are experiencing constitutes a mental health problem.[11]

This means that strategies to prevent or even mitigate discrimination against people with mental health problems that rely primarily on disclosure will be partial at best. For instance, one option to limit discrimination could be that there are additional layers of checks - either human-led or not - that would kick into action when there is a flag related to a protected characteristic or a vulnerability. But with disclosure of mental health problems being so limited, this would be ineffective.

That is not to say, however, that firms and supervisory authorities cannot do anything with disclosed information. For instance, an insurer we have spoken to does not offer automated quotes to people who declare a mental health problem in cases when this increases the price, leads to exclusions or means they choose to decline them altogether. While the shortcomings of this ‘human in the loop’ process noted in the DP are fair, this additional step does at least create the opportunity for potentially discriminatory decisions to be acknowledged and logged - in order for the firm to take action or for regulators to ask to see when carrying out monitoring.

The more firms do to encourage disclosure of mental health problems, the more this could improve these methods. Money and Mental Health in collaboration with Money Advice Trust published a three-part guide on how firms can encourage, respond to and record disclosures effectively.[12]

The inference of potential mental health problems from data is another opportunity. If firms use this data as part of their monitoring, this could help to overcome the very limited picture they have of mental health problems among their customers. That said, caution would need to be taken on how that inference is made, stored and used in future, for instance on lending decisions. To protect privacy, firms should provide customers with choice and control over

[10] https://www.moneyandmentalhealth.org/wp-content/uploads/2021/11/The-State-Were-In-Report-Nov21.pdf

[11] https://www.moneyandmentalhealth.org/money-and-mental-health-facts/

[12] https://www.moneyandmentalhealth.org/wp-content/uploads/2022/11/Discosure-Guide-1-Disclosure-Environments.pdf

whether and how their data is analysed to identify potential vulnerability or protected characteristics. This should include the option to opt out, and allowing customers to choose what types of potential vulnerability they are comfortable with the firm looking for.

From the point of view of supervisory authorities, an ability to understand exactly what firms are doing when it comes to their data, models and governance is already an issue. Returning to the example of insurance, there is little to suggest that the FCA has a clear picture of what data insurers draw upon in relation to mental health problems or how exactly it is used. Being able to effectively regulate existing practices will require more of an active approach and confronting the trade-off between commercial sensitivities and the protection of consumers.

To help some of the potential harms raised in our response, supervisory bodies should involve people with mental health problems in the design of AI-driven solutions, and also in considerations about how they are communicated and explained to customers. Testing and co-producing interventions with customers would help firms to develop interventions that reach a balance between being effective and unintrusive.

### Q7: What metrics are most relevant when assessing the benefits and risks of AI in financial services, including as part of an approach that focuses on outcomes?

With the focus of regulation of AI being on the outcomes it has for consumers, the metrics used need not be fundamentally different to those currently in use. For instance, unfair treatment of customers with mental health problems might be identified through metrics related to prices charged or fees incurred. But the vast wealth of data and the increased potential to analyse it opens up more opportunities to explore longitudinal impacts - for instance whether communications from creditors leads over the long-term to worse health outcomes. This is already possible, as Muggleton et al's research on gambling has demonstrated.[13] But the deployment of AI offers the chance to greatly expand firms' knowledge.

Whichever metrics on outcomes are used, they should be employed to understand differences in outcomes for different groups of vulnerable customers - to understand if they are worse for any one particular group (as the FCA's vulnerability guidance suggests). This means that mental health data must be collected to at least some extent in order to be able to draw comparisons.

## Regulation

### Q9: Are there any regulatory barriers to the safe and responsible adoption of AI in UK financial services that the supervisory authorities should be aware of, particularly in relation to rules and guidance for which the supervisory authorities have primary responsibility?

[13] Muggleton N, Parpart P, Newall . et al. The association between gambling and financial, social and health outcomes in big financial data. Nat Hum Behav 5, 319–326 (2021). https://doi.org/10.1038/s41562-020-01045-w

Contact: conor.darcy@moneyandmentalhealth.org

The overlapping nature of AI means collaboration will be essential. Nimble regulatory responses that act first, rather than dwelling too long on exactly whose responsibility it is, will be needed. This will require careful thought about the objectives of the varying organisations, the skillsets of staff, the metrics by which success is measured and the incentives within organisations.

It will necessarily also require increased knowledge of the methods used in firms. With large financial institutions having massive budgets, far in excess of the UK's supervisory bodies, collaboration and inventiveness will be required. As noted, AI systems are likely to stretch beyond single-purpose uses, or even remain solely within financial services. Ensuring that the supervisory authorities work closely with partners across the government and regulatory space, including beyond the UK, will be important to manage the scale of the challenge. From the point of view of protecting customers with mental health problems, collaboration with partners in health will be vital. See also our response to Q11 on the Equality Act.

### Q11: How could current regulation be simplified, strengthened and/or extended to better encompass AI and address potential risks and harms?

While stretching beyond regulation, a major weakness in the current legislative system is how the Equality Act 2010 is (not) enforced in financial services. As our 2022 paper on the issue outlined, the FCA has the skills but not the remit to take action on apparent breaches of the Equality Act in financial services, while the Equality and Human Rights Commission has the remit but not the access or the budget to be able to effectively oversee equalities issues in the sprawling and complex world of financial services.[14]

A realism about the limits of financial regulation is also needed. If the market isn't working for some groups even with firms adhering to the rules, it will require supervisory authorities to make clear to government when this is occurring and what responses might be needed. This could be a prohibition or restriction of certain tools but it could instead be a social policy response, for instance acting as an insurer of last resort as has been the case with Flood Re.

[14] https://www.moneyandmentalhealth.org/wp-content/uploads/2022/02/MMH-Time-to-Act-Report-WEB-030222.pdf

Contact: conor.darcy@moneyandmentalhealth.org