

*Analysis*

Does the law allow non  
data subjects to challenge  
algorithmic harms?

Alex Lawrence-Archer & Ravi Naik

October 2023



---

## **Analysis: does the law allow non-data subjects to challenge algorithmic harms that affect them?**

---

<b>A. Summary of analysis, background, and instructions .....</b>	<b>2</b>
I. Summary of analysis .....	2
II. Background and instructions .....	3
<b>B. The Equality Act and the EHRC .....</b>	<b>3</b>
I. Indirect discrimination .....	<i>Error! Bookmark not defined.</i>
II. Public Sector Equality Duty .....	6
III. The EHRC's powers and enforcement approach.....	<i>Error! Bookmark not defined.</i>
<b>C. Scenario 1: police deployments.....</b>	<b>3</b>
I. Decision-making context.....	9
II. Voluntary guidance on the use of AI in the public sector.....	10
III. PSED duty .....	<i>Error! Bookmark not defined.</i>
IV. General public law principles.....	14
V. Challenge by way of judicial review .....	20
VI. The IOPC .....	24
VII. Summary .....	<i>Error! Bookmark not defined.</i>
<b>E. Scenario 2: 'dynamic' rail pricing.....</b>	<b>25</b>
I. Decision-making context.....	27
II. Consumer protection and competition law.....	28
III. Equality Act.....	39
IV. Summary .....	<i>Error! Bookmark not defined.</i>
<b>F. Scenario 3: content moderation .....</b>	<b>43</b>
I. Decision-making context.....	44
II. Terms of service .....	45
III. Self-regulation .....	46
IV. Online Safety Bill.....	47
V. Indirect discrimination under the Equality Act.....	52
VI. Summary .....	<i>Error! Bookmark not defined.</i>
<b>G. Conclusion: gaps in groups' ability to challenge uses of AI that may harm them.....</b>	<b>26</b>

## A. Summary of analysis, background, and instructions

### I. Summary of analysis

1. We consider three Scenarios which show how automated decision-making can lead to harms to groups of individuals without their personal data being processed.
2. For an algorithm used by the Metropolitan Police (the 'Met') to allocate police patrols in a way that leads to over-policing of areas with above-average ethnic minority populations, public law places some constraints on automation, but does not prevent it. With good evidence, the use of the tool could be challenged on the basis that it takes account of irrelevant considerations. However the Scenario demonstrates how public law is only beginning to be applied in the field of automated decision-making, making any challenge rather uncertain, which is compounded by the costs risks involved and challenges in obtaining permission to bring a challenge within the very limited time period allowed. Most clearly, the law would require the Met to carry out an Equality Impact Assessment regarding the use of the tool, which would at least improve transparency and might curb the worst excesses of self-reinforcing over-policing of certain communities.
3. For a train company setting rail fares using real-time dynamic pricing, consumer, competition, and equality law do not prevent automation. Competition and consumer law put strong guardrails in place as to *how* automation may be implemented, ensuring that customers are not misled, and that systematic errors in pricing should eventually be dealt with. This is an area where enforcement – both through regulation and through collective proceedings – is notably strong, providing protection beyond just the letter of the law against automation gone awry. Whilst this kind of area-based dynamic pricing could in theory be indirectly discriminatory, we consider it likely to be objectively justifiable and therefore lawful, even if it has some differential impacts on protected groups.
4. For a social media platform using automated content-moderation which excessively removes lawful content and penalises non-English language posts, the forthcoming Online Safety Bill<sup>1</sup> makes some efforts towards protecting freedom of expression online and ensuring equal treatment according to platforms' terms of service. These provisions are weak, however, and likely counterbalanced by the majority of the Online Safety Bill, which creates strong incentives to automate and expand content moderation, and to

---

<sup>1</sup> At the time of writing the Online Safety Bill was in final form and soon to receive Royal Assent.

make terms of service more restrictive and open to platform interpretation. A claim for indirect discrimination could theoretically succeed, but would entail one or more individuals taking a significant risk in bringing civil proceedings against the platform. Thus the law *encourages* this kind of automation harm rather than constraining it.

5. Taken together we find that these Scenarios show that it is (even) harder to challenge algorithmic harms where those affected are not data subjects for the purposes of the UK General Data Protection Regulation. Even to the limited extent that the law does set limits on automation, obtaining the evidence to mount challenges is likely to be very difficult, and retrospective litigation does not give those affected a voice in how technology impacting them develops in the long-term.

## II. Background and instructions

6. We are instructed to consider three examples of automated decision-making through the use of data which may lead to harm to individuals or communities.
7. Often, when automated decision-making causes harm, those affected will be having their personal data processed by the decision-maker (the ‘controller’) within the meaning of the UK GDPR. Much case law, regulatory enforcement and analysis has therefore focused on how individuals’ rights as data subjects under the UK GDPR may enable them to challenge and seek redress for automation harms<sup>2</sup>. There has been less analysis of the legal position where those affected by automated decision making are *not* also data subjects for the purposes of the UK GDPR. The three scenarios we analyse are designed to test the extent to which the law enables such individuals or groups to challenge, limit and/or prevent harms which may be caused by automated decision-making.

## B. The Equality Act and the EHRC

8. The Equality Act 2010 (the ‘**Equality Act**’) has relevance to all three Scenarios, dealing as it does with unfair treatment of certain groups. We therefore introduce it here. The Equality Act regulates discrimination on the basis of a specified range of protected characteristics (§4-12):

---

<sup>2</sup> See by way of example Mezaros & Ho, *AI research and data protection: Can the same rules apply for commercial and academic research under the GDPR?* (2021) *Computer Law & Security Review* 41, and Abrusci & Mackenzie-Gray Scott, *The questionable necessity of a new human right against being subject to automated decision-making* (2023) *International Journal of Law and Information Technology* 31:2

- age;
- disability;
- gender reassignment;
- marriage and civil partnership;
- pregnancy and maternity;
- race (which includes nationality – s.9);
- religion or belief;
- sex;
- sexual orientation.

I. Direct and indirect discrimination

9. *Direct* discrimination (s.13) is defined as follows:

*“A person (A) discriminates against another (B) if, because of a protected characteristic, A treats B less favourably than A treats or would treat others.”*

10. In the context of data-driven automated decisions, where less favourable treatment is associated with a variable that is – or is a proxy for - a protected characteristic, this will be direct discrimination, but only where there is ‘*exact correspondence*’ between the variable/proxy and the protected characteristic<sup>3</sup>.

11. *Indirect* discrimination (s.19) is defined as follows:

*“(1) A person (A) discriminates against another (B) if A applies to B a provision, criterion or practice which is discriminatory in relation to a relevant protected characteristic of B’s.*

*(2) For the purposes of subsection (1), a provision, criterion or practice is discriminatory in relation to a relevant protected characteristic of B’s if—*

*(a) A applies, or would apply, it to persons with whom B does not share the characteristic,*

*(b) it puts, or would put, persons with whom B shares the characteristic at a particular disadvantage when compared with persons with whom B does not share it,*

*(c) it puts, or would put, B at that disadvantage, and*

---

<sup>3</sup> The limitations of this approach are described and critiqued in Adams-Prassl, Binns & Kelly-Lyth, *Directly Discriminatory Algorithms* (2022) *Modern Law Review* 86:1

(d) *A cannot show it to be a proportionate means of achieving a legitimate aim.*"  
(emphasis added)

12. The scope of indirect discrimination is therefore much broader: there need not be 'exact correspondence' between variables leading to less favourable treatment and protected characteristics, provided that those with protected characteristics are at significant risk of being treated unfavourably due to an automated decision-making system<sup>4</sup>. However unlike direct discrimination, indirect discrimination is not always unlawful; it may be justified where the provision, criterion, or practice ('PCP') is a proportionate means of achieving a legitimate aim.
13. Whether a discriminatory PCP can be justified will depend on the facts of each specific case. Of particular note is that the need to save costs *alone* cannot justify indirect discrimination<sup>5</sup>. However, the courts have loosened this constraint by consistently finding that the need to save costs in combination with some other factor(s) (sometimes known as the 'costs-plus' approach) will be enough to show that a defendant has a legitimate aim in indirectly discriminating<sup>6</sup>. Academic commentators have noted that English courts have taken a relatively flexible approach to the question of proportionality and whether indirect discrimination may be justified<sup>7</sup>. The Competition and Markets Authority has itself described the enforcement situation for indirect discrimination as 'difficult'.<sup>8</sup>
14. The Equality Act makes both direct and indirect discrimination unlawful in certain circumstances including, relevantly to this analysis, where a person provides a service to the public or a section of the public (for payment or not) (s.29 EA). A breach of s.29 gives rise to a private right of action for the individual(s) affected (Part 9 Chapter 2 of the Act).
15. The Equality Act does not create 'group rights': any breach would need to be enforced by one or more individuals directly affected by the discrimination complained of. It does however create a duty for public authorities which is relevant to more indirect impacts on groups: the Public Sector Equality Duty.

---

<sup>4</sup> *Ibid*

<sup>5</sup> *Woodcock v Cumbria Primary Care Trust* [2012] EWCA Civ 330

<sup>6</sup> *Heskett v Secretary of State for Justice* [2020] EWCA Civ 1487

<sup>7</sup> Lane & Ingleby, *Indirect Discrimination, Justification and Proportionality: Are UK Claimants at a Disadvantage?* (2018) *Industrial Law Journal* 47:4

<sup>8</sup> <https://www.gov.uk/government/publications/algorithms-how-they-can-reduce-competition-and-harm-consumers/algorithms-how-they-can-reduce-competition-and-harm-consumers>

## II. Public Sector Equality Duty

16. s.149 Equality Act places a duty (the 'PSED') on all public authorities, in the exercise of their functions to:

*“have due regard to the need to—*

*(a) eliminate discrimination, harassment, victimisation and any other conduct that is prohibited by or under this Act;*

*(b) advance equality of opportunity between persons who share a relevant protected characteristic and persons who do not share it;*

*(c) foster good relations between persons who share a relevant protected characteristic and persons who do not share it.”*

17. What is required in order to comply with the PSED will be context specific. The Equality and Human Rights Commission ('EHRC') describes compliance with the PSED in the following terms:

*“The general equality duty therefore requires organisations to consider how they could positively contribute to the advancement of equality and good relations. It requires equality considerations to be reflected into the design of policies and the delivery of services, including internal policies.”*

18. Often, public authorities (attempt to) comply with the PSED by carrying out an 'Equality Impact Assessment' of a policy or practice which may raise Equality Act issues, although the law does not explicitly require this<sup>9</sup>. An Equality Impact Assessment seeks to identify ways in which the policy creates relevant risks, and to mitigate them.

## III. Enforcement

19. Scenarios 2 and 3 raise the prospect of indirect discrimination in the context of the provision of services, contrary to the Equality Act §19 and 29. For the reasons given in those sections, it is rather unlikely that the conduct described in Scenarios 2 and 3 constitutes unlawful discrimination. We therefore do not explore enforcement in detail, but it could be achieved through action by the EHRC or, failing that, through (a) civil claims by (an) affected individual(s).

---

<sup>9</sup> See Hickman *Too hot, too cold or just right?*, Public Law, April 2013 and (e.g.) R (D) v Worcestershire County Council [2013] EWHC 2490

a) The EHRC

20. The Equality and Human Rights Commission ('EHRC') is tasked by the Equality Act 2006 (s.3; as amended by the Equality Act 2010) with:

*“encouraging and supporting the development of a society in which—*

*(a) people’s ability to achieve their potential is not limited by prejudice or discrimination,*  
*(b) there is respect for and protection of each individual’s human rights*  
*(c) there is respect for the dignity and worth of each individual,*  
*(d) each individual has an equal opportunity to participate in society, and*  
*there is mutual respect between groups based on understanding and valuing of diversity and on shared respect for equality and human rights.”*

21. As well as having general powers to undertake research and publish guidance (§13-19 Equality Act 2006), the EHRC is empowered to:

- Carry out investigations into whether someone has breached equality law (s.20) including a power to compel information (s.20 and Sch 2);
- Enter into binding agreements with entities regarding compliance (s.23);
- Issue a notice requiring that person to prepare an action plan or recommending action to be taken, where a breach of equality law has been found (s.21); and
- Institute or intervene in judicial review proceedings (s.30)<sup>10</sup>.

22. The EHRC cannot directly impose a legal requirement on a person to take steps to comply with the Equality Act or levy fines. Rather, it must instead apply to the court to secure compliance with an action plan (s.22) or for an injunction to restrain an unlawful act (s.24).

23. There is no formal mechanism by which individuals can bring complaints about infringements of the Equality Act to the EHRC, and organisations are not required to report on infringements. Further, there is some evidence that the EHRC makes relatively sparing use of the limited enforcement powers it has. In 2019 a Select Committee Report<sup>41</sup> stated:

*“As an organisation [the EHRC] must overcome its timidity and be bolder in using the existing powers that only it has.”*

---

<sup>10</sup> The EHRC may also intervene in other proceedings under this section.



24. Since that report, enforcement action by the EHRC does not appear to have significantly increased. The most recent legal intervention listed on its site is from 2020 and the most recent investigation was launched in 2021. The EHRC's *Human Rights Legal Cases* page has not been updated since 2016, and the most recent agreement under s.23 listed is from 2018<sup>11</sup>. The EHRC itself has emphasised the challenges it faces to the Government's AI White Paper<sup>12</sup>, stating:

*"While the Commission is committed to the regulation of AI under both equality and human rights law, these additional duties are outside our current business plan commitments and unfunded. The Government must invest in the Commission and other regulators to ensure that the regulatory community can build the capacity and expertise needed to support safe, responsible and ethical innovation in AI."*

b) Civil claims

25. Any civil claim for indirect discrimination in the provision of services would be brought in the county court (s.114 Equality Act), a process characterised by significant barriers:
- i) On top of court fees, any claimant would likely need legal advice, which would be expensive unless *pro bono* advice can be found. This is especially likely for Scenarios 2 and 3 in which any claim of indirect discrimination would be novel and evidentially complex. The need for injunctive relief in both cases suggests any claims would need to be brought on the county court's 'fast track' (or possibly even the intermediate or multi track), which is not explicitly designed for unrepresented claimants to use (unlike the 'small claims' track). Whilst legal aid is in theory available for claims under the Equality Act, this is means-tested, and the income threshold is very low. Support with legal costs may be available from the EHRC under §28-9 of the 2006 Equality Act, but this is rare in practice<sup>13</sup>.
  - ii) Any claimant would face the risk of being made to pay the Defendant's (i.e. the train company in Scenario 2 or the social media platform in Scenario 3) legal costs if unsuccessful, which could run into the tens or even hundreds of thousands of pounds, depending on the track used and how the cases proceed<sup>14</sup>.

---

<sup>11</sup> <https://www.equalityhumanrights.com/en/legal-casework/enforcement-work>

<sup>12</sup> <https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper> and <https://www.equalityhumanrights.com/en/file/43396/download> at [45] for the response.

<sup>13</sup> The EHRC lists only 14 instances of this ever having happened.

<sup>14</sup> Costs risk would be lower on the fast or intermediate tracks, which are subject to fixed recoverable costs, although even here there is a risk of higher adverse costs being awarded in certain situations. The rules on track allocation and cost recoverability are complex and not explored in detail here.

## C. Scenario 1: police deployments

*The Metropolitan Police introduces a new data-driven tool to determine how and where physical police patrols will be allocated. The system uses fully anonymised statistics on past crimes and arrests to ‘predict’ where offences such as public disorder, possession of controlled substances, and possession of bladed articles etc. are most likely to take place. In line with the Home Office’s urging<sup>1</sup>, stop-and-search powers will be used proactively during these patrols.*

*The system results in increased allocations of patrols to neighbourhoods with historically high recorded arrest rates. These areas also happen to be areas with a proportion of ethnic minority residents significantly higher than the national average.*

**Processing involved:** *The system does not process any personal data since all inputs are fully anonymised.*

**Harm:** *The system results in continued over-policing of historically over-policed communities which poses a risk to rights to life, dignity, liberty (among other rights) as well as rights to non-discrimination for individuals living in them. This over-policing risks being seriously biased and inaccurate in that it merely replicates arrest patterns of the past rather than reflecting current patterns and developments, reflecting the police’s greater motivation to use stop-and-search powers in certain areas.*

### I. Decision-making context

#### a) Responsibility

26. There is no statutory scheme which governs decisions on where to place police patrols. It is an operational decision for each force (ultimately, for the Chief Constable and in the case of the Met the equivalent role is the Commissioner of the Police of the Metropolis, the ‘**Commissioner**’). The decision to patrol – and where to do so – is taken under police force’s common law powers to execute its duty to prevent and detect crime.<sup>15</sup> As a starting point, this means that the Met in this example has a relatively wide ambit of discretion in deciding where to patrol – and indeed in deciding *how to decide* (i.e. whether to use algorithmic tools) since there is no explicit law setting out which factors must (or must not) be taken into account. Ultimately the Commissioner is accountable to the Mayor of London and the Home Secretary.

---

<sup>15</sup> Halsbury’s Laws, Vol 84 (Police and Investigatory Powers), para 1 and 40; Rice v Connolly [1966] 2 Q.B. 414

b) Transparency

27. There is no general duty for the police to explain how operational decisions such as patrol placement are made. The Met is a public authority for the purposes of the Freedom of Information Act 2000 ('FOIA'). In theory this means it could be asked to disclose information it holds on how patrols are allocated under s.1. However, s.31 FOIA exempts information from disclosure in the law enforcement context if disclosure would prejudice police core functions. Past FOIA requests show how forces have applied this exemption to refuse to provide information on how patrols are allocated<sup>16</sup>.
28. The s.31 exemption would likely apply in this scenario, meaning anyone challenging the use of the tool in this Scenario would struggle to obtain detailed evidence about how the tool works and how it is used to inform or replace human decision-making on patrol allocation<sup>17</sup>.

II. Voluntary guidance on the use of AI in the public sector

29. There is a wide range of voluntary public sector guidance on the use of algorithmic decision-making including the Guide to Using AI in the Public Sector<sup>18</sup> and the Data Ethics Framework<sup>19</sup> which are relevant to the tool in this Scenario<sup>20</sup>. However, as this guidance is voluntary and there is no central body with responsibility for monitoring compliance with it, it cannot form the basis of a challenge to the use of the patrol-allocation tool in this Scenario.

III. Equality Act

a) Direct and indirect discrimination likely not relevant

30. Whilst the protected characteristic of race is clearly at play (as explicitly stated in the Scenario), direct and indirect discrimination (see section B.**Error! Reference source not found.** above) are likely not relevant to this Scenario. Both require 'a person' to be treated less favourably than another without the relevant protected characteristic (which

---

<sup>16</sup> FOI Request F-2021-02781 to the Police Service of Northern Ireland: <https://www.psnipolice.uk/sites/default/files/2022-10/02781%20PSNI%20Patrols.pdf>

<sup>17</sup> The Government's Algorithmic Transparency Standard - <https://www.gov.uk/government/collections/algorithmic-transparency-reports> - is potentially relevant to the tool in this Scenario, but it is entirely voluntary and we assume for the purposes of this analysis that the Met has not provided any public information on the use of the patrol-allocation tool.

<sup>18</sup> <https://www.gov.uk/government/collections/a-guide-to-using-artificial-intelligence-in-the-public-sector>

<sup>19</sup> <https://www.gov.uk/government/publications/data-ethics-framework>

<sup>20</sup> See *Effective Protection from AI Harms*, AWO (2023) from §231 for a more comprehensive list: <https://www.awo.agency/files/AWO%20Analysis%20-%20Effective%20Protection%20against%20AI%20Harms.pdf>

comparator may be hypothetical). The harm that we are instructed to consider in this Scenario does not concern unfavourable treatment of one – or even a group of – individuals. Rather it arises from an automated practice which might be expected, indirectly, to have a negative impact on groups over the longer-term due to the feedback loops at play, which many would consider unfair and socially undesirable.

31. Even if it were possible to argue that there was a ‘person’ who could bring a discrimination claim (in the sense of being affected as an individual), the mere fact of there being more patrols in that person’s community is very unlikely – without more – to meet the definition of either ‘*less favourable*’ treatment (direct discrimination) or ‘*particular disadvantage*’ (indirect discrimination)<sup>21</sup>. Treatment meeting these thresholds could only arise from a direct interaction with the patrols, in which case the decision to allocate to them to the area would be no more than background to the substantive issue.

b) The PSED in this scenario

32. s.150 Schedule 19 Equality Act provides that the Met is a public authority for the purposes of the PSED. The Met therefore must have “due regard”<sup>22</sup> to equality and discrimination when adopting and using the patrol-allocation tool.
33. The Scenario does not state whether the Met carried out an Equality Impact Assessment or took any other steps in an attempt to comply with the PSED. For the purposes of this analysis we assume that any Equality Impact Assessment that was conducted **did not** consider the risk that the algorithmic tool would entrench over-policing in areas with above-average ethnic minority populations. A failure to consider that risk could well constitute a breach of the PSED.
34. In this Scenario, the protected characteristic (s.4 Equality Act) is stated to be race. The Met must have due regard to the need to eliminate discrimination (including indirect) relating to race when allocating patrols (s.149(1)(a)). Implementing a system which systematically over-allocates patrols to areas with above-average ethnic minority populations without putting any safeguards in place would seem to be a failure to have due regard to that need.

---

<sup>21</sup> Indeed, many would argue that having more patrols in an area is *more* favourable treatment.

<sup>22</sup> The EHRC has provided guidance on the meaning of ‘due regard’: <https://www.equalityhumanrights.com/en/corporate-reporting/public-sector-equality-duty#:~:text=The%20Act%20helpfully%20explains%20that,the%20needs%20of%20other%20people>

35. A relatively small difference in treatment of people with a protected characteristic is enough to suggest indirect discrimination<sup>23</sup>. Thus the over-policing through extra patrols could suffice even if the overall policing of different areas remains substantially the same. It has been argued<sup>24</sup> that the courts' affirmation of the '*exact correspondence*' test (see B. above)<sup>25</sup> severely limits the application of the Equality Act to the use of automated decision-making. However, the '*exact correspondence*' test is primarily relevant to determining whether there is *direct* discrimination under the Equality Act. Indirect discrimination is broader, and the factors to be considered for the PSED broader still. The PSED may be engaged:
- i. Where there is a risk of indirect discrimination (which may be the case here); or
  - ii. Where there is a need to consider the advancement of '*equality of opportunity between persons who share a relevant protected characteristic and persons who do not*' (s.149(1)(b) – this includes minimising disadvantages and encouraging participation in public life); or
  - iii. Where there is a need to consider the fostering of '*good relations between persons who share a relevant protected characteristic and persons who do not*' (s.149(1)(c)).
36. Thus even if the risk of indirect discrimination as a result of the use of the patrol-allocation tool is slight, we conclude that a failure by the Met to consider in any way the harm described in the Scenario would likely be a breach of the PSED, as §149(1)(b) and (c) are engaged.
37. The PSED is a continuing duty<sup>26</sup>, meaning it is not only the presence or absence of an Equality Impact Assessment at the outset which is relevant. If discrimination and equality issues with the patrol-allocation tool become evident over time, there would be a duty on the Met to either carry an assessment out, or update it to take account of those new findings.

---

<sup>23</sup> London Underground v Edwards (No 2) [1999] ICR 494 (CA)

<sup>24</sup> Casale, *Around the Black Box: Applying the Carltona Principle to Challenge Machine Learning Algorithms in Public Sector Decision-Making*, (2021) LSE Law Review Vol VII

<sup>25</sup> Notably in R (Coll) v Secretary of State for Justice [2017] UKSC 40 at [28]

<sup>26</sup> Forward v Aldwyck Housing Group Ltd [2019] EWCA Civ 1334, [37]

38. An Equality Impact Assessment by the Government in 2021<sup>27</sup> regarding the expansion of stop and search shows (i) that the potential for over-policing of people with the protected characteristic of race engages the PSED, and (ii) gives a sense of the extent of consideration of the issue required in order to discharge the duty. Although impact assessments are often published, those affected in this would not necessarily know whether one had been carried out, since there is no legal requirement to publish them. A range of ways may be imagined in which the assessment could become public (beyond voluntary publication by the Met) such as:
- i. Through FOIA;
  - ii. Through disclosure other proceedings by an individual relating to a specific action by the police (or brought by someone challenging the *lack* of police patrols in their area) on patrol in one of the areas chose by the patrol-allocation tool; or
  - iii. Through actions by police officers themselves (e.g if they were to challenge the use of the tool on employment grounds).
39. *R (Bridges) v South Wales Police* [2020] EWCA Civ 1058 (from [173]) demonstrates how the PSED will be engaged where the use of an algorithmic tool may lead to indirect discrimination in the context of the use of a facial recognition system used to identify suspects in a crowd. It shows that in this Scenario the PSED would extend to a full consideration of the possible equality impacts of the use of the tool. It would not be enough for the Met to ignore issues merely because there is no *direct* evidence of discrimination (at [199]).
40. *Bridges* also shows that a lack of human oversight of algorithmic recommendations may breach the PSED, since it is a non-delegable duty. Thus if the Met does not have human officeholders properly scrutinising the patrol-allocation tool's recommendations, this would be further evidence that the PSED has not been complied with. In *Bridges*, the mere presence of a human officer viewing and approving a recommendation by the facial recognition algorithm was held not to be enough to discharge the PSED, showing that in this Scenario the Met would need to demonstrate that officers are genuinely using their own judgment in combination with any recommendations from the patrol-allocation tool.

---

27

[https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/1079780/Final\\_EIA\\_for\\_s60.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1079780/Final_EIA_for_s60.pdf)

41. A successful challenge (or EHRC enforcement) might lead to improved practices by the Met – for example more human review of the tool’s recommendations or tweaking of the tool’s operation. Academic commentators have emphasised the potential of the PSED in this regard:

*“Although the PSED is a procedural rather than substantive obligation and decision-makers are not prevented from simply making the same decision again after it has been quashed, the duty may produce substantive outcomes indirectly due to its potential to combat algorithmic opacity.”<sup>28</sup>*

42. But this is not guaranteed. Indeed, it is a significant limitation of the PSED that any successful enforcement of the duty (whether by the EHRC – see **B.Error! Reference source not found.** above – or by way of judicial review) would not rule out the use of the algorithmic tool or even necessarily lead to significantly different outcomes for the communities affected in this Scenario.
43. In this Scenario the PSED duty requires the Met to have regard to equality issues. Provided it does so – e.g. through a comprehensive Equality Impact Assessment – it is for the Met to determine what *weight* those issues should be given in its overall decision-making<sup>29</sup>. It would be open to the Met – having taken relevant equality considerations into account – to stick to their broad judgment that past crime statistics are the best basis on which to allocate patrols, even if this leads to some ‘over-policing’ of areas with above-average ethnic minority populations. Such an approach would be in compliance with the PSED.

#### IV. General public law principles

44. Public law allows the courts to review a decision made by a public authority on a relatively limited number of grounds. One is that the decision is unlawful due to a breach of a statutory duty. Thus a breach of the PSED would be grounds for judicial review of a decision. There are also a number of more general grounds on which the decision(s) described in this Scenario might be challenged:

---

<sup>28</sup> Seglias, *Bias and Discrimination in Opaque Automated Individual Risk Assessment Systems: Challenges for Judicial Review under the Equality Act 2010* (2021) The Oxford University Undergraduate Law Journal Issue X

<sup>29</sup> See e.g. *R (Harris) v London Borough of Haringey* [2010] EWCA Civ 703 at [40] and *Bridges* at [175]

a) Locating the decision

45. Analysing the possibility of a challenge on general public law grounds requires clarity on the ‘decision’ by the Met which is being challenged. In this Scenario one might look to challenge the overall decision to introduce and use the patrol-allocation tool (this would be closely analogous to the challenge in *Bridges*). Alternatively an individual decision to allocate one or more patrols (presumably allocations are reviewed and remade periodically) could be challenged.
46. The level of human involvement in the decisions may be relevant in deciding which is most amenable to challenge. Australian case law suggests that a *fully* automated determination cannot be a ‘decision’ for public law purposes<sup>30</sup>. However there is no similar authority in England and Wales, and it is difficult to see how this is sustainable given the increasing use of automated decision-making<sup>31</sup>.
47. A more important factor may be timing. Judicial review is ordinarily subject to strict time limits, which may make challenging a specific patrol allocation more realistic, assuming it takes time for the affected communities to become fully aware of the Met’s use of the tool and how it appears to be unfairly impacting them. We discuss possible challenges to *both* the introduction of the tool in the first place, and to individual decisions made/supported by the tool, below.

b) Unlawful delegation and/or fettering of discretion

48. Where legislation names a particular person or officeholder charged with making a decision in exercise of a public function, it will be unlawful for the officeholder to ‘delegate that decision to someone else’<sup>32</sup>. This might be thought to create a barrier to automation of decisions, particularly where a decision is ‘fully’ automated, as opposed to merely providing recommendations to decision-makers (or where recommendations from algorithmic tools are relied upon excessively rigidly)<sup>33</sup>.
49. However, it is arguable that the ‘*Carltona*’ principles – principles established regarding delegation of decision-making by ministers<sup>34</sup> – allow for the ‘devolution’ of decisions

---

<sup>30</sup> *Pintarich v Deputy Commissioner of Taxation*

<sup>31</sup> See also analysis from *Linklaters*: <https://techinsights.linklaters.com/post/102idcd/do-androids-dream-of-judicial-review-challenges-to-automated-decision-making-pro>

<sup>32</sup> *Noon v Matthews* [2014] EWHC 4330 (Admin)

<sup>33</sup> The Scenario does not state precisely how the tool’s recommendations are used and whether they are reviewed by human decision-makers. Any challenge based on unlawful delegation or fettering would require either no human oversight of patrol allocations or a very rigid following of the allocation tool’s recommendations.

<sup>34</sup> *Carltona Ltd v Commissioners of Works* [1943] 2 All ER 560 (CA)



through automation *even by named officeholders* (provided they have sufficient oversight over the decision-making tools): this being analogous to the devolution of ministerial decisions to civil servants in their departments<sup>35</sup>.

50. In this Scenario, a more fundamental difficulty is that the decision – deriving as it does from common-law powers – has not been expressed by Parliament as being required to be taken by any particular person.
51. The same issue likely prevents a challenge on the basis that fully automating patrol allocation decisions is an unlawful fettering of discretion. Whilst the issue is complex, case law<sup>36</sup> strongly suggests that the exercise of common-law powers cannot be challenged on this basis, since a public authority exercising common-law powers is lawfully able to decide whether and how to fetter its discretion in exercising them.

c) *Irrationality and proportionality*

52. Traditionally, a decision by a public authority will be unlawful where it is unreasonable in the sense of being irrational: a decision which ‘*no sensible person [...] could have arrived at*’<sup>37</sup>. There is no reason in principle why an algorithmic decision could not meet this standard, if a tool were set up – or malfunctioned – so as to come to a truly bizarre decision. However there is nothing in this Scenario that suggests the use of the patrol-allocation tool is – or its recommendations are – irrational. It may lead to outcomes which trouble the affected communities, but it is clearly open to a sensible person to base patrol allocations on past crime statistics (and likewise it is not irrational for the Met to decide to use the tool in the first place).
53. The test in *Wednesbury* – while still relevant<sup>38</sup> – has to some extent been replaced in judicial review by a broader consideration of the proportionality of a public authority’s actions or decisions, particularly where human (or other fundamental) rights are engaged.<sup>39</sup> Proportionality reviews requires a consideration of:

*(i) whether the public authority’s objective is sufficiently important to justify the limitation of a fundamental right;*

*(ii) whether the limitation of rights is rationally connected to the objective;*

*(iii) whether a less intrusive measure could have been used; and*

---

<sup>35</sup> See *Casale (2021)*

<sup>36</sup> *R (Elias) v Secretary of State for Defence* [2006] 1 WLR 3213

<sup>37</sup> *Associated Provincial Picture Houses Ltd. v Wednesbury Corporation* [1948] 1 KB 223

<sup>38</sup> *Keyu v. Foreign Secretary* [2015] 3 WLR 1665

<sup>39</sup> Taggart, *Proportionality, Deference, Wednesbury* (2008) NZ L Rev 423

(iv) whether, having regard to these matters and to the severity of the consequences, a fair balance has been struck between the rights of the individual and the interests of the community.<sup>40</sup>

54. Here, the Met has an important objective in allocating patrols effectively and efficiently. Its actions are rationally connected to that objective: it is clearly rational to consider past crime statistics in order to allocate patrols. How the ‘intrusiveness’ and ‘fair balance’ questions are answered is complicated by it being far from obvious how in this Scenario – merely by allocating patrols to an area – the Met interferes with individuals’ rights<sup>41</sup>. The awkwardness of proportionality review in such cases has been pointed out<sup>42</sup>. Suffice it to say that in this Scenario the interference with rights that *directly* results from the use of the patrol-allocation tool (if any) is slight. The Met’s actions are therefore unlikely to be in breach of the requirements of proportionality.
55. The interactions and overlap between the *Wednesbury* proportionality tests are complex and the subject of much academic debate<sup>43</sup>. Since the Met’s actions in this scenario do not appear to breach either standard, we do not explore this further.

d) (Ir-)relevant considerations

56. Public authorities have a duty to take into account relevant considerations in exercising their public functions<sup>44</sup> and, conversely, not to take into account irrelevant considerations<sup>45</sup>. A challenge in this vein would be more straightforward where legislation specifies which considerations are relevant, though the question may also be addressed objectively in the case of the exercise of a common law power as in this Scenario. Past crime statistics are a relevant consideration in this decision. However it might be argued that by focusing *only* on past crime statistics, the patrol-allocation tool leaves out relevant factors such as qualitative intelligence, an understanding that much crime goes unreported, changing social demographics and so on.

---

<sup>40</sup> Bank Mellat v HMT [2013] UKSC 38 per Lord Sumption at [20]

<sup>41</sup> Whilst accepting that *once on patrol*, officers clearly may interfere with individuals’ rights.

<sup>42</sup> Taggart (2008) and Williams, *Structuring Substantive Review* (2017) Public Law, 2017:1

<sup>43</sup> See Williams (2017), Lam, *Proportionality vs Rationality Review: A False Dichotomy?* (2021) online [<https://blogs.kcl.ac.uk/kslr/2021/07/01/1860/>] and Daly, *Wednesbury and Proportionality — Where are We Now?* (2016) online [<https://www.administrativelawmatters.com/blog/2016/11/28/wednesbury-and-proportionality-where-are-we-now/>]

<sup>44</sup> Tesco Stores v Secretary of State for the Environment [1995] 1 WLR 759.

<sup>45</sup> R v Rochdale MBC ex p Cromer Ring Mill [1982] 3 All RR 761

57. An individual patrol-allocation decision could therefore be challenged on this basis<sup>46</sup>, but this would be a relatively novel application of this principle:

*“There is no question that the ground of [relevant considerations] has the capacity to address and guide the use of [algorithmic] systems on this front, but it will need a great deal of more detailed, technical and legal development before it is able to do so.”<sup>47</sup>*

58. A challenge based on irrelevant considerations would face practical difficulties in that the communities affected do not have complete information on the statistical factors taken into account by the patrol-allocation tool. At best, they may have a high-level description from the Met about how it works, making developing an argument about the factors which it does/not consider very challenging (see further on transparency in section B.V below).

e) Duty to give reasons

59. For some public decisions, there is a duty on the public authority to provide reasons to the person(s) affected<sup>48</sup>. However, there is no general duty to do so for all public decisions<sup>49</sup>. The duty will only arise in the kinds of case where it is necessary to make effective the right to judicial review. In particular this will be the case where a decision is being made about an individual, especially where that decision is quasi-judicial and there is no public interest against the giving of reasons. This decision does not meet these criteria: it is not about an individual and it is not quasi-judicial, determining one or more individual’s rights or entitlements. Further – as shown by the PSNI FOIA case referred to above (see para 27) – there are some public interest reasons militating *against* the giving of reasons for this kind of decision. It’s therefore very unlikely that a challenge against either the introduction of the tool or an individual patrol allocation decision could be brought on this basis.

---

<sup>46</sup> The decision to introduce the tool could not be, since considerations such as saving money and increasing the efficiency of patrol allocations are clearly relevant to the decision as to whether or not to introduce a degree of automation.

<sup>47</sup> Edwards, Williams & Binns, *Legal and regulatory frameworks governing the use of automated decision making and assisted decision making by public sector bodies* (2021), accessed at <https://research.thelegaleducationfoundation.org/wp-content/uploads/2021/07/FINAL-Legal-and-Regulatory-Frameworks-Governing-the-use-of-Automated-Decision-Making-and-Assisted-Decision-Making-by-Public-Sector-Bodies-1.pdf>

<sup>48</sup> R v Secretary of State for the Home Department, Ex p Doody [1994] 1 AC 531 and Dover District Council v CPRE Kent [2017] UKSC 79

<sup>49</sup> R v Civil Service Appeal Board, ex parte Cunningham [1992] ICR 816

f) Human rights

60. A public authority's decision or act will be challengeable if it breaches the Human Rights Act 1998 (the 'HRA'). Whilst clearly there are many ways in which Met officers may breach the HRA once they are on patrol, this would be linked inextricably to the specific actions of a specific officer. The fact that the patrol was in an area as a result of the use of the patrol-allocation tool would merely be a background factor. The use of the tool in and of itself in the circumstances described in this Scenario could not constitute a breach of the HRA<sup>50</sup>.

g) Breach of policy or legitimate expectation

61. If there were a clear Met or government-wide policy against the use of algorithmic tools to determine patrol allocations (or a policy which required the decision to be made in a way which effectively excluded such tools), this Scenario would be a breach of that policy and potentially challengeable. Whilst some 'policies' have some relevance to this Scenario – such as the Police Code of Ethics,<sup>51</sup> which addresses discrimination, and the voluntary guidance referred to at para 29 above, they are far from specific enough to rule out the use of the patrol-allocation tool in the way described.
62. More significantly, not all policies are enforceable through judicial review:

*"In our judgement, public law has not reached the stage at which all administrative policies have become enforceable as a matter of law. Policies come in various forms and their content is wide-ranging. Some policies, such as those in the present context, are essentially inward facing and govern the way in which a public authority will conduct its own affairs. They do not concern the exercise of public powers."<sup>52</sup>*  
(emphasis added)

63. The policies of potential relevance to this Scenario – in particular internal public sector guidance on the use of algorithmic tools – appear to be of the type that are not susceptible to judicial review, even if they were clear enough for the use of the tool in this Scenario to be a breach.
64. Somewhat similarly, had the Met created a 'legitimate expectation' that it would *not* decide patrol allocations using an algorithmic tool, this could create a ground for

---

<sup>50</sup> Though to the extent any proceedings under the HRA were brought in relation to specific police actions, this might be a way of obtaining information through disclosure about the use of the patrol-allocation tool as part of the factual background to the HRA case.

<sup>51</sup> <https://www.college.police.uk/ethics/code-of-ethics>

<sup>52</sup> R (All the Citizens) v SoS (2022) EWHC 960 (Admin) at [102]

challenge<sup>53</sup>. However there is nothing in the Scenario to suggest statements or conduct on the Met's part that could give rise to such an expectation.

*h) Summary*

65. General public law principles could provide a route to challenge the use of the patrol-allocation tool in this Scenario, but it is not certain. The most promising avenue would be a challenge to an individual patrol allocation decision on the basis of irrelevant considerations, but even this would be a novel application of the principle and would likely face significant difficulties. This Scenario demonstrates how more generally:

*“while administrative law does in principle have grounds capable of dealing with [questions of automation], it is very likely that those grounds are not yet nuanced and sophisticated enough to be able to do so in practice.”<sup>54</sup>*

V. Practical challenges of judicial review

66. Assuming any breach of the PSED is not enforced by the EHRC (see **Error! Reference source not found.** above) then it – as well as/instead of any challenge based on general public law grounds would have to be enforced through an application for judicial review brought by an individual or group affected by the patrol-allocation decisions. Aside from making out the legal grounds of challenge, there are a number of practical difficulties in doing this:

*a) Evidencing grounds of challenge*

67. Any successful judicial review depends on evidence. In this Scenario, evidence of how the algorithmic tool works, how it influences or replaces human decision-makers, the factors it takes into account, the training data used, and the long-term effect it has on the distribution of patrols (in particular as regards to those with the protected characteristic of race). Ideally anyone challenging the use of the patrol-allocation tool would also be able to demonstrate where patrols *should* be being directed, on the basis of factors other than past crime statistics which are said to be relevant.

---

<sup>53</sup> See *R (Coughlan) v North and East Devon Health Authority* [1999] EWCA Civ 1871 and *United Policyholders Group v Attorney General of Trinidad and Tobago* [2016] UKPC 17, [2016] 1 WLR 3383. The requirements for a legitimate expectation to be enforceable are complex but not explored in detail here since it is not directly relevant to the Scenario.

<sup>54</sup> Edwards, Williams & Binns (2021)

68. Much of this evidence will be difficult to obtain. As we have seen,<sup>55</sup> the information will not necessarily be available through FOIA. There may be additional barriers in the form of:
- i. (Claimed) commercial confidentiality regarding the development of the tool (assuming it is contracted in by the Met from a private sector provider). This was an issue in *Bridges*.
  - ii. A lack of understanding by the Met themselves as to how the patrol-allocation tool makes its recommendations/decisions due to e.g. a lack of training, technical understanding etc.; and/or
  - iii. Depending on the complexity of the tool, ‘built-in’ barriers to *anyone* understanding how the tool makes its recommendations/decisions.
69. These are issues which will arise with many automated tools, illustrating a wider issue in evidencing public law challenges to their use, which has been widely discussed by academic commentators<sup>56</sup>.
70. There is no formal disclosure process in judicial review proceedings. But, to an extent, these transparency difficulties can be overcome by the duty of candour on public authorities in such proceedings, which requires them to:
- “assist the court with full and accurate explanations of all the facts relevant to the issue the court must decide.”<sup>57</sup>*
71. The duty of candour could assist those affected by the Met’s conduct once judicial review proceedings are in train. The nature of some of the potential grounds of challenge provide further assistance in that case law suggests they place a positive obligation on public authorities to apprise themselves of all relevant facts.<sup>58</sup> That is, if the Met’s conduct were challenged on the basis of the PSED, it would be no defence for them to say they did not have access to the data necessary to look into the equality concerns:

---

<sup>55</sup> See para 27

<sup>56</sup> Burrell, *How the machine ‘thinks’: Understanding opacity in machine learning algorithms* (2016) and Cobbe, *Administrative law and the machines of government: judicial review of automated public-sector decision-making* (2019) Cambridge University Press Online

<sup>57</sup> Treasury Solicitor’s Department Guidance, 2010, accessed at [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/285368/Tsol\\_discharging\\_1\\_.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/285368/Tsol_discharging_1_.pdf)

<sup>58</sup> See *Bridges* in particular.

*“the Court of Appeal [in Bridges] took a more expansive view of the PSED as mandating a public authority to mitigate algorithmic opacity rather than being itself impeded by it.”<sup>59</sup>*

72. Transparency and evidence-gathering remain a significant barrier, however, since anyone challenging the use of the patrol-allocation tool would want some solid evidence *before* launching the challenge. That is, *before* the Met has any duty of candour to explain how the tool works.

b) Permission, timing and standing

73. A judicial review is not brought in the same way as a civil claim, by simply issuing proceedings. The Administrative Court’s permission must first be sought<sup>60</sup>. The main matters which the court would consider in this Scenario<sup>61</sup> present difficulties to communities seeking to challenge the Met’s use of the tool:

- i. **Timing:** permission must be applied for within 3 months from the date of the decision<sup>62</sup>. Although this can be extended by the court<sup>63</sup> where it is just to do so, this is very much an exception and cannot be taken for granted. Given the material that needs to be prepared for the application for permission such as grounds, documentary and witness evidence, 3 months is an extremely tight timeframe, even assuming the affected communities become aware of all the relevant issues regarding the patrol-allocation tool when it is first introduced or used.
- ii. **Standing:** permission will only be granted to a person with ‘sufficient interest’ in the decision being reviewed. A person living in an area affected – especially someone with a history related to over-policing – would likely have standing. But such individuals are rarely willing to take on the burden of being the face of a judicial review. A local community or civil society group could have standing in this Scenario. But recent case law<sup>64</sup> has shown that a real, demonstrable link to the issue is required. The current chain of case law suggests that it would not be

---

<sup>59</sup> Seglias (2021)

<sup>60</sup> Administrative Court Guide 2022, particularly Section 9, accessed at [https://www.judiciary.uk/wp-content/uploads/2022/11/14.130\\_HMCTS\\_Administrative\\_Court\\_Guide\\_2022\\_FINAL\\_v06\\_WEB\\_2.pdf](https://www.judiciary.uk/wp-content/uploads/2022/11/14.130_HMCTS_Administrative_Court_Guide_2022_FINAL_v06_WEB_2.pdf)

<sup>61</sup> The question of whether there is an adequate alternative remedy also goes to the granting of permission, but it is not relevant in this Scenario.

<sup>62</sup> s.31(3) Senior Courts Act 1981, Administrative Court Guide 2022 and CPR 54.5

<sup>63</sup> Administrative Court Guide, Section 6

<sup>64</sup> R (Good Law Project and Runnymede Trust) v Secretary of State for Health [2022] EWHC 298 (Admin)

enough to be a civil society organisation with a general concern about discrimination or policing.

c) Costs

74. Any person seeking to bring a judicial review will need to fund their own legal costs. Whilst legal aid is in theory available (if the challenge is brought by an individual), eligibility and means-testing thresholds make it unrealistic for most. The individual or entity bringing a challenge will also take on ‘adverse costs risk’ – that is, the risk of being made to pay the defendant’s costs if the claim is unsuccessful<sup>65</sup>.
75. s.88 of the Criminal Justice and Courts Act 2015 allows the Court to make a ‘costs capping order’ (**CCO**) limiting this adverse costs risk where:
- “(a) the proceedings are public interest proceedings,  
(b) in the absence of the order, the applicant for judicial review would withdraw the application for judicial review or cease to participate in the proceedings, and  
(c) it would be reasonable for the applicant for judicial review to do so.”*
76. This will include a consideration of whether lawyers for the claimant are acting for free or at reduced rates, and whether the applicant is an “appropriate person to represent the interests of other persons or the public interest generally” (s.89 Criminal Justice and Courts Act 2015).
77. A CCO cannot be guaranteed, but depending on the source of legal representation used by any challenger in this Scenario, and the funds available to the challenger (i.e. assuming they are limited), it is reasonably likely that one would be granted if permission for judicial review itself were granted. This would not remove adverse costs risk entirely, only limit it to a lower level.
78. Further, costs capping cannot obviate the risk of being made to pay for the Met’s costs of responding to the application for permission,<sup>66</sup> since by definition a CCO will not be available if permission is refused.
79. All told, anyone seeking to bring a judicial review against the Met’s use of the patrol-allocation tool will – unless their lawyers are acting fully pro bono – need to account for

---

<sup>65</sup> CPR 44.2(2)(a) and R (M) v Croydon London Borough Council [2012] EWCA Civ 595

<sup>66</sup> Administrative Court Guide Section 25.4 and R (Ewing) v Office of the Deputy Prime Minister [2005] EWCA Civ 1583, [2006] 1 WLR 1260, [47]



at least £20,000 - 30,000 to cover their own costs and adverse costs risks (assuming a CCO is obtained).

d) Remedy

80. If a claim is successful, the remedy likely to be sought is for the decision to be quashed, requiring it to be remade,<sup>67</sup> this time lawfully. Damages would not be available. It is difficult to predict exactly how the Met would respond to a successful challenge – say, one based on the PSED – since it would depend on how the challenge is formulated and the court’s judgment.
81. However, if we assume that the Met wants to continue using the tool (e.g. because it has invested money in it) then it would likely either reintroduce it (if the challenge was to the use of the tool *per se*) or remake the allocation decision (if the challenge was to an individual patrol allocation made using the tool), but seeking to cure the deficiency that gave rise to the challenge. For a PSED challenge, this might involve carrying out an Equality Impact Assessment, tweaking the algorithm, or adding in greater human involvement. The latter two changes could also be made in response to a public challenge on the basis of irrelevant considerations.
82. But when curing the unlawfulness in this way, the Met would retain substantial discretion in its new decision. It is for the Met, not the courts, to determine what weight to give relevant considerations, provided they are taken account of. That is, the end result of any judicial review might lead to *some* improvement in – and transparency around – the Met’s practices, curbing the worst excesses of the over-policing tendencies of the patrol-allocating tool, but it would by no means guarantee a substantially different outcome for communities affected.

VI. Complaining to the Met and/or the IOPC

83. It would of course be open to those troubled by the Met’s use of the patrol-allocation tool to complain directly to the force. The Met does consider complaints about policing policy<sup>68</sup>. The Met would follow guidelines set by the Independent Office for Police Conduct<sup>69</sup> (‘IOPC’) in handling the complaint. This does not represent a freestanding legal route to force change to the Met’s use of the algorithmic tool, but – given the possibility of a later judicial review challenge – bringing evidence-based concerns

---

<sup>67</sup> In rare cases damages may be awarded but that would not be relevant in this Scenario since it is not possible to demonstrate that anyone has been harmed as a direct result of the unlawful decision.

<sup>68</sup> <https://www.met.police.uk/fo/feedback/complaints/complaints/>

<sup>69</sup> <https://www.policeconduct.gov.uk/complaints/guide-to-complaints-process>

regarding over-reliance on the tool and consequent over-policing could be enough to prompt a change in policy without the need for legal action.

84. Under the Police Reform Act 2002 and Policing and Crime Act 2017, the IOPC investigates complaints where complainants are not satisfied with the outcome provided by the force complained to. There is no reason in principle why the IOPC could not look into the Met's policy on the use of the patrol-allocation tool, but:
- i. The IOPC does not investigate every complaint, only '*the most serious matters, including deaths following police contact*<sup>70</sup>', making it difficult to prompt the IOPC to investigate this issue; and
  - ii. It is unclear whether anything described in the Scenario constitutes a breach of policing standards as set out in the Police Code of Ethics, which set out only relatively vague requirements regarding non-discrimination.

## VII. Scenario 1 conclusion

85. The Scenario describes the Met using the patrol-allocation tool without an Equality Impact Assessment or considering how it could be ignoring other predictors of crime which should be relevant to their decisions. This is would likely be challengeable through judicial review, most obviously on the basis that the PSED has not been complied with. An alternative would be for the EHRC to enforce the PSED, although such enforcement is not guaranteed, and the communities affected have few ways to encourage it.
86. But those affected by the Met's automation face significant barriers:
- i. PSED and public law principles are only just beginning to be applied by the EHRC and courts to the use of algorithmic tools by public bodies, making enforcement and challenges uncertain and risky;
  - ii. Bringing a judicial review is challenging, with particular difficulties in building an evidence base, and in covering costs (including adverse costs); and
  - iii. Even if successful, decisions on patrol allocations would be remitted to the Met, which would likely continue to use the tool with some modifications – curbing the worst excesses and improving transparency, but perhaps leaving communities in substantially the same position as they started in.

---

<sup>70</sup> <https://www.policeconduct.gov.uk/about-us>

#### D. Scenario 2: 'dynamic' or 'demand-led' rail pricing

*To increase overall revenue, railway companies introduce 'demand-led pricing' with the backing of the Department for Transport<sup>71</sup>. Under the system, an algorithm is used to optimise offered fares for maximum occupancy on an hourly basis, decreasing fares during quiet periods to encourage travel, and increasing fares during busy ones to shift demand to other, quieter times of the day/week. The AI system requires only data on available routes and anonymised statistics on capacity.*

*The system results in significantly increased fares on key commuter routes in and out of major cities during rush hour, as well around major public events.*

**Processing involved:** *The AI system does not process any personal data since all inputs are fully anonymised.*

**Harm:** *The system limits price transparency, making it difficult for consumers to plan and to compare modes of travel. This has the potential to lead to mistakes and overpayments, and to make less carbon-intensive travel less attractive. It may also place extra costs on those who have less choice about how and when they travel, i.e. those on lower incomes who are less likely to own a car<sup>72</sup> and those living further from work without the option of working from home (e.g. key workers).*

*There could be negative impacts on consumers during particular events such as emergencies where the need to travel suddenly surges.<sup>73</sup> The system could also conceivably malfunction, generating high prices which are not linked to demand, but whose underlying logic is difficult to challenge, e.g. due to the opacity of the algorithm<sup>74</sup>.*

---

<sup>71</sup> <https://www.telegraph.co.uk/news/2023/02/07/cost-rail-tickets-could-fluctuate-based-commuter-demand/#:~:text=In%20times%20of%20high%20demand,place%20across%20Britain's%20rail%20network.>

<sup>72</sup>

[https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/784685/future\\_of\\_mobility\\_access.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/784685/future_of_mobility_access.pdf)

<sup>73</sup> See by way of analogy Uber's application of surge pricing to taxi fares during the 2017 London Bridge attacks: <https://www.bbc.co.uk/news/newsbeat-40158459>

<sup>74</sup> See by way of analogy the Horizon Post Office scandal: <https://www.postofficehorizoninquiry.org.uk/>

I. Decision-making context

87. The institutional structure for the provision of rail services in England and Wales is complex<sup>75</sup>. It is not the purpose of this analysis to give a full account of it, but in short<sup>76</sup>:

- i. Network rail – an ‘*arm’s-length*’ public body – owns and manages the infrastructure of the railway network.
- ii. Train operating companies (‘**TOCs**’) lease (most) stations from Network Rail and run passenger services (including setting prices for those services and selling tickets). They are granted area-based franchises to do this. TOCs are not public authorities, although some are publicly owned. We assume in this scenario that the demand-led pricing<sup>77</sup> relates to tickets sold by a privately-owned TOC.
- iii. Section 28(2) of the Railways Act 1993 provides the statutory basis for fares and ticketing regulation, which is determined through a combination of regulation and the franchise agreements between the franchising authority and individual TOCs. In England and Wales, the franchising authority is the Department for Transport (‘**DfT**’). For simplicity we assume the dynamic fares in this Scenario are all within England and Wales and all within one train company’s network.

88. In exercise its public functions – including granting franchise agreements - DfT is required to:

*“make sure that rail fares are reasonable; in determining what is reasonable the franchising authority may consider the interests of rail users and potential rail users and the financial situation including the amount of funding required to operate, maintain, renew and upgrade the railway”<sup>78</sup>*

89. The TOC in this Scenario will therefore have a franchise agreement with DfT which sets out the kinds of fares which must be offered (in terms of their conditions, but not their prices). Further, some specific fares may be regulated in terms of both their conditions

---

<sup>75</sup> We do not consider the impact of any changes that could arise from the implementation of the *Williams-Shapps Plan for Rail* - [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/994603/gbr-williams-shapps-plan-for-rail.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/994603/gbr-williams-shapps-plan-for-rail.pdf) - since legislation to implement it has not yet been laid before Parliament.

<sup>76</sup> See e.g. <https://www.networkrail.co.uk/running-the-railway/> for further details.

<sup>77</sup> To some extent, rail pricing is already demand-led, in the sense that prices for many routes are set on the basis of willingness to pay. For simplicity we use the term ‘demand-led pricing’ to refer to the automated, real-time, demand-led pricing envisaged in this Scenario.

<sup>78</sup>House of Commons Library, *Rail fares, ticketing & prospects for reform* (2019) accessed at <https://researchbriefings.files.parliament.uk/documents/CBP-8552/CBP-8552.pdf>

and the maximum amount which may be charged. These constraints are set by the government (the Department for Transport) through the industry-wide Ticketing Settlement Agreement. Again, the system of regulated fares is complex and not the focus of this analysis but, in short, fare regulation focuses on commuter fares around and in/out of London, especially season tickets. The main restriction is that fares cannot increase by more than a set amount, calculated by reference to the retail price index measure of inflation.

90. Plainly, it would be unlawful for demand-led pricing to set fares that exceeded the regulated maximum. For the purposes of this analysis we assume that any dynamic fares either:
  - i. Are not regulated fares; or
  - ii. Are only varying within the bounds permitted by the Ticketing Settlement Agreement.
91. Revenue from ticket sales does not fully fund the cost of operating railways in the UK. The difference is made up by a substantial public subsidy<sup>79</sup>. One aim of demand-led pricing is likely to be to increase the overall amount of revenue collected through ticket sales. The correct balance between ticket revenue and government subsidy is a political question on which views differ widely. Therefore, and for the avoidance of doubt, we do not consider an *increase in the overall amount paid by passengers* towards the railways to be a 'harm' for this Scenario, as opposed to consumers being misled, paying more than they need to for a specific ticket, or being unfairly treated on a group basis.

## II. Judicial Review of the Department for Transport

92. The TOC in this Scenario is a private company. It is not carrying out a public function in setting fares and its decision(s) in that regard therefore cannot be judicially reviewed.
93. The DfT could be judicially reviewed in respect of its decision to permit the TOC to set fares in this way. At least some provision for demand-led pricing would likely need to be made in the TOC's franchise agreement. Alternatively, if no changes to the franchise agreement are required, the DfT could be judicially reviewed in light of its general duty to ensure that fares are 'reasonable'.

---

<sup>79</sup> Office for Rail and Road most recent statistics to March 2022: <https://dataportal.orr.gov.uk/media/2162/rail-industry-finance-uk-statistical-release-202122.pdf>

94. To avoid repetition, the key principles for judicial review are set out in relation to Scenario 1 above (section **Error! Reference source not found..Error! Reference source not found.**). This Scenario is different as the public authority – DfT – is not *itself* using any kind of automation. There are therefore fewer potential grounds for judicial review, since (for example) there will be no *technical* difficulties in the DfT articulating the reasons for its decision, or leading to it inadvertently taking into account irrelevant considerations.
95. If we assume that the DfT has carried out an Equality Impact Assessment<sup>80</sup> regarding its decision to permit automated real-time demand-led pricing, we do not consider that there are strong merits to any ground of judicial review:
- i. The decision to permit demand-led pricing cannot be described as ‘irrational’. Nor does it appear to fail the test of proportionality, since the DfT’s aim (a sustainable service) is important; pricing trains according to demand is rationally connected to it; and the interference with fundamental rights to be balanced against the DfT’s aim is slight (if any). We assume here that the dynamic pricing model works accurately overall, i.e. that it does indeed link prices to demand. The Scenario raises the prospect of malfunctions, although not indicating that these are common. If it were demonstrated that the pricing algorithm was *systematically* generating high prices unconnected to demand, but the DfT persisted in permitting (or requiring) the TOC to use it, there would be a prospect of successfully challenging that persistence, since (a) the DfT’s action would no longer be linked to its aim, and (b) DfT would arguably be failing to ensure that fares are ‘reasonable’. In the long-run this should prevent DfT from insisting on using a faulty algorithm (if for some reason it wished to), although it might well be challenging to gather sufficient evidence to prove that prices are systematically wrongly-generated.
  - ii. The DfT will have a wide ambit of discretion as to which factors to take into account in making the decision.
  - iii. The decision does not threaten to interfere with individuals’ human rights or breach any specific policy or legitimate expectation.
96. It is important to note that – as in Scenario 1 – the fact that the DfT is likely required to carry out an Equality Impact Assessment (given the considerations in section D.IV below) would have value in increasing transparency (to the extent it is published or made

---

<sup>80</sup> One which is carried out competently and therefore complies with the PSED.

available e.g. through FOIA) about how the demand-led pricing is expected to work and impact different groups, which would assist those concerned in determining whether any challenges under Equality Law could be made out.

### III. Consumer and competition law

#### a) Unfair trading regulations

97. The Consumer Protection from Unfair Trading Regulations 2008 (the ‘**UTR**’s) prohibit ‘*unfair commercial practices*’ (Reg 3) in ways that are of potential relevance to this Scenario. The (now-defunct) Office for Fair Trading published guidance on the UTRs<sup>81</sup> which has been adopted by the competition regulator the Competition and Markets Authority (‘**CMA**’). The definition of ‘commercial practice’ under the UTRs sets their scope. It is broad and the marketing and sale of tickets by the TOC in this Scenario are in scope of the UTRs:

*“any act, omission, course of conduct, representation or commercial communication (including advertising and marketing) by a trader, which is directly connected with the promotion, sale or supply of a product to or from consumers, whether occurring before, during or after a commercial transaction (if any) in relation to a product<sup>82</sup>.”* (Reg 2)

98. The UTRs define ‘unfair’ – i.e. prohibited – practices in a number of ways:

- i. A commercial practice is ‘misleading’ and therefore unfair under Reg 5(2) if:

*“[...] its overall presentation in any way deceives or is likely to deceive the average consumer in relation to any of the matters in that paragraph, even if the information is factually correct; and (b) it causes or is likely to cause the average consumer to take a transactional decision he would not have taken otherwise.”*

The Scenario states that the demand-led pricing model ‘*has the potential to lead to mistakes and overpayments*’, which, depending on its severity, could meet the definition of ‘misleading’. This would in practice be fact specific. Whilst the UTRs do not define specific practices which are misleading, the words ‘*likely to deceive the average consumer*’ suggest that any practice would need to lead to mistakes often, or in a significant percentage of transactions. By contrast, the Scenario

---

<sup>81</sup>

[https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/284442/oft1008.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/284442/oft1008.pdf)

<sup>82</sup> A product is broadly defined and includes services.

merely states that there is the '*potential*' for '*mistakes*', suggesting that consumers making unwanted transactional decisions is uncommon. Nor does the Scenario suggest that the demand-led pricing is presented in a way which is 'deceptive'. Thus whilst this does not prevent the demand-led pricing described, it does constrain the presentation of the demand-led prices so that they are not misleading overall. A similar analysis applies to Reg 6, which prohibits 'misleading omissions', so that the demand-led pricing would need to be presented to customers with full information (e.g. the fact that the price of the ticket might later change based on shifts in demand).

- ii. A commercial practice is prohibited if it is listed in Sch 1 to the UTRs. This list includes "*falsely stating that a product will only be available for a very limited time [...] in order to elicit an immediate decision*" (Sch 1 Para 7). Again, the Scenario does not directly suggest that the demand-led pricing involves any such false statements, but the UTRs constrain the way that demand-led prices are presented to users in a way that is truthful. For example it would not be lawful for the TOC to accompany very low demand-led prices for quiet routes with untrue suggestions that such low prices were likely to 'sell out' in an effort to further stimulate purchases<sup>83</sup>.

It is an offence for a trader to engage in these practices (Regs 9, 10 and 12) as well as being a civil wrong.

99. We do not consider that what is described in the Scenario breaches the UTRs, including where the system generates inaccurate prices or high prices during emergencies. The Office for Fair Trading Guidance is clear that the bar for a commercial practice to be unfair is relatively high: it "*must be unacceptable when measured against an objective standard*". Put another way: "*the Unfair Trading Regulations impose on traders a general statutory duty to behave honestly and in good faith with consumers*"<sup>84</sup>. The Scenario indicates only that the demand-led pricing introduces complexity with some potential for overpayment or error, which could not be considered 'unacceptable' or not in good faith. High prices during emergencies might be considered ethically questionable or in bad taste, but they are not commercially 'unacceptable'. Were the demand-led pricing implemented in such a way that overpayment and mistaken/unwanted purchases were

---

<sup>83</sup> The CMA has set out how such false urgency tactics are likely to breach the UTRs: <https://competitionandmarkets.blog.gov.uk/2023/03/31/urgency-and-price-reduction-claims-are-your-online-tactics-legal/>

<sup>84</sup> House of Commons Library, *Consumer protection: Unfair Trading Regulations* (2021) <https://researchbriefings.files.parliament.uk/documents/SN04678/SN04678.pdf>



*systematic* – e.g. through the information provided or the user interface –, this would suggest that the TOC was in breach of the UTRs, though this would not strictly speaking be a harm arising from the *automation* of the demand-led pricing.

b) Abuse of a dominant position

100. The Competition Act 1998 (the '**Competition Act**') and the Enterprise Act 2002 (as amended in 2013 by the Enterprise and Regulatory Reform Act 2013; the '**EA 2002**') together create a legal framework for regulating and governing competition in UK markets. Of relevance to this Scenario are the rules against the abuse of a dominant position. S.18 of the Competition Act provides:

*“(1) [...] any conduct on the part of one or more undertakings which amounts to the abuse of a dominant position in a market is prohibited if it may affect trade within the United Kingdom.*

*(2) Conduct may, in particular, constitute such an abuse if it consists in—*

*(a) directly or indirectly imposing unfair purchase or selling prices or other unfair trading conditions [...]*” (emphasis added)

101. 'Dominant position' is not further defined in the Competition Act, but the concept has been developed through regulatory decisions and case law. It requires a consideration of the market in which the undertaking operates (in this Scenario it would be the market for rail travel in the area covered by the TOC's franchise), and whether the undertaking has:

*“a position of economic strength enjoyed by an undertaking which enables it to prevent effective competition being maintained on the relevant market by affording it the power to behave to an appreciable extent independently of its competitors, customers and ultimately of its consumers.”<sup>85</sup>*

102. Whether an undertaking has a dominant position is fact specific. There is not enough information in the Scenario to say conclusively whether this TOC has a dominant position, but the structure of the UK rail market makes it very likely, since each TOC, through its franchise, provides the vast majority of services in the area covered by the franchise. Indeed, the fact that fares are highly regulated (both through DfT's granting of franchise agreements and through regulated fares) is in recognition of the fact that

---

<sup>85</sup> Case 27/76 United Brands v Commission [1978] ECR 207, [1978] 1 CMLR 429

individual TOCs can behave independently of their customers and competitors, as the provision of rail services in any one region approximates a natural monopoly.

103. Recent case law<sup>86</sup> shows that TOCs will often be in a dominant position for the purposes of s.18 Competition Act. We assume therefore that the TOC in this Scenario does have a dominant position.

104. Traditionally, abuse of a dominant position has been seen through a lens of the prices charged for goods. Guidance from the CMA<sup>87</sup> states:

*“anti-competitive conduct which exploits consumers or tends to have an exclusionary effect on competitors is likely to constitute an abuse. Examples of the type of conduct that may fall into this category for a dominant business include: charging prices so low that they do not cover the costs of the product or service sold [...]”*

105. This Scenario raises the prospect of:

- i. **Excessively high, exploitative prices:** i.e. where very busy routes’ prices rise considerably, or more rarely where prices are very high without high demand due to an error in the system. In *United Brands*<sup>88</sup> a two-stage test was set out: (i) is the price excessive? (ii) is it unfair in itself or by comparison to other products? Other EU case law<sup>89</sup> suggests that prices must be ‘*persistently*’ elevated to be excessive. English precedent suggests that pricing must be truly exceptionally excessive to qualify as abusive for the purposes of the Competition Act.<sup>90</sup> The Scenario does not state how high prices can rise under demand-led pricing, but it seems unlikely that they would be so high as to be in breach of s.18 Competition Act where the system is operating normally – i.e. where prices are linked to demand – since higher prices reflect the demand-driven value of the tickets sold. This remains the case regardless of what is driving demand – such as an emergency. Again, whilst charging high prices for rail travel during an emergency might be considered

---

<sup>86</sup> Gutmann v First, Stagecoach and London & South Eastern [2021] CAT 31 and [2022] EWCA Civ 1077 on appeal, collectively the ‘**Boundary Fare Litigation**’

<sup>87</sup> CMA Guidance 2014: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/306899/CMA19.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/306899/CMA19.pdf)

<sup>88</sup> A European precedent, though note that the Competition Act’s provisions effectively derive from EU law and mirror the wording in the relevant articles of the Treaty on the Functioning of the European Union.

<sup>89</sup> Case C-177/16 *Latvian Copyright Society*

<sup>90</sup> CMA v Pfizer and Flynn [2020] EWCA Civ 339 in which price increases of over 2,500% were found not to be abusive. See in particular Sections E, F and H.

unbecoming of a good corporate citizen, it would not be exploitative for the purposes of competition law.

Consistently high prices which are *not* linked to demand (but result from an error in the system) could well be regarded as exploitative and therefore unlawful, since the crucial connection between the value of the tickets and their price will have been broken by the error in the algorithm.

- ii. **Excessively low prices** (sometimes known as ‘destroyer’ or ‘predatory’ pricing, i.e. having an ‘exclusionary effect’): such as where very quiet routes have prices set to near free in an effort to increase occupancy on trains that would otherwise run close to empty, which could be seen as a threat to wipe out competition to the TOC’s product (e.g. bus tickets). We do not consider that very low demand-led prices in this Scenario could constitute abusive pricing, since by definition the low prices are for seats which would otherwise be empty. The marginal cost to the TOC of the product (the otherwise-empty seat) is effectively nil, meaning a very low price cannot be lower than the ‘*cost of the [...] service sold*’<sup>91</sup>.

106. The list of examples of abuse of a dominant position given in s.18 is not exhaustive<sup>92</sup>. The Boundary Fare Litigation is an important example in the context of rail pricing of how it is not only excessively high or low prices which can breach s.18. These cases relate to the failure of TOCs to make sufficiently prominent the availability of cheaper fares to destinations near, but outside of London, taking into account customers’ ability to travel to the edge of the London travel zone where they already have travelcards. That failure is said to have resulted in customers buying more expensive tickets than were necessary. The certification of the Boundary Fare Litigation (that is, the acceptance of it in principle) by the Competition Appeals Tribunal, confirmed by the Court of Appeal, shows how misleading commercial practices – similar to those covered by the UTRs – might breach s.18 Competition Act when carried out by an undertaking in a dominant position. As the CAT has noted:

---

<sup>91</sup> Cf. the CMA’s guidance and the (now-defunct) UK Competition Commission’s report *The supply of bus services in the north-east of England* 1995: [https://webarchive.nationalarchives.gov.uk/ukgwa/20120120004948mp\\_/http://www.competition-commission.org.uk//rep\\_pub/reports/1995/371bus.htm#full](https://webarchive.nationalarchives.gov.uk/ukgwa/20120120004948mp_/http://www.competition-commission.org.uk//rep_pub/reports/1995/371bus.htm#full)

This comparison is known as the ‘Areeda-Turner Test’. See e.g. Hovenkamp *Predatory Pricing under the Areeda-Turner Test* (2015) Faculty Scholarship 1825.

<sup>92</sup> Case C-6/72 Continental Can (the wording of the relevant EU legislation is substantially the same as s.18 Competition Act).

*“The law on what constitutes unfair trading conditions, in particular, is in a state of development [...] we do not regard it as an extraordinary or fanciful proposition to say that for a dominant company to operate an unfair selling system, where the availability of cheaper alternative prices for the same service is not transparent or effectively communicated to customers, may also constitute an abuse.”<sup>93</sup>*

*“A lack of transparency can be an important factor in rendering unlawful that which might otherwise be lawful”<sup>94</sup>*

c) Enforcement: the CMA and other regulators

107. The Competition Act (as amended by Enterprise and Regulatory Reform Act 2013) empowers the CMA to enforce competition law. That is, to investigate and take enforcement action in respect of the abuse of a dominant market position. A full account of the CMA’s enforcement powers is not required for this analysis. It suffices to say that they are extensive and strong<sup>95</sup>, including:

- i. Investigatory powers including in some cases entry without a warrant (§25-29 Competition Act);
- ii. The power to give directions to remedy a breach of s.18, and to have those directions enforced by the court (§33 and 34 Competition Act);
- iii. The power to levy fines of up to 10% of turnover (s.36 Competition Act).

108. The most likely breach in this Scenario would be to a systematic error in the pricing algorithm causing prices to become disconnected from demand for some routes. The biggest barrier in remedying such a breach would be evidencing the error in the algorithm, whose logic may be opaque or protected commercial property. In this regard, the CMA’s investigatory powers are significant, and could be used in response to a report from one or more individuals affected by the error<sup>96</sup>.

109. The CMA also has a responsibility to enforce breaches of the UTRs alone (that is, when the UTRs are breached but not in the context of a dominant position) pursuant to s.213

---

<sup>93</sup> Gutmann [2021] CAT 31 at [60] and [64]

<sup>94</sup> Gutmann [2022] EWCA Civ 1077 at [101]

<sup>95</sup> Note also that the CMA is widely regarded as a very active regulator: <https://www.whitecase.com/insight-alert/green-light-more-active-uk-antitrust-regulator-snapshot-overview-uk-governments>

<sup>96</sup> <https://www.gov.uk/guidance/tell-the-cma-about-a-competition-or-market-problem>

Enterprise Act 2002<sup>97</sup>. Under that section and the Consumer Rights Act 2015, local Trading Standards Services also enforce the Unfair Trading Regulations in England and Wales under the coordination of the National Trading Standards Board<sup>98</sup>. The Office for Rail and Road is also able to enforce the UTRs under s.213 Enterprise Act 2002 and has done so in the past.<sup>99</sup> These additional regulators have information-gathering powers under the Enterprise Act 2002 and Consumer Protection Act 2015 enabling them to carry out this enforcement activity.

110. The CMA in particular as the lead regulator for this area carries out a range of intelligence-gathering and market engagement work to inform its understanding of potential areas for further investigation. This is guided by its priorities, which for 2023/4 include '*People can be confident they are getting great choices and fair deals*'<sup>100</sup>, of relevance to this Scenario. Thus the CMA could investigate the TOC's conduct in this Scenario of its own accord, without the need for a complaint by an affected individual.

d) *Enforcement: the Competition Appeals Tribunal*

111. Section 12(1) of the Enterprise Act 2002 established the Competition Appeal Tribunal<sup>101</sup> the ('CAT'), a specialist court dedicated to deciding cases involving competition law. Any claim that the TOC in this Scenario had breached s.18 of the Competition Act (whether through excessive or misleading implementation of demand-led pricing) could be brought before the CAT<sup>102</sup>.
112. A full account of the CAT's role and procedures is not necessary for this analysis. Its most important feature is that collective proceedings may (subject to certain conditions) be brought before it on an 'opt-out' basis (s.47B Competition Act). That is, an individual can act as a representative of a class of persons (there is no numerical limit on the size of such a class provided the relevant legal tests for certification by the CAT are met, meaning a class could easily number in the 10s of millions) who are said to have been affected by a breach of competition law. That individual would be able to bring

---

<sup>97</sup> Although its specific enforcement powers differ and are not (yet) as strong as for enforcement of breaches of the Competition Act 1998. See <https://www.clearlyantitrustwatch.com/2022/07/cma-to-be-given-new-powers-to-enforce-consumer-law-and-impose-significant-fines/> for a discussion of the proposed changes in this area.

<sup>98</sup> <https://www.nationaltradingstandards.uk/what-we-do/>

<sup>99</sup> <https://www.orr.gov.uk/monitoring-regulation/rail/passengers/consumer-law-investigation-and-enforcement/consumer-law>

<sup>100</sup> See the CMA's current annual plan: <https://www.gov.uk/government/publications/cma-annual-plan-2023-to-2024/cma-annual-plan-2023-to-2024#areas-of-focus-for-2023-to-2024>

<sup>101</sup> <https://www.catribunal.org.uk/>

<sup>102</sup> Competition Act 1998 (as amended by the Consumer Rights Act 2015) s.47A

proceedings for the damages due to the entire class, save for any members who opt out of the proceedings.

113. The scope for opt-out proceedings overcomes major barriers to bringing civil claims in this area because there is the prospect of the recovery of a very large amount of damages, even where the loss for each individual is low, and without the cost of directly representing every individual in the class. This large potential pot of money incentivises the involvement of well-resourced third party litigation funders, who can fund the up-front legal costs of bringing proceedings, as well as the cost of purchasing after-the-event insurance to cover adverse costs risks. This financing allows claims for breaches of competition law to be brought on behalf of large numbers of individuals that would never otherwise be brought<sup>103</sup>.
114. Opt-out collective proceedings are in theory available in other areas of civil law such as through the procedure in CPR 19.8. However, the rules applied by the courts on whether they may proceed are much more restrictive than for cases brought before the CAT, particularly in relation to the requirement that every member of the class has the 'same interest'. In practice this makes bringing opt-out collective proceedings otherwise than before the CAT prohibitively expensive in most circumstances<sup>104</sup>. The availability of collective proceedings in this area makes the enforcement of rules against the abuse of a dominant position significantly more realistic. The Boundary Fare Litigation demonstrates this in the context of the sale of rail tickets.

e) Enforcement: UTR right to redress

115. Reg 3 of the Consumer Protection (Amendment) Regulations 2014 provides for a freestanding civil right of redress for individuals where a trader has engaged in a 'misleading' commercial practice (Reg 5 of the UTRs – see para 98), which could in theory apply to a ticket bought through TOC's demand-led pricing depending on how it is implemented. If the misleading practice was a 'significant factor' in the purchase, an individual buying a demand-led priced ticket could seek an unwinding of the contract, a discount, or in certain circumstances damages. Although bringing such a claim would face similar challenges as bringing a civil claim under the Equality Act (see para 25, although the adverse costs position would be significantly less challenging, since most

---

<sup>103</sup> This is a very high-level description. See e.g. <https://cms.law/en/gbr/publication/opt-out-class-actions-in-the-uk-are-we-entering-a-new-era-in-litigation> and <https://my.slaughterandmay.com/insights/briefings/uk-collective-actions-regime-where-are-we-now> as starting points for further information on opt-out proceedings in competition law.

<sup>104</sup> See e.g. *Lloyd v Google LLC* [2021] UKSC 50

claims for breaches of the UTRs would be brought on the small claims track) this provides an additional potential layer of enforcement to the UTRs.

f) Summary

116. Whilst pricing and selling rail tickets are closely regulated by consumer and competition law, these areas of law do not prevent the automation of demand-led pricing per se. There is nothing in the Scenario that suggests the TOC would be in breach of consumer or competition law simply by algorithmically determining rail prices and allowing them to fluctuate on a more real-time basis in response to demand. Even where prices rise in response to an emergency, the constraint on such pricing would be public opinion rather than consumer or competition law.
117. This conclusion does however depend on precisely *how* demand-led pricing is implemented. One can easily imagine more extreme approaches to implementation (not explicitly contemplated in the Scenario) such as very significant price variations or misleading interfaces which *would* be unlawful, most significantly as a breach of s.18 of the Competition Act.
118. Perhaps most realistically, dynamic pricing by the TOC could be challenged (or enforced against) under competition law, or by way of judicial review if it was systematically wrong, and this known to the DfT. This could become necessary if the TOC or DfT are unresponsive to evidence of the error being put to them by those affected – a common issue in the context of algorithmic systems which have a veneer of scientific objectivity.
119. In either case, there would be extensive preparatory correspondence before proceedings were issued. In the case of a report made to the CMA, there would be an investigation before any enforcement action. Presumably any significant errors in the pricing algorithm would be resolved by the TOC without the need for a claim to be heard by the courts, or for the CMA to take formal regulatory action, since it is difficult to see why the TOC would want to persist in using a faulty system.
120. Thus the law provides a range of guardrails to how the demand-led pricing may lawfully work in practice. Those guardrails are very strong: there are multiple realistic routes to enforcement, including almost uniquely in English law through opt-out collective proceedings. But whilst the guardrails are strong, they are also low: the type of conduct they protect against is more related to clear and systematic errors and commercial bad faith, rather than to the automation per se.

#### IV. Equality Act

##### a) Potential for indirect discrimination

121. Setting fares through demand-led pricing is a PCP which could in theory lead to indirect discrimination, contravening §19 and 29 of the Equality Act<sup>105</sup>. Depending on how demand-led pricing is implemented, it could:

- i. Put those with the protected characteristic of **race** at a particular disadvantage, if specific routes more likely to be used in areas with higher (or lower)-than-average ethnic minority populations became significantly more expensive;
- ii. Put those with the protected characteristic of **disability** at a particular disadvantage, if prices rise significantly for routes (or times) which they are unable to switch away from, because (for example) they cannot drive;
- iii. Put those with the protected characteristic of **sex** at a particular disadvantage, where (for example) routes more likely to be used by women because of their working patterns become significantly more expensive.

122. The CMA itself has recognised this in the context of dynamic pricing:

*“People with protected characteristics are unevenly distributed geographically. As a result, even simple policies implementing regional pricing or varying services available in different areas could potentially result in indirect discrimination.”<sup>106</sup>*

123. Socio-economic status is not a protected characteristic, so any ‘discrimination’ against those on lower incomes generally, or working in particular sectors, could not be unlawful under the Equality Act. Even where one or more people of lower socio-economic status use a route which also *happens* to be disproportionately used by those with a protected characteristic, this would not support a claim by those of lower socio-economic status unless they *also* have that protected characteristic.

124. Demonstrating a ‘particular disadvantage’ compared to the comparator group (i.e. rail passengers in the TOC’s region who do not share the protected characteristic) would

---

<sup>105</sup> It could not constitute direct discrimination since the basis on which fares are altered does not sufficiently closely proxy for any protected characteristic. Note that Sch 3 Para 34A exempts anything covered by Regulation ( EC ) No 1371/2007 (which relates to rail travel) from s.29 Equality Act. However, that regulation does not regulate *pricing* of rail tickets, so §19 and 29 remain relevant to this Scenario.

<sup>106</sup> <https://www.gov.uk/government/publications/algorithms-how-they-can-reduce-competition-and-harm-consumers/algorithms-how-they-can-reduce-competition-and-harm-consumers#fn:51>



likely be challenging. Statistical evidence would be required showing systematic effects for specific routes and groups. s.136 Equality Act provides some assistance here, providing that in any court proceedings:

*“If there are facts from which the court could decide, in the absence of any other explanation, that a person (A) contravened the provision concerned, the court must hold that the contravention occurred [unless A can show that it did not]”*

125. That is, if a claimant could adduce statistical evidence of systematic disadvantage in rail pricing for those with one or more protected characteristics, the burden of proof to show that there is no indirect discrimination would then shift to the TOC<sup>107</sup>. The Scenario does not provide enough information to conclusively say whether those with a protected characteristic are indeed put at a particular disadvantage, but it is at least possible, depending on how the demand-led pricing is implemented.

b) Objective justification for particular disadvantage

126. Even if those with protected characteristics are placed at a particular disadvantage, this may be objectively justifiable as a proportionate means for the TOC to achieve a legitimate aim (see B.I):

- i. The TOC has a legitimate aim beyond maximising revenue in smoothing out demand to run an efficient and sustainable service. Indeed the importance of smoothing demand is well-recognised as vital to the long-term viability of the railways<sup>108</sup>.
- ii. The TOC’s aim is an important one. Whether some disadvantage affecting those with protected characteristics is proportionate will be highly fact-specific, depending on (at least):
  - a) The closeness of the relationship between negative price effects (i.e. increases) and routes more likely to be used by (e.g.) women or those with the protected characteristic of race.

---

<sup>107</sup> It is not enough to merely show a difference – evidence of a disadvantage compared to a comparator is required – see *Madarassy v Nomura* [2007] EWCA Civ 33 and *Efobi v Royal Mail Group Ltd* [2021] UKSC 33

<sup>108</sup> House of Commons Library, *Rail fares, ticketing & prospects for reform* (2019) accessed at <https://researchbriefings.files.parliament.uk/documents/CBP-8552/CBP-8552.pdf>

- b) The extent of any price increases beyond typical prices prior to the introduction of demand-led pricing (which may need to be analysed on average over a period of time).
  - c) Other options or measures to ameliorate any disadvantage. For example, season tickets are likely to remain available to enable commuters to avoid being subject to demand-led pricing on a day-to-day basis, and disabled persons railcards are likely to continue to be available to reduce fares.
- iii. It would be open to the TOC to argue that other means of achieving its aim are not adequate, given the well-documented issues with existing, more binary distinctions between peak and off-peak services<sup>109</sup>.

127. If a systematic error in the system happened to affect those with a protected characteristic more than others, then it would be much more likely to be challengeable as indirect discrimination, since there could be no objective justification for the high, erroneous prices. However since any error would presumably be random, it would be highly coincidental if it happened to affect a route or station in a way which overlapped with one or more protected characteristics.

128. Perhaps most importantly, the Scenario describes an *extension* of a system of pricing which already exists. It is already more expensive to travel at peak times and on busy routes. We are not aware of any suggestion – let alone a successful challenge on the basis – that this existing practice constitutes unlawful discrimination. It seems unlikely that merely automating the process – increasing the extent and frequency of price variation – would make a qualitative difference such that the demand-led model was unlawful. Price variations under the new system would have to be both (i) extreme; a genuine step-change from the current system, and (ii) clearly disadvantaging those with protected characteristics.

c) Enforcement

129. For the reasons given above, the Scenario does not suggest that the TOC's demand-led pricing constitutes unlawful indirect discrimination. If the demand-led pricing *were* implemented in an extreme way, leading to a particular disadvantage that could not be objectively justified, this breach could be enforced by the EHRC or through a civil claim. EHRC enforcement might be more realistic here than in Scenario 3 below, as this would

---

<sup>109</sup> See e.g. <https://www.buinessstraveller.com/business-travel/2022/02/14/goodbye-to-cliff-edge-rail-fares/> for a discussion of this issue of 'cliff-edge' pricing and the associated capacity issues.

be a high-profile issue affecting many people. A civil claim would be challenging and unappealing for the reasons given in section B.III above but could be worthwhile depending on the facts.

d) Summary

130. The Scenario does not directly suggest that the TOC's demand-led pricing would constitute unlawful discrimination. Indeed, for any realistically foreseeable implementation of demand-led pricing (excluding systematic errors), it is difficult to see how it would do so, since (i) it would be difficult to evidence particular disadvantage for those with protected characteristics, and (ii) any disadvantage would be likely to be objectively justifiable.

131. The corollary of this is that the Equality Act does provide some protection against extreme implementations of demand-led pricing which involve very significant price differences affecting those with protected characteristics. Enforcement would depend on the EHRC taking action or on one or more individuals taking on the substantial risk of bringing a civil claim, and this would not address any perceived unfairness to those on lower incomes/with lower socio-economic status, since this is not a protected characteristic.

V. Scenario 2 Conclusion

132. The combination of private and public entities and the presence of effective natural monopolies make rail pricing a complex and highly regulated area. However there is nothing in the law that would render the conduct described in this Scenario directly challengeable, or that would in general prevent the automation of demand-led pricing per se, provided it works as intended. After all, in many ways this would simply be a technologically enabled extension of how rail prices are already set. However:

- i. The role of the DfT means it would likely be easier for concerned groups to understand how demand-led pricing is likely to work (through an Equality Impact Assessment);
- ii. Consumer protection, competition and Equality Law all constrain demand-led pricing from being implemented in ways that are more extreme/obviously unjust than described in the Scenario; and
- iii. Were the TOC or DfT to ignore evidence of systematic errors in prices, regulation or threatened legal action would likely lead to their being rectified.

## E. Scenario 3: content moderation

A social media company (the '**Platform**') introduces a new AI tool to automate the process of text- image- and video-content moderation on its platform. Trained on previous content removal cases globally, as well as information about sensitivities in a range of countries and languages, it pre-emptively blocks content with a sufficiently high 'risk score', which can only be reinstated (if judged not in breach of the terms of service by a human reviewer) after a lengthy appeals process which must be initiated by the poster.

Overall, the sensitivity of the system and reluctance on users' part to use the appeals process significantly reduces the quantity of not only TOS-breaching and illegal content, but also legitimate content, such as:

- LGBT content<sup>110</sup>
- Criticism of certain world leaders<sup>111</sup>
- Content critical of government policy or documenting human rights abuses which is judged to be negative in tone or 'offensive'/'distressing'.

Further, the over-sensitivity of the system is more pronounced in languages other than English<sup>112</sup>, where the social media company invests fewer resources in algorithm training and has less past content removal cases to draw on.

**Processing involved:** There will be processing of personal data involved in both the training of the AI system and the use of it to block content. However, this will be processing of the personal data of those posting the offending content. No personal data of those who might otherwise view legitimate, but blocked, content is processed, meaning their GDPR rights are not engaged.

**Harm:** Users of the platform who would otherwise benefit from seeing the legitimate, but blocked content (e.g. those seeking information and discussion about government policies they disagree with or investigating and documenting human rights abuses) have their free expression rights (i.e. including the right to receive ideas as well as impart them) interfered with by no longer being able to access it. This is worse for users looking for content in languages other than English, creating a 'two-tier' system in terms of the kind of content available to users.

---

<sup>110</sup> <https://www.theguardian.com/technology/2019/sep/26/tiktoks-local-moderation-guidelines-ban-pro-lgbt-content>

<sup>111</sup> <https://www.wsj.com/articles/when-chatbots-run-up-against-chinas-censorship-f7ee1cea>

<sup>112</sup> <https://apnews.com/article/the-facebook-papers-language-moderation-problems-392cb2d065f81980713f37384d07e61f>

I. Decision-making context

133. As a private company providing a service on contractual terms, it is for the Platform in this Scenario to decide whether and how to moderate content<sup>113</sup> on its service. We assume that (i) the Platform is a large one with a significant public profile, (ii) that it qualifies as a Category 1 regulated user-to-user service<sup>114</sup> (a ‘**Cat 1 Service**’) for the purposes of the Online Safety Bill 2023<sup>115</sup> (the ‘**OSB**’) and (iii) the OSB has come into force as an Act of Parliament.
134. Whilst platforms are not liable in English law for unlawful<sup>116</sup> content on their services unless they become specifically aware of it,<sup>117</sup> they already have strong motivations for identifying and removing a range of content, whether illegal or not:
- i. There may be significant public pressure on major platforms to proactively identify and remove content which is either illegal and/or thought to be grossly offensive (misogynistic or racist content, or that which promotes eating disorders, for example<sup>118</sup>);
  - ii. Major platforms are likely to be motivated to demonstrate to governments that they are taking proactive steps to ‘clean up’ content on their service;
  - iii. Platforms available internationally may seek to comply with more restrictive laws in other countries, and find it cheaper or easier to do this on a blanket basis (rather than merely geofencing content) in some cases; and/or
  - iv. Platforms may rationally take the view that minimising the amount of ‘abusive’ or ‘unpleasant’ content – even if it is not illegal – is likely to improve user experience overall, increasing engagement and revenues.

---

<sup>113</sup> We consider only the moderation of user-generated content; not the approach taken by the Platform to determining which content advertisers can promote on the service.

<sup>114</sup> Clauses 86-88 and Schedule 11 Online Safety Bill. The designation is by reference to both size and functionality. Threshold conditions have not been clarified but this appears designed to capture the very largest social media platforms, such as Facebook, X, and TikTok.

<sup>115</sup> We consider draft of the Bill published 19 July 2023: <https://bills.parliament.uk/publications/52368/documents/3841>

<sup>116</sup> Content might be illegal under (for example) s.127 Communications Act 2003, s.15(a) Sexual Offences Act 2003, s.1 Protection of Children Act 1978, or the Terrorism Act 2000. The list of potentially relevant statutes is long: clause 53 and Schedules 5-7 of the draft Online Safety Bill provide a good starting point for further reading.

<sup>117</sup> e-Commerce Directive 2000/31/EC

<sup>118</sup> See for example campaigns like the Molly Rose Foundation: <https://mollyrosefoundation.org/>

135. Further, the OSB creates a range of new duties<sup>119</sup> for the Platform regarding certain types of illegal content, as well as content which is not illegal but is deemed to be harmful to children. These duties (see E.IV below) are likely to significantly increase the Platform's motivation to proactively moderate user content.

## II. Terms of service

136. The Platform is in a contractual relationship with each of its users – including those affected by the reduced availability of legitimate content in this Scenario. Contractual terms are set effectively unilaterally by platforms, typically incorporating both a main 'terms of service' document<sup>120</sup> and other, more detailed 'community' standards or rules<sup>121</sup>.

137. A high-level review of contractual terms for major platforms Meta, X (formerly Twitter) and TikTok shows that they do address content moderation. However the terms focus on:

- i. What content is in breach of the terms of service;
- ii. How platforms seek to identify such content; and
- iii. The process by which they remove it, including whether and how users whose content is removed (as opposed to other users) can appeal such decisions.

138. That is, contractual terms focus on users' rights (such as they are) in relation to platform action to remove *their* content.

139. For those affected by the harm in this Scenario, platform terms of service (unsurprisingly) do not contain any undertaking to ensure the availability of any specific (type of) content, or to carry out content moderation in a way which ensures fairness between interest groups or those sharing protected characteristics<sup>122</sup>. In effect, users signing up to the Platform (let us assume it has 'typical' contractual terms) are not contracting for access to any content in particular – only for access to such content as happens to be uploaded

---

<sup>119</sup> But without creating legal *liability* for content.

<sup>120</sup> E.g. <https://m.facebook.com/legal/terms> and <https://twitter.com/en/tos>

<sup>121</sup> E.g. <https://transparency.fb.com/en-gb/policies/community-standards/?source=https%3A%2F%2Fwww.facebook.com%2Fcommunitystandards> and <https://help.twitter.com/en/rules-and-policies#twitter-rules>

<sup>122</sup> We do not consider that such a lack of a contractual entitlement could constitute an unfair term for the purposes of the Unfair Terms in Consumer Contract Regulations 1999 since it cannot be said to be 'contrary to the requirement of good faith', and relates to the definition of the main subject matter of the contract.

by others, subject to the Platform’s removal of content which it considers in breach of the terms of service<sup>123</sup>.

### III. Self-regulation

140. In the absence of existing regulation of their approach to content moderation and removal, some platforms have made high-profile moves towards a form of ‘independent’ self-regulation: most notably Meta/Facebook in the establishment of its ‘Oversight Board’<sup>124</sup>.

141. The Oversight Board is funded by Meta and established by a Charter<sup>125</sup> setting out its governance, composition, policies, and procedures. It focuses on assessing individual pieces of content (whether or not already removed by Meta) brought forward by individual users (Article 2 of the Oversight Board Charter). Pursuant to the Oversight Board Charter, Meta commits to implementing decisions of the Board.

142. The importance of freedom of expression is a strong focus for the Oversight Board and might therefore be quite receptive to complaints regarding the excessive or unequal removal of content. Indeed its purpose is stated in its Charter to be:

*“to protect free expression by making principled, independent decisions about important pieces of content and by issuing policy advisory opinions on Facebook’s content policies.”*

143. Whilst the focus on individual content decisions would appear to rule out a complaint to the Oversight Board about the *general* lack of (or uneven) availability of certain types of content, the Board does issue ‘policy advisory opinions’<sup>126</sup> which address wider issues in how Meta’s platforms are run. It is conceivable that a complaint about a specific piece of content could therefore be used to draw the broader harms in this Scenario to the Oversight Board’s attention, inviting them to issue such an advisory opinion that would address them. Such a course has the advantage of being free and not incurring any adverse costs risks.

144. Important barriers to bringing a challenge before the Oversight Board include:

---

<sup>123</sup> We do not consider whether an individual who has content removed - or their access terminated - would have a contractual claim, since this Scenario is about harm to *others* who might benefit from viewing the content.

<sup>124</sup> <https://www.oversightboard.com/>

<sup>125</sup> [https://about.fb.com/wp-content/uploads/2019/09/oversight\\_board\\_charter.pdf](https://about.fb.com/wp-content/uploads/2019/09/oversight_board_charter.pdf)

<sup>126</sup> <https://www.oversightboard.com/decision/>

- i. Many Board members and some staff have a legal background, and its decisions are judgment-like. This suggests that complainants would benefit from having advice in preparing their complaint (though it is not formally required), which would need to be paid for.
- ii. In this Scenario, the issue is that content is *not being shown* to the general population of users, who will therefore by definition face challenges in understanding that there is an issue to complain about.

145. Self-regulation has a further fatal limitation. Despite its branding, the Oversight Board is not truly independent. The trust which runs the Oversight Board is funded by Meta and has its trustees appointed by Meta (Article 5 of the Oversight Board Charter). Meta's commitment to respect the Board's decisions is voluntary and can be withdrawn at any time or on a case-by-case basis if it appears expedient to Meta to do so<sup>127</sup>. Meta is the only major platform to have set up such a model of self-regulation. For these reasons, we do not consider that this type of self-regulation offers a meaningful way for the groups affected by this Scenario to protect freedom of expression in the face of automated content moderation.

#### IV. Online Safety Bill: promoting freedom of expression?

146. The OSB creates significant new obligations for Cat 1 Services, which would be binding on the Platform in this Scenario. Some of these obligations aim explicitly to promote freedom of expression, which could help address the harms in this Scenario:

a) 'Particular regard' to freedom of expression & impact assessments

147. Clause 22 OSB obliges all regulated user-to-user services:

*"When deciding on, and implementing, safety measures and policies, a duty to have particular regard to the importance of protecting users' right to freedom of expression within the law."*

148. For Cat 1 services there is a further duty (Clause 22(4)) to carry out an impact assessment on the (potential) impact of safety measures and policies. This must be done both prospectively, and once any measures are adopted. Impact assessments

---

<sup>127</sup> See for example this recent case in which Meta ignored a clear recommendation by the Oversight Board to suspend a prominent account in relation to hate speech: <https://transparency.fb.com/en-gb/oversight/oversight-board-cases/cambodian-prime-minister-video/>



must be published and kept up to date (Clause 22(6)), including a public explanation of steps taken to protect freedom of expression (Clause 22(7)).

149. Safety measures and policies are those designed to comply with a range of other duties created by the OSB regarding the treatment of content that is illegal and/or harmful to children, and the handling of complaints.
150. The automation described in the Scenario quite likely meets the definition of a ‘*safety measure [or] policy*’, meaning that the Platform would have to have a ‘*particular regard*’ to the importance of freedom of expression. But that does not necessarily mean the Platform’s approach is in breach of the duty. Presumably (although there is of course no case law or regulatory decision on this point), as long as the Platform can show that it has *considered* freedom of expression, the *weight* to be given to the issue is for the Platform to determine as against other factors such as cost, speed, and efficiency of content removal.
151. The Platform would also be obliged to carry out and publish an impact assessment. Similarly, this does not in and of itself prevent or circumscribe the automation described. However it could conceivably incentivise the Platform to curb the worst excesses of its over- and inconsistent moderation. It would also increase transparency about the Platform’s use of information, which could help those affected bring other challenges or advocate for change extra-legally.

b) Removal of content in accordance with terms of service

152. Clause 72 OSB obliges the Platform, as a Cat 1 Service to:

*“operate the service using proportionate systems and processes designed to ensure that the provider does not— (a) take down regulated user-generated content from the service, [...] except in accordance with the terms of service.”*

153. Clauses 73(4)-(8) OSB require terms of service for Cat 1 Services to be clear and applied ‘consistently’, and to facilitate complaints about content moderation decisions.
154. The automation in this Scenario could be in breach of these provisions in a number of ways:
  - i. The fact that the content-moderation algorithm is removing content which does not *in fact* breach the terms of service puts the Platform in breach of Clause 72(1);

- ii. The automation used likely does not meet the definition of a '*proportionate system [or] process*' since it is clearly over-sensitive relative to the terms of service, breaching Clause 72(1); and
- iii. The fact that content in non-English languages is more likely to be removed strongly suggests that the terms of service are not being applied 'consistently', breaching Clause 73(4).

155. However, these provisions do not protect freedom of expression as strongly as it first appears:

- i. The duty in Clause 72 does not prevent the Platform from taking down content where that is done '*to comply with*' duties regarding content which is illegal or harmful to children (Clause 72(2)). Similarly, the duty in Clause 73(4) to apply terms '*consistently*' does not apply where those terms make provision for protecting individuals from illegal – and children from harmful as well as illegal – content. Much of the Platform's system of automated content moderation – and related terms of service – will at least *in part* be aimed at dealing with content which is illegal or harmful to children. Even if the automated processes are of questionable accuracy, the fact that they have these as their *purpose* presumably relieves the Platform of its duties under Clauses 72(1) and 73(4)<sup>128</sup>.
- ii. The duties in Clauses 72 and 73 refer only to the Platform's terms of service, rather than to any objective standard of what kind of content can be removed. The Platform can modify its terms of service at any time, and it would be open to the Platform to make them more restrictive or vaguer to ensure that the automated content moderation system matched them. That is, rather than make the automated content moderation more accurate, the Platform's terms of service could be fitted to the system, bringing the Platform's practices into compliance with clauses 72 and 73.

156. Relevant to both these limitations is the fact that the bulk of the OSB – rather than protecting freedom of expression – creates extensive new regulatory obligations for the Platform to monitor, identify and remove content. A full account of the obligations created

---

<sup>128</sup> This is a complex question of statutory interpretation which will likely require clarification in time by the regulator or the courts. But suffice it to say that it is not at all clear that Clauses 72 and 73 provide protection against the harm in this Scenario.

by the OSB is beyond the scope of this analysis<sup>129</sup>, but they include stronger obligations to identify and removal illegal content and content which breaches Platform terms of service, as well as duties to prevent children from accessing content which – despite being legal – may be harmful to them.

157. Platforms in breach of the OSB face fines of £18m or 10% of global revenue (whichever is higher; Sch 13 para 4 OSB). Content removal duties are clear and specific, making them easier for Ofcom to enforce (see below), whereas the OSB's limited freedom of expression duties are vaguer. Whilst it is difficult to predict precisely how platforms will respond to the OSB, its overall impact appears likely to create strong incentives for stricter standards of content removal, actively *encouraging* the kind of automation described in this Scenario rather than guarding against it.

c) Enforcement

158. Nothing in the OSB creates private or group rights of action against the Platform in this Scenario, even if it were in breach of freedom of expression provisions in the Bill. The OSB will be enforced by Ofcom, whose enforcement powers are set out in Part 7, Chapters 4 and 6. A full account of those powers is not necessary for this analysis, but they are extensive and include:

- i. Powers to require information and powers of entry and inspection (Clauses 101-104 and 108-109);
- ii. Powers to commission expert reports into specific issues (Clause 105);
- iii. Powers to require a platform to take specific steps to comply with the OSB (Clause 134); and
- iv. Powers to levy substantial fines for non-compliance (see para 157 above).

159. The strength and extent of Ofcom's enforcement powers cuts both ways. On the one hand, they could in theory be used to enforce freedom of expression duties to restrain the harm described in this Scenario. On the other, as noted above, they create strong incentives to remove more content more efficiently – likely through automation.

160. The OSB's new provisions are substantial in scope. Ofcom will have a duty to ensure:

---

<sup>129</sup> See <https://www.cyberleagle.com/> for a full account of the extensive new duties placed on platforms by the OSB.

*“adequate protection of citizens from harm presented by content on regulated services, through the appropriate use by providers of such services of systems and processes designed to reduce the risk of such harm.”*

161. It will however be for Ofcom to decide how to discharge that duty. In turn, there is no guarantee that there would be any enforcement against the Platform in this Scenario, even if it were clear that it was in breach of the OSB’s freedom of expression requirements by excessively and inconsistently automating content removal.

162. The communities affected in this Scenario might be able to influence Ofcom in its approach to regulation through a ‘super-complaint’, which may be made by an ‘eligible entity’ where a platform’s conduct appears to be (at risk of):

*“significantly adversely affecting the right to freedom of expression within the law of users of the services or members of the public, or of a particular group of such users or members of the public.”* (Clause 170(1)(b))

163. However:

- i. It is not yet clear which (kinds of) entities will be ‘eligible entities’, whom the communities affected in this Scenario would have to persuade to use their status to bring a super-complaint; and
- ii. Where a complaint relates to only one service (as this one against the Platform would), it is only admissible if Ofcom considers it to be ‘*of particular importance*’ or affecting ‘*a particularly large number of users*’ (Clause 170(2)).
- iii. The procedural steps involved in bringing a super-complaint and Ofcom’s duties to respond are yet to be determined (Clauses 171 and 172).

d) Summary

164. The OSB creates some freedom of expression duties for platforms like the one in this Scenario. They do not prevent the kind of automation described but might restrain some of the worst impacts. This is particularly true for the duty in Clause 22 of the OSB, which could at least improve transparency about the automated content moderation and its impact. Duties to moderate content ‘consistently’ and in accordance with terms of service are more subjective – tied to platform-determined terms of service – and may not apply at all if the Platform could show that its automated moderation is aimed at complying with illegal and harmful content provisions of the OSB.

165. Even the limited protection offered by the OSB against the harm in this Scenario depends very much on the approach to enforcement taken by Ofcom (taking into account any super-complaint by those affected if they are able to bring one). Given the balance of provisions in the OSB, we conclude that it is just as likely to encourage the kind of automated and excessive content moderation described in this Scenario, as it is to uphold minority communities' interests in accessing lawful content online on an equal basis with others<sup>130</sup>.

V. Indirect discrimination under the Equality Act

a) Is there indirect discrimination?

166. It is perhaps arguable that the Platform's inconsistent application of content moderation standards in languages other than English amounts to indirect discrimination (see **B.Error! Reference source not found.** above).

167. Someone whose native language is not English, and who is troubled by the excessive moderation of content in their language could argue:

- i. The Platform has in place a PCP: the automated approach to content removal which is more sensitive or content in languages other than English.
- ii. The PCP arguably puts those who have the protected characteristic of race (and perhaps religious belief depending on the specific kind of content being removed) at a 'particular disadvantage' compared to those who do not share the protected characteristic (native English speaking users of the service), since there is less content available on the service in their native language. This is not certain, however. We are not aware of any case law in which the lack of availability of content in this way has been shown to create a 'particular disadvantage' for indirect discrimination purposes, and it would be a somewhat novel application of the term, since the issue does not directly relate to any direct *treatment* received by the potential claimant in this example.

168. It has been argued that indirect discrimination can arise in relation to the kind of content shown to social media users.<sup>131</sup> However such arguments have focused on the

---

<sup>130</sup> It is notable that the EU's Digital Services Act (Chapter III) deals with many of these same issues. Although the legislative requirements differ, it may nonetheless be instructive to see how that regulation on content moderation, transparency etc. is enforced in practice now that the DSA has come into force.

<sup>131</sup> Global Witness Submissions to the Equality and Human Rights Commission (2021) accessed at [https://www.globalwitness.org/documents/20185/Global\\_Witness\\_EHRC\\_Submissions\\_7\\_September\\_2021.pdf](https://www.globalwitness.org/documents/20185/Global_Witness_EHRC_Submissions_7_September_2021.pdf)

presentation of *different content* to different (types of) user(s), such as the failure to present job adverts to all users, regardless of their sex. A claim in the context of this Scenario would need to go further, since the Platform presents all of the same content to all of its users, albeit that some of them have a better experience since more content is available in English.

169. Any claimant would also need to show that the PCP is *not* a proportionate means of achieving a legitimate aim – i.e. that the Platform’s automated content moderation approach cannot be objectively justified despite its disparate impact on those with protected characteristics. As in Scenario 2, this is likely to be challenging:

- i. The Platform has a legitimate aim in moderating content – including to meet its legal obligations – efficiently.
- ii. Although cost-savings alone cannot justify indirect discrimination, there is something more at play for the Platform, in that it is more technically challenging to achieve higher levels of accuracy in languages other than English<sup>132</sup>.
- iii. The aim pursued by the Platform in automating content moderation is an important one. Whether or not the PCP is *proportionate* will therefore be highly fact-specific and dependent on the extent of the disadvantage suffered by those sharing the protected characteristic. The Scenario does not state *how much more* pronounced the sensitivity is in non-English languages, but there would have to be very significant difference to begin to argue that the automated system was disproportionate to the Platform’s aims. This is especially the case since the Platform could well argue that there isn’t a realistic alternative to using the automated system, since employing human moderators in every language would be costly, logistically difficult, and slow – possibly so slow as to make compliance with legal duties under the OSB impossible. Any claimant would need to meet a very high bar to persuade a court to go behind the highly technical and commercial decisions made by the Platform in implementing the automated content moderation system.

170. Thus whilst indirect discrimination could in theory apply to this Scenario, we consider that there would be very significant evidential challenges in bringing a claim against the Platform. It would be difficult for an ordinary user of the Platform (as opposed to

---

<sup>132</sup> I.e. this is likely to be a case where the Platform can argue that it is ‘compelled’ to balance priorities, as in the case of *Heskett*.

someone who has had their content removed) to demonstrate both that the Platform's PCP puts them at a particular disadvantage, and that it is not objectively justified. Further, even if such a claim were successful, this would only deal with one aspect of the harm in this Scenario – the unequal treatment between languages – leaving the broad impact of lower content availability for everyone unaddressed.

b) Enforcement

171. If (which is rather doubtful) the Platform's approach constitutes indirect discrimination, the EHRC could take enforcement action against it (see section B.**Error! Reference source not found.**). Since this is somewhat unlikely, the alternative would be for one or more individuals to bring a civil claim under §19 and 29 of the Equality Act – for either damages or injunctive relief (i.e. a court order for the Platform to stop discriminating) or both. For the reasons given in section B.III this would be very challenging for an individual, given the speculative nature of any claim.
172. In sum, anyone seeking to use the Equality Act to address the harm in this Scenario would – in the presumed absence of enforcement action by the EHRC – face a very uncertain and risky prospect, making it unrealistic for the vast majority of those affected.

VI. Scenario 3 Conclusion

173. Users of social media platforms like the one in this Scenario are not contractually entitled to view any particular kind of content, or even necessarily to content being moderated consistently. Until the recent legislative developments such as the OSB and the EU's Digital Services Act, content moderation has been largely unregulated.
174. Despite some gestures towards the protection of freedom of expression, the OSB creates at least as much incentive *away* from that goal it does *towards* it. The provisions restraining platforms from excessively or inconsistently moderating content are relatively weak, have important exceptions, and leave platforms with significant discretion, in particular as to how they draft their terms of service.
175. While there could be a theoretical claim for indirect discrimination, it is doubtful and would likely rely on an individual taking substantial risk to bring it with no certainty of a favourable outcome. As has been noted by academic commentators:

*“Such rights [against discrimination and in relation to content removal] are [...] structurally incapable of representing all of the diffuse interests at stake: users whose content is removed might occasionally get it reinstated, but these remedies don't*

*protect the potentially millions of users prevented from seeing the content, or the broader public interest in free online discourse.*<sup>133</sup>

176. In light of the powerful and growing commercial incentives to extend and automate content moderation online, it appears that the law – even should the OSB pass – does not address the kind of automation harm described in this Scenario.

## **F. Conclusion: indirect and group-level harms are less amenable to challenge**

Taken together these Scenarios show that whilst individuals whose personal data is being processed have relatively clear rights of action in respect of harms resulting from automated decision-making<sup>134</sup>, there is less protection where those affected by automated decisions are *not* data subjects within the meaning of the GDPR:

- Scenario 2 shows the greatest level of protection, with consumer, competition and equality law forming a patchwork of guardrails which would prevent automation from being implemented in an extreme or deceptive way in a consumer pricing context (with consumer and competition law being notable for their very strong enforcement and the ability to bring opt-out collective proceedings).
- Scenario 1 shows some degree of protection although this is mainly through law requiring public authorities to go through particular steps in decision-making (such as due regard to equality law, non-delegation etc.). It may be hoped that such requirements – with or without the need for formal enforcement – would encourage public bodies to implement automation both transparently and in ways which avoid the worst excesses of automation harms. The Scenario clearly shows however that public bodies have significant latitude about the ultimate substantive decisions they make in using automation, provided they have due regard to relevant (including equality) considerations.
- Scenario 3 shows the least protection. Self-regulation offers no meaningful protection for users, whilst enforcement of discrimination claims against major platforms is rather unrealistic. Whilst the OSB attempts to address the harms envisaged by this Scenario, in fact it is more likely to increase and entrench them, given the incentives it will create

---

<sup>133</sup> Griffin, *Rethinking Rights in Social Media Governance: Why fundamental rights are not enough to remedy the injustices of contemporary social media* (2022, Sciences Po Verfassungsblog; accessed at <https://www.sciencespo.fr/public/chaire-numerique/en/2022/03/02/article-rethinking-rights-in-social-media-governance/>)

<sup>134</sup> Although that is not to say that there is complete or effective protection against such harms, even in those cases.



for platforms to expand and further automate content moderation in order to avoid breaching the bulk of the OSB's provisions.

Across the Scenarios there are common themes which highlight limitations on the role played by law and regulation in improving algorithmic decision-making, even in the relatively limited circumstances where challenges can be brought:

- Legal challenges are risky and difficult to mount. Strong evidence of problems with algorithmic systems (e.g. that the TOC's dynamic prices are erroneous in Scenario 2, or that content moderation is discriminatory in Scenario 3) is required before one can be brought. That evidence can be very challenging to obtain where algorithms have opaque logic, are protected by commercial confidentiality or are (partially) exempt from FOIA in the public sector context. Any role played by the law in moderating group-level algorithmic harms is likely to be heavily reliant on extensive investigatory work by civil society.
- Where the law provides some protection in terms of *how* automation is implemented, it does not prevent automation per se, and tends not to address the more indirect, diffuse harms that can flow from negative feedback loops which are particularly characteristic of automated decision-making systems.
- Even a successful challenge may not direct lead to changes in how an algorithm operates. Any claim would by definition be retrospective, leaving it to a significant extent up to the decision-maker how to respond to a court ruling or regulatory action. This process may be particularly complex in the public sector context where even a successful judicial review may simply leave the authority free to remake the decision to automate in the same way, but with more documented justification for doing so. That is, whilst law and regulation place some constraints on the extremes of automation harms to groups and for redress when they are transgressed, they do not give affected communities a 'voice' in how things change over time.