Title: The Online Safety Act

RPC Reference No: RPC-DSIT-4347(6)

Lead department or agency: Department for Science, Innovation and Technology

Other departments or agencies: Home Office

Impact Assessment (IA)

Date: 23/10/2024

Stage: Enactment

Source of intervention: Domestic

Type of measure: Primary

Contact for enquiries: Alex Huth
alex.huth@dsit.gov.uk

Summary: Intervention and Options

RPC Opinion: Fit for purpose

Cost of Preferred (or more likely) Option (in 2019 prices)

Total Net Present Social Value	Business Net Present Value	Net cost to business per year	Business Impact Target Status
See below	-£2,800 million	£280 million	Qualifying provision

What is the problem under consideration? Why is government action or intervention necessary?

The internet is a powerful force for good, but illegal and harmful content is widespread online. A lack of transparency, perverse incentives, and an inconsistent voluntary approach towards fighting harm online has limited the effectiveness of market solutions. Therefore, the government must act to protect users online.

What are the policy objectives of the action or intervention and the intended effects?

The policy objectives are as follows:

- **to increase user safety online**
- **to preserve and enhance freedom of expression (FoE) online**
- **to improve law enforcement's ability to tackle illegal content online**
- **to improve users' ability to keep themselves safe online**
- **to improve society's understanding of the harm landscape**

What policy options have been considered, including any alternatives to regulation? Please justify preferred option (further details in Evidence Base)

- **Option 0 - do nothing**: no clear regulatory framework to tackle illegal content and content harmful to children.
- **Option 1 - online safety framework**: a new regulatory framework which places duties on companies to improve the safety of their users online, overseen and enforced by an independent regulator.

Option 1 is the government's preferred option as it is likely to achieve reductions in online harm while maintaining a proportionate and risk-based approach.

Is this measure likely to impact international trade and investment?		Yes		
Are any of these organisations in scope?	**Micro** Yes	**Small** Yes	**Medium** Yes	**Large** Yes
What is the CO2 equivalent change in greenhouse gas emissions? (Million tonnes CO2 equivalent)		**Traded:** n/a	**Non-traded:** n/a	

Will the policy be reviewed? It will be reviewed. **If applicable, set review date:** within 5 years

I have read the Impact Assessment and I am satisfied that, given the available evidence, it represents a reasonable view of the likely costs, benefits and impact of the leading options.

Signed by senior analyst: Alex Rubin Date: 14/08/2024

Signed by the responsible minister: mJones Date: 14/08/2024

Summary: Analysis & Evidence **Policy Option 1**: online safety framework

Description: a new regulatory framework establishing duties on companies to improve the safety of their users online, overseen and enforced by an independent regulator.

FULL ECONOMIC ASSESSMENT

Price Base Year 2019	**PV Base Year** 2020	**Time period** **10**	**Net Benefit (Present Value (PV)) (£m)**		
			Low: See below	**High:** See below	**Best Estimate:** See below

COSTS (£m)	**Total Transition** (Constant Price)	Years	**Average Annual** (excl. Transition) (Constant Price	**Total Cost** (Present Value)
Low	64	1	243	**2,160**
High	179		427	**3,850**
Best Estimate	107		332	**2,970**

Description and scale of key monetised costs by 'main affected groups'

Businesses (in aggregate) are expected to incur the following transition costs (all in 10-year 2020 PV, 2019 prices): reading and understanding the regulations (£40m-£74m), ensuring they have a user reporting mechanism in place (£17m-£32m), updating terms of service (£15m-£23m), and reflecting the illegal content judgement in their internal guidance (£16m-£86m).

Businesses are expected to incur the following ongoing compliance costs: producing risk assessments (£13m-£43m), potential additional content moderation (£1,340m-£2,500m), employing age assurance systems (£18m-£92m), transparency reporting (£1m-£10m), conducting due diligence on advertisers (£53m-£187m), offering optional user verification (£9m-£14m), producing FoE and privacy IAs (£1m - £11m), and paying regulators' fees (£539m).

Government is expected to incur the following costs (all in 10-year 2020 PV, 2019 prices): justice impacts (£99m - £241m, best £170m).

Other key non-monetised costs by 'main affected groups'

The following costs to businesses have not been monetised: fines for non-compliance (out of scope), cost to internet service providers (ISPs) and payment service providers (PSPs) of business disruption measures, cost to industry and government stemming from the requirement to report online child sexual exploitation and abuse (CSEA). Where possible, this IA provides an indication of the likely scale of these impacts.

There are several indirect costs and wider impacts on society which have not been monetised, these include potential pass through from the fraudulent advertising duty, innovation impacts, competition impacts, freedom of expression implications, privacy implications, and trade impacts - these have all been thoroughly assessed qualitatively.

BENEFITS (£m)	**Total Transition** (Constant Price)	Years	**Average Annual** (excl. Transition) (Constant Price)	**Total Benefit** (Present Value)
Low				**See below**
High				**See below**
Best Estimate				**See below**

Description and scale of key monetised benefits by 'main affected groups'

Based on a subset of quantified online harm,[1] this IA estimates that this option would need to reduce online harm on an average annual basis by between 0.9%-1.7% (best 1.3%) to break even. This equates to between £251 million - £448 million (best £345 million) average annualised benefit over the appraisal period. Given the difficulties in monetising the impact of online harm, this represents a very conservative approach to benefit estimation and the break-even point is likely much lower. These potential benefits are included only for the break-even analysis and have not been included in the illustrative Net Present Social Value.

Several illustrative scenarios are also estimated to indicate how the benefit-cost ratio (BCR) would change if different illustrative assumptions were made about the effectiveness of Option 1 in reducing harm. Under a

[1] This includes contact Child Sexual Exploitation and Abuse (CSEA), modern slavery, hate crime, illegal sale of drugs online, cyberstalking, fraud facilitated by user generated content, and cyberbullying.

scenario in which harm is reduced by 3%, the benefit-cost ratio (BCR) for the Online Safety framework is estimated to be 2.3, and under a scenario in which harm is reduced by 5%, the BCR for the Online Safety framework is 3.8.

Other key non-monetised benefits by 'main affected groups'

This proposal is expected to accrue the following non-monetised benefits:

- improved efficacy of law enforcement and crime prevention for illegal content and behaviour online, expected to accrue as either cost savings or improved outcomes, for example, through minimising the creation and spread of illegal harm
- increases in levels of media literacy or the ability of users to keep themselves safe online
- an increase in the evidence base on online harms
- a reduction in the non-monetised impacts of online harms (namely, those not captured in the break-even analysis)

Key assumptions/sensitivities/risks	**Discount rate (%)**	3.5%

The key assumptions for this option are:

- the number of platforms in scope of the framework
- the risk categorisation of in-scope platforms (used as a proxy for proportionate requirements stemming from future codes of practice)
- the incremental cost of potential changes to content moderation practises
- growth rate of online harm over the appraisal period

All key assumptions are tested.

BUSINESS ASSESSMENT (Option 1)

Direct impact on business (Equivalent Annual) £m: -263		**Score for Business Impact Target (qualifying provisions only) £m:** 1,310	
Costs: **263 million**	**Benefits:** **-**	**Net:** **-263 million**	**EANDCB:** **263 million**

***Please note:** this enactment-stage impact assessment (IA) reflects changes to the legislation that have occurred over the course of the passage of the Online Safety Bill (now Online Safety Act 2023, abbreviated as OSB and OSA respectively) through Parliament.*

*To help the reader, the text highlights significant changes, particularly to the appraisal, with break-out boxes labelled "**Break-out Box # - since the final-stage IA:"** It should be possible to understand the difference between the two documents with reference to those boxes only. Less significant clarifying updates and corrections have been made inline. Figures in tables have also been updated unless otherwise noted. However, inline illustrative stats that are used in the policy rationale have not been updated (unless otherwise noted), to best reflect the rationale at the time the OSB was introduced; developments since then will be reported as part of post-implementation reviews.*

To avoid the appearance of spurious accuracy, estimates made by DSIT officials have been rounded to three significant digits.

This IA uses the original estimates of number of platforms in Categories 1, 2A and 2B. A separate IA will be produced to reflect any changes in costs when the secondary legislation that will set the definition of categories of platforms is laid.

Policy rationale

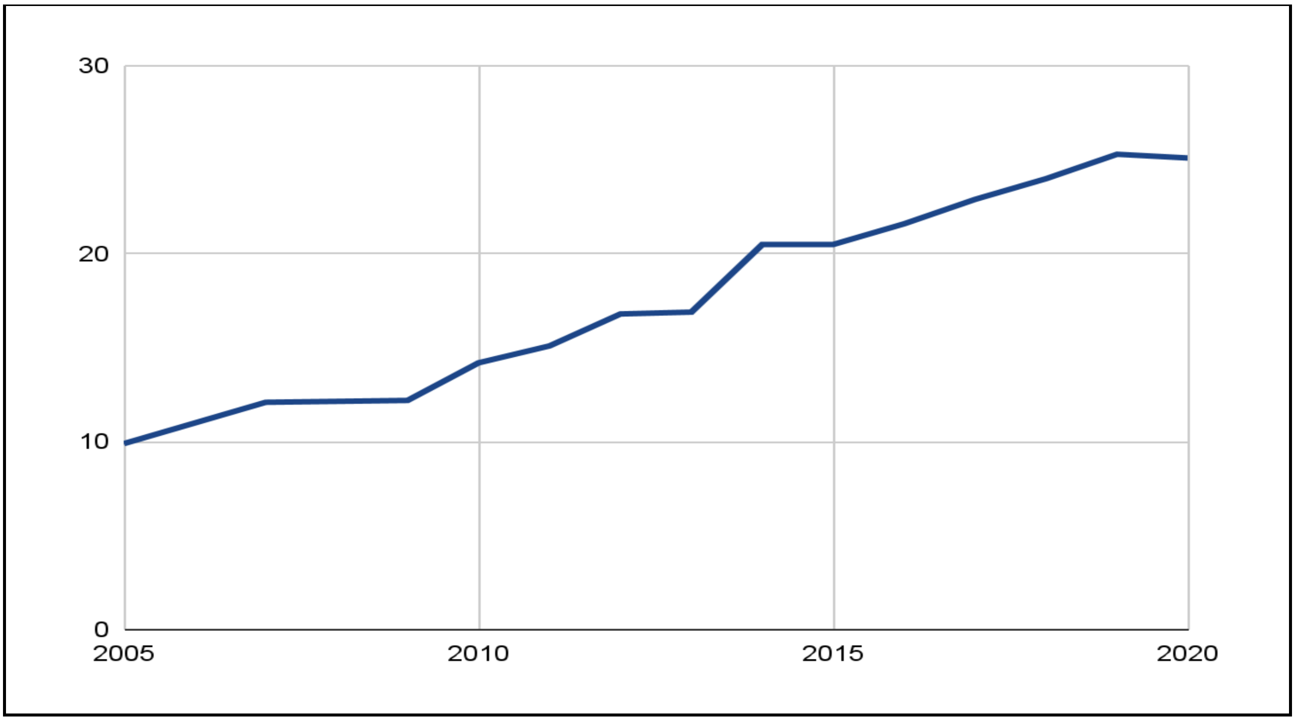
Problem under consideration

1. **Individuals in the UK are spending an increasing proportion of their time online.** This has been part of a long-term trend over the last fifteen years. Between 2005 and 2020, the weekly time spent online by UK adults has significantly increased from 9.9 hours to 25.1 hours. Being online is an integral part of everyday life for most people in the UK. For both children aged 5-15 years and adults aged 18-54 years, going online is almost universal.[2] During the pandemic, internet use among adult internet users in the UK was most pronounced in April 2020, averaging 4 hours and 2 minutes online each day. Similarly, three-quarters of British parents reported that their child's screen time averaged nine hours per day at the height of the first lockdown – nearly double the screen time prior to the outbreak.[3]

[2] Online Nation - 2021 report (Ofcom). The percentage of individuals who go online in each age group: 18-24 (97%), 25-34 (98%), 35-44 (99%) and 45-54 (92%).

[3] Study suggests lockdown could have permanently altered families' tech habits (October 2020)

Figure 1: Weekly time spent online by UK adults (hours)[4].



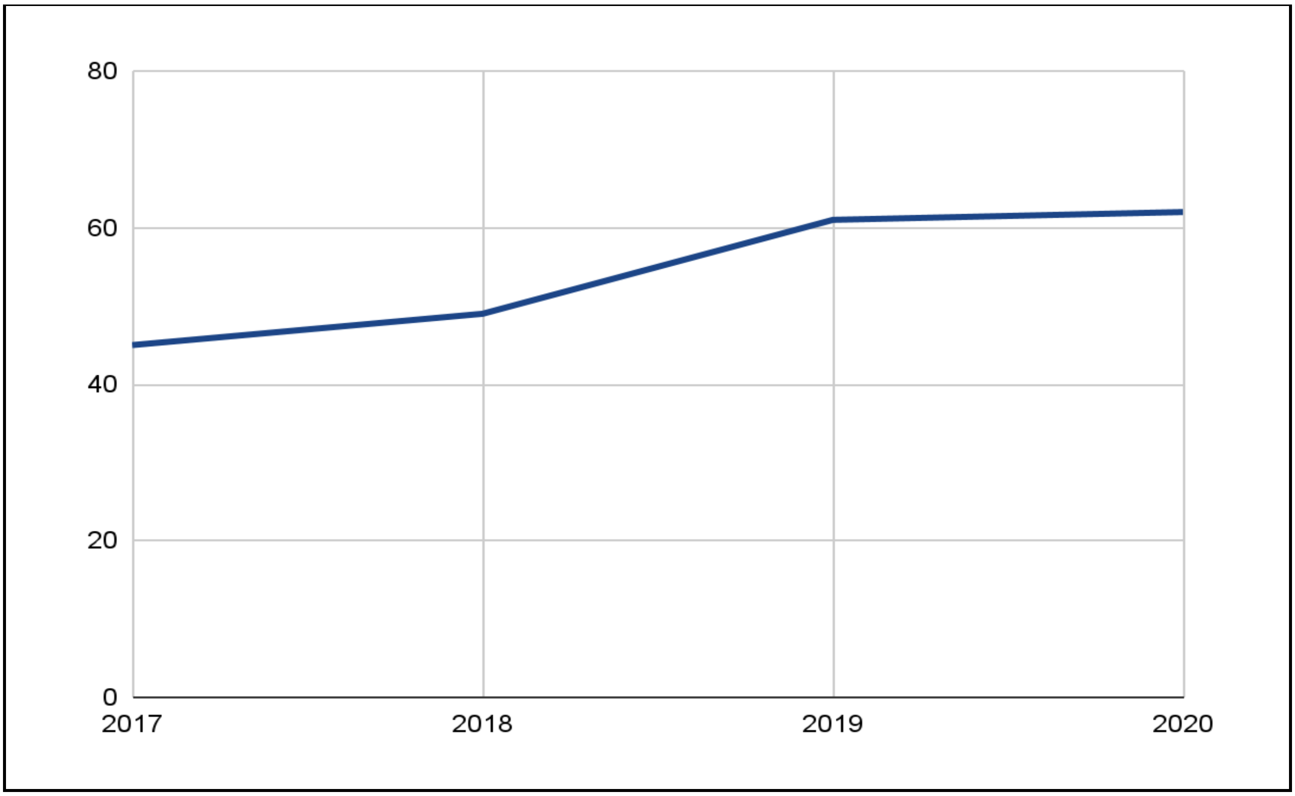
This line graph illustrates the increasing amount of time spent online each week by UK adults, from 9.9 hours in 2005 to 25.1 hours in 2020.

2. **The internet is a vital part of so many everyday activities.** Over the years, reliance upon the internet for communication, access to information, entertainment, and e-commerce has dramatically increased in the UK. The internet is a place for socialising with 92% of UK internet users going online to communicate with others and 82% of them having a social media profile, in a 2020 Ofcom report. The internet also acts as a source of entertainment with 74% of internet users watching TV content online.[5] Findings presented in the Reuters Digital News Report 2022 suggest that in the week prior, 73% of respondents (all of whom are news users) sourced news online, including social media, compared to 53% accessing it through TV.[6] However, while the internet is a powerful force for good, illegal and harmful content and activity is widespread online.

3. **UK users are becoming increasingly concerned about the content they interact with and their experiences online.** According to a 2020 Ofcom and ICO report, 62% of adult internet users have had at least one potentially harmful online experience in the last 12 months - this figure increases to over 80% for 12–15-year-olds.[7]

Figure 2: Adult internet users that have had at least one potentially harmful experience online in the past 12 months (per cent)[8].

[4] Adults' Media use and Attitudes report - Ofcom (2005-2020)
[5] Adults' Media Use and Attitudes report - Ofcom (2020/21)
[6] Reuters Institute Digital News Report 2022 - Reuters
[7] Internet users' experience of online harms - Ofcom and ICO (2020)
[8] Internet users' concerns about and experience of potential online harms - Ofcom (2017-2020)



This line graph illustrates the increasing percentage of adult internet users that have had at least one potentially harmful experience online in the past 12 months, from 45% of adults in 2017 to 65% of adults in 2020.

4. **Parents are also increasingly concerned about nearly all aspects of their child's online use.[9]** According to a 2022 Ofcom report, most parents (63%) felt that their child had a good balance between screen time and doing other things, however 40% of respondents said they struggled to control their child's screen time.[10] Ofcom's research reveals that there has been a steady decline over the past few years in the proportion of parents of online 5-15 year olds who agree that 'the benefits of the internet for my child outweigh any risks'; just over half agreed with this in 2019, compared to two-thirds in 2015.[11] [12] Whilst most parents are aware of the available parental safety controls, there is limited use of such features. For example, 91% of parents are aware of content filters via parental control software though only 73% of parents actually use them.[13] In addition, despite being under the minimum age requirement (13 for most social media sites), in 2022 33% of 5–7-year-olds and twice as many 8-11 year olds (60%) said they had a social media profile.[14]

5. **Research commissioned by 5Rights[15] claims that the current design of social media platforms does not always prioritise the safety of its users, particularly that of children.** The findings indicate that the designers' objectives focus on increasing time spent, users, and activity on the platform. This is in part done using algorithms which amplify the type of content that a profile appears to show interest in, potentially resulting in the promotion of harmful content. The report used child-aged avatars to assess which content the algorithms amplified which included sexualised images, content promoting eating disorders or weight loss and self-harm, despite the platforms

[9] Children and parents: media use and attitudes report 2020 - Ofcom

[10] Children and parents: media use and attitudes report 2022- Ofcom

[11] Children and parents: media use and attitudes report 2019 - Ofcom

[12] Children and parents: media use and attitudes report 2015 - Ofcom

[13] Children's Media Literacy Survey 2021 - Parents Data Tables - Ofcom, 2024. This figure is updated because the previous figure deviated strongly from the best current data and misrepresented parents' engagement with safety tools

[14] Children and parents: media use and attitudes report 2022 (Ofcom, 2022)

[15] 5Rights is a charity focusing on the protection of children's privacy and data online.

recognising that these accounts were registered as children.[16]

6. **The scale of illegal child sexual exploitation and abuse (CSEA) content online is significant.** In 2021, there were 29.1 million reports of suspected CSEA content[17] referred to the US National Center for Missing and Exploited Children (NCMEC), an increase of 34% from 2020.[18] Because of the large number of user-to-user platforms based in the US, who report CSEA content to the NCMEC, this figure has global significance. Reports of CSEA content online also appear to have increased through the pandemic, with the Internet Watch Foundation (IWF) recording a record number.[19]

7. **Online platforms are also used as a tool to promote extremist content.** A 2020 Ofcom report found, 3% of UK adults and 5% of children aged 12-15 have encountered material online promoting terrorism/radicalisation.[20] Evidence from a Ministry of Justice sample of extremist prisoners found that for cases prior to 2005, 83% were radicalised face-to-face with only 17% radicalised using a mixture of online and face-to-face. For cases from 2015 to 2017 this had increased dramatically to 56% radicalised using a mixture of online and offline, 27% purely online and 17% just offline, showing a significant remapping of the pathways to extremism.[21]

8. **Online fraud facilitated by UGC (user-generated content) continues to pose a major threat to UK users with large sums being lost to criminals each year.** Fraud is the UK's most common crime type: in the year ending March 2022 there were 4.5 million instances of fraud against adults in England and Wales[22], and over half of these had some online element.[23, 24] Online fraud not only has a significant financial impact on the victim but can also take an emotional toll, this is particularly relevant for romance scams online. In the year ending February 2020, the National Crime Agency (NCA) reported victim losses of over £60 million from romance fraud alone.

9. **As spend on digital advertising increases and consumers shift their purchasing online, victims of scam adverts are incurring significant financial losses.** Tackling fraudulent advertising is vital as more people see scam adverts than fraudulent UGC, with 63% having seen a scam advert and almost half seeing them at least monthly, according to a 2020 report. One in four (23%) people who have experienced a mental health problem have been victim to an online scam, three times the rate among people who have never experienced a mental health problem (8%).[25]

10. **Content and activity that is harmful to children but not illegal is also widespread online.** One study of children between 8-18 years old presenting to hospital following self-harm found that 26% of them had viewed self-harm and suicide content online.[26] Online advocacy of self-harm poses a clear threat to people's wellbeing - according to an online study of European children, 10% of children aged 11-16 years had visited pro-eating disorder sites and 5% had visited suicide sites.[27]

[16] Pathways: How digital design puts children at risk - 5Rights (2021)

[17] Reports can contain multiple pieces of content including images, videos or other files. In 2019, 16.9 million reports totalled 69.1 million pieces of suspected CSA material and other incident related content. In 2020, 21.7 million reports included 64.5 million pieces of suspected CSA material and other incident related content.

[18] By the numbers - NCMEC (2020-2021)

[19] IWF has record month as public reports of child sexual abuse surge - IWF (2020)

[20] Internet users' concerns about and experience of potential online harms - Ofcom (2020)

[21] Exploring the role of the Internet in radicalisation and offending of convicted extremists (MoJ, 2021)

[22] Crime Survey for England and Wales - ONS (year ending March 2022)

[23] Nature of crime: fraud and computer misuse - ONS (year ending March 2022)

[24] Cyber fraud represents cases where the internet or any type of online activity was related to any aspect of the offence.

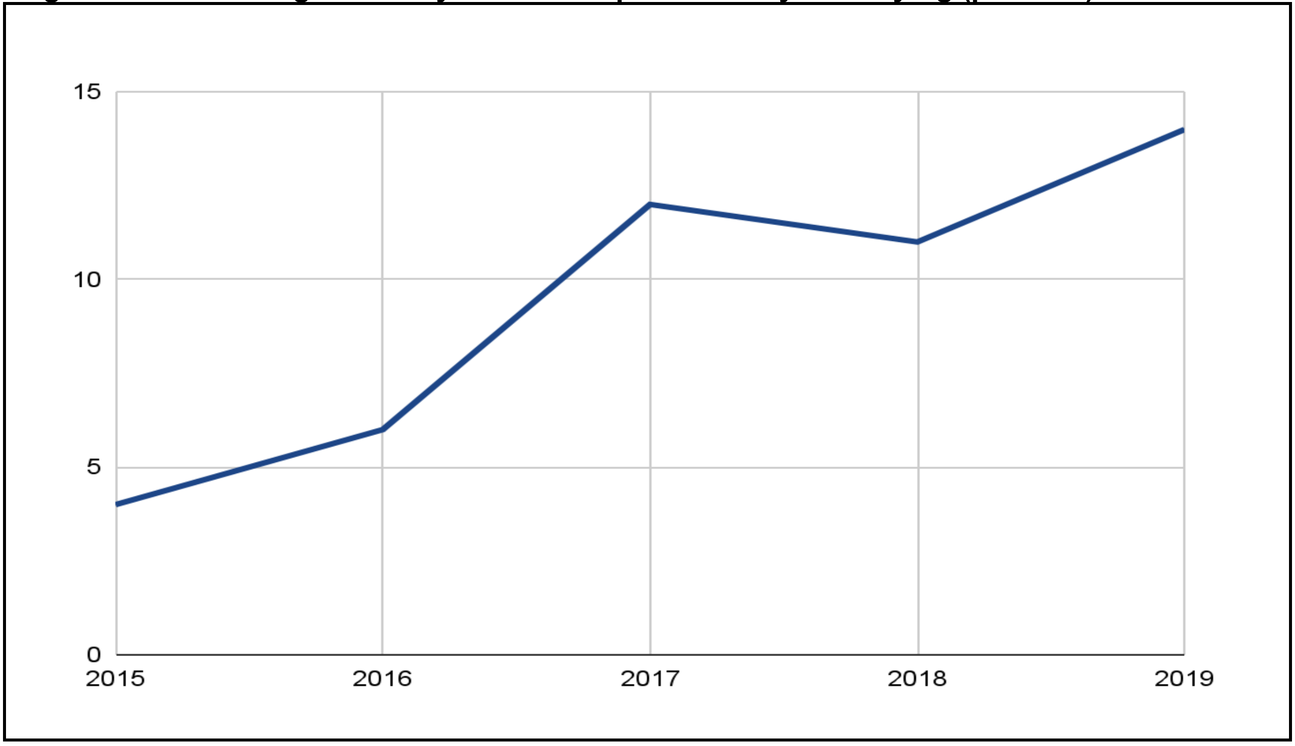
[25] Caught in the Web - Online Scams and Mental Health (Money and Mental Health Policy Institute, 2020)

[26] Suicide and Self-Harm Related Internet Use - Padmanathan et al. (2018)

[27] Risks and safety on the internet: the perspective of European children: full findings and policy implications from the EU Kids Online survey of 9-16 year olds and their parents in 25 countries - LSE (2011)

11. **Children are experiencing cyberbullying at concerning rates.** Based on figures from the Office for National Statistics (ONS), one in five children aged 10-15 years in England and Wales (19%) experienced at least one type of online bullying behaviour in the year ending March 2020. The prevalence of online bullying is significantly higher for children with a long-term illness or disability (26%) than those without (18%). 22% of children said that these incidents affected them a lot emotionally and one in four young people now have anticipatory anxiety about being abused online.[28]

Figure 3: Children aged 12-15 years that experienced cyberbullying (per cent)[29]



This line graph illustrates the increasing percentage of children aged 12-15 years that have experienced cyberbullying, from 4% of children in 2015 to 14% of children in 2019.

12. **A significant proportion of children access pornography online both inadvertently and intentionally.** Although legal age restrictions on offline pornography exist, under the status quo, children can easily access pornography across a range of online platforms. 51% of children aged 11-13 years old have seen pornography and this number is likely conservative.[30] Many children - some as young as 7 years old - stumble upon pornography online, with 61% of 11–13-year-olds describing their viewing as mostly unintentional.[31]. There are not enough safeguards to protect children from accessing pornographic material online.

13. **Online abuse is widespread online.** The nature of the internet provides a channel through which abuse and hate speech can spread anonymously and instantly. Research conducted by Ofcom in 2022 highlights that 11% of UK users have encountered hateful, offensive, or discriminatory content that targets a group or person based on specific characteristics in a four-week period.[32] The prevalence of hate speech online is particularly concerning among individuals with protected

[28] Online bullying in England and Wales - ONS (year ending March 2020)
[29] Children and Parents: media use and attitudes report - Ofcom (2015-2019)
[30] Young People and Pornography, BBFC, Revealing Reality, 2020
[31] Ibid.
[32] Online Nation 2022 Report - Ofcom

characteristics. In the first six months of 2018, 22% of reported antisemitic incidents in the UK involved social media.[33] Also in 2018, a report by Tell MAMA stated that one third of verified anti-Muslim/Islamophobic incidents were online.[34] Experiences of online abuse also vary depending on gender identity and sexual orientation. Findings of the Online Hate Crime Report 2020 indicate that 78% of LGBT+ respondents[35] had experienced online abuse, resulting in 38% of these individuals reducing the use of their online accounts and 15% left the site altogether.[36] In 2021 a survey found for 17% of girls and young women (15-18), fear of sexual harassment limits or stops them from using social media or going online.[37]

14. **Recent events have shone a spotlight on the targeting of certain public figures.** A study by the Alan Turing Institute[38] found nearly 60,000 abusive tweets directed towards Premier League players in the first half of the 2021-22 season. Around one in twelve personal attacks (8.6%) targeted a victim's protected characteristic, such as their race or gender. Over the period, 68% of players received at least one abusive tweet, and one in fourteen (7%) received abuse every day. In addition, there were increasing levels of abuse towards UK MPs during the pandemic with those from minority ethnic backgrounds and women frequently targeted during this period.[39]

15. **Many adults do not want to see certain types of content.** Polling by Ipsos shows over four in five (84%) adults in the UK are concerned about seeing harmful content such as racism, misogyny, homophobia and content that encourages self-harm.[40] According to Ofcom's Adult Media Literacy Tracker, in 2022 78% of adult users agreed that internet users must be protected from seeing inappropriate or offensive content, an increase from 65% in 2021.[41]

Rationale for intervention

Negative externalities

16. **Online harm encompasses several negative externalities, harmful content has consequences beyond that of the direct impact on the consumer, including consequences for those beyond the direct victim.** Internet users are less likely to validate online information sources,[42] and with 66% of adults in the UK using the internet as their main source of news,[43] this increases the likelihood of individuals falling victim to misleading and manipulative content. One survey found that 4.6% of people changed their mask-wearing behaviour during the COVID-19 pandemic after exposure to misinformation. The study ascribed an estimated 2,187 additional hospitalisations and 509 additional deaths in 2020 to this change in behaviour.[44]

17. **Similar negative externalities occur with exposure to content and activity that is potentially harmful but not illegal.** Secondary effects of cyberbullying include depression, self-harm and life-

[33] Adult online hate, harassment and abuse: a rapid evidence assessment - LSE (2019)

[34] Normalising Hatred - Tell MAMA, 2018

[35] Lesbian, gay, bisexual and transgender and related communities.

[36] Online Hate Crime Report 2020 - Galop, 2020

[37] Research Briefing: It Happens all the time - Girlguiding, 2021

[38] Tracking abuse on Twitter against football players in the 2021 – 22 Premier League Season - The Alan Turing Institute, 2022

[39] MP Twitter Engagement and Abuse Post-first COVID-19 Lockdown in the UK: White Paper - Farrell et al (2021)

[40] DCMS/Ipsos Polling - 12 July 2022

[41] Adults' Media Use and Attitudes Report - Ofcom, 2023

[42] Adults' Media Use and Attitudes report - Ofcom (2020/21)

[43] News Consumption in the UK: 2022 - Ofcom (2022)

[44] The Cost of Lies - London Economics (2020), modelling refers to the period between 1 April and 10 November 2020.

long impacts for the victims. An estimated 36% of victims of bullying reported feeling depressed as a result and 44% reported feeling anxious.[45] In some cases, children can develop long-term behavioural difficulties including alcohol consumption and substance abuse.[46]

18. **Pornography can result in significant short and long-term impacts on children.** Children - many who have stumbled upon pornography - feel a range of negative emotions after seeing this content, such as feeling shocked, confused, disgusted, sick, scared, and upset.[47] Exposure to this content at such a young age can negatively impact how they view their own body, for example by comparing themselves to the people featured in pornography, or seeing the people in pornography as examples of what a normal naked body looks like.[48] Several longitudinal studies have found an association between adolescents' pornography consumption and subsequent body dissatisfaction (as well as increased sexual and relational dissatisfaction).[49]

19. **Worryingly, children's exposure to pornographic content has also been shown to affect attitudes and sexual behaviour.** Evidence suggests that pornography can influence young people's sexual behaviours and expectations towards more "rough" and "forceful" sexual encounters.[50] In one longitudinal study, 10-15 year olds who consumed violent pornography were six times more likely to be sexually aggressive than those who did not consume it, or than those who consumed less aggressive pornography.[51] Also, in another study, 29% of children who intentionally access pornography did not think consent was needed if "you knew the person really fancies you", in comparison to only 5% of those who had mostly seen pornography by accident.[52]

20. **Online abuse can also influence people's willingness to speak out, fundamentally impacting on a functioning democracy**.[53] An international survey of female journalists[54] found that nearly three quarters (73%) had experienced online violence in the course of their work, and threats of physical violence, including death threats, were identified by 25%. Physical threats associated with online violence caused 13% of women survey respondents to increase their offline security, 30% said they had self-censored on social media in the face of this abuse and 20% described how they withdrew from all online interaction.

Information Asymmetry

21. **In addition to negative externalities, information asymmetries exist between users and online platforms.** While efforts have been made to increase the transparency between the two, there remains a lack of clarity among users with regards to the risk of exposure to harmful online content. This extends to information provided on platforms' websites. One survey indicates that many people do not engage with this information and those that do engage struggle to understand the information

[45] The Annual Bullying Survey 2020 - Ditch the Label

[46] The Relative Importance of Online Victimization in Understanding Depression, Delinquency, and Substance Use - Mitchell et al. (2007)

[47] Researching the Affects That Online Pornography Has on U.K. Adolescents Aged 11 to 16 (Martellozzo et al. 2020)

[48] Young People and Pornography, BBFC, Revealing Reality, 2020

[49] What is the IMPACT of pornography on young people? A RESEARCH BRIEFING for educators, PSHE, 2020

[50] Young People and Pornography, BBFC, Revealing Reality, 2020

[51] X-rated material and perpetration of sexually aggressive behavior among children and adolescents: is there a link? (Ybarra et al., 2011)

[52] Young People and Pornography, BBFC, Revealing Reality, 2020

[53] Online abuse against women MPs chilling - Amnesty International, 2020

[54] The Chilling: A global study of online violence against women journalists (International Centre for Journalists, 2022)

provided.[55] 36% of young people aged 12-18 never read the terms and conditions when signing up to new apps or social media networks, while 20% sometimes do and 14% have done once.[56]

22. **The gap between companies' terms of service and what they do in practice creates uncertainty for users and reduces trust in the relevant companies.** Where companies do have terms of service in place, concerns have been raised that these policies are not always properly enforced, for example policies about prohibiting abuse and other harmful content. The Lords Digital and Communications Committee report on freedom of expression online[57] provided evidence of content and user accounts being mistakenly removed and then reinstated. In 2022, Meta identified that its original decision was incorrect in 32 of the 50 cases shortlisted for review by the Oversight Board.[58]

23. **Children's attitudes towards the privacy of their online profiles highlights their lack of understanding of potential exposure to online harm**. Over half (53%) of children aged 12-17 are aware of settings in social media so that fewer people could see their profile, however less than one third (30%) use this feature. Similarly, when asked which measures they used to protect themselves online, more than a third of children (35%) reported using measures which might in fact have put them more at risk, because they could enable them to come across potentially harmful content.[59] It is therefore difficult for users to make an informed decision as to how they use online platforms and what content they access.

Government Intervention

24. **Online platforms have failed to effectively address online harm and ensure the safety of their users.** In the absence of regulations, harmful online content is addressed on a voluntary basis. Measures taken to improve the safety of online platforms can be delayed and reactive, resulting from governmental or societal pressure. There remains significant work to be done by social media platforms to ensure they have robust policies and tools in place to tackle the increase in harmful content which circulates at crisis moments, such as disinformation during the COVID pandemic and following Russia's invasion of Ukraine.

25. **The need for regulation is recognised by the sector itself.** Nearly half of tech industry workers (45%) believe that the industry is currently under-regulated and only 2% see voluntary commitment as the most effective way of mitigating potential harm.[60] In 2021, 65% of UK adult internet users believe that individuals must be protected from seeing inappropriate or offensive content online, an increase from 61% in 2020.[61] Fewer than one in four UK users (23%) believe that the current level of online safety measures is sufficient.[62] There is international recognition for the need to regulate online platforms (see international context section below). One study has found that 60% of US consumers support more government regulation of platforms.[63] The findings also suggest that there are significant concerns about the practises of online platforms and the power that the larger platforms hold.[64]

[55] Understanding how platforms with video sharing capabilities protect users from harmful content online - Ernst & Young (2021)
[56] The Wireless Report 2021 - Ditch the Label, 2021
[57] Free for all? Freedom of expression in the digital age - Lords Digital and Communications Committee (2021)
[58] Annual Report - Oversight Board, 2022
[59] Children and parents: media use and attitudes report 2022 - Ofcom, 2022
[60] People, Power and Technology: The Tech Workers' View - DotEveryone (May 2019)
[61] Online Nation 2022 report - Ofcom 2021
[62] Ibid.
[63] Platform Perceptions, Consumer Attitudes on Competition and Fairness in Online Platforms - CR Consumer Reports, Digital Lab (2020)
[64] Ibid.

26. **Absent regulation, there is the potential for a trade-off between encouraging traffic to a site and ensuring the safety of all users.** For example, there is a potential economic incentive for platforms not to address content such as fake news. Research suggests that false news is 70% more likely to be re-tweeted than real news.[65] The high levels of interaction with fake news, including the anti-vaxx rhetoric, generates a higher profit for platforms. Between July and August 2020, interactions on posts criticising COVID-19 vaccines on six UK Facebook pages increased by 350%.[66] Removing this content could therefore result in a short-term loss of profit and reduced user engagement.

27. **Some platforms will face incentives to address harmful content to maintain advertising revenue.** Online advertising has seen consistent growth and is forecast to account for nearly three quarters (74.3%) of all advertising spend in 2022.[67] Advertising has become the primary source of revenue for many online businesses and underpins the provision of online services such as search and social media.[68] Major platforms may have an incentive to reduce harmful content to maintain their advertising revenue but the effect is likely to be muted because advertisers have an inelastic demand for social media advertisements on the largest platforms. This is a result of smaller platforms being unable to offer advertisers such a large and engaged user base that is provided by the more popular social media platforms. Advertisers rely on the popularity of online platforms with young consumers. The findings from Ofcom's Children's Media Lives research showed that an increasing proportion of the content children consume was produced by either companies or organisations, or people who were monetarily benefiting from their content[69] and in 2021 33% of 10–17-year-olds said they trusted influencers.[70] The main social media platforms also provide a unique method of marketing, namely user-generated content (UGC)[71] by influencers. UGC has been shown to have a significantly stronger impact than marketing generated content on consumer behaviour[72] and there are a limited number of platforms through which this form of marketing can take place. Therefore, given the limited options available, advertisers are unlikely to migrate away from platforms should they not address harmful content.

28. **There are few or insufficient legal incentives for firms to take a comprehensive approach to protecting their users from harm and to protect their rights.** Although an individual could bring a claim against an internet platform to seek redress, the government is not aware of any cases having been brought on contractual or negligence grounds (whether successful or otherwise). This likely reflects the challenges of bringing such claims and the inevitable costs involved in legal action. On top of this, the existing legal frameworks relating to online content are primarily focussed on the intermediaries' liability for individual items of content, and do not give providers duties to implement comprehensive safety and freedom of expression-related systems and processes, designs and services to protect users, and children specifically

29. **A clear, proportionate and predictable regulatory framework will encourage businesses to start up, grow and invest.** Many other countries are also planning to introduce online regulation. By acting we will be able to provide certainty to platforms. There is an opportunity to set global standards, unlock investment and influence the global approach.

30. **Existing UK laws concerned with communication have not kept up to date with changes in**

[65] The spread of true and false news online - Vosoughi et al. (2018)

[66] Online Nation 2021 report - Ofcom (2021)

[67] IAB analysis of Advertising Association/WARC Expenditure Report - IAB, 2022

[68] Online Nation Report 2021 - Ofcom, 2021

[69] Children's Media Lives 2022 Summary Report - Ofcom, 2022

[70] Children and parents: media use and attitudes report 2022 - Ofcom, 2022

[71] Content, such as images, videos, text, and audio, that has been posted by users on online platforms.

[72] Social Media Brand Community and Consumer Behavior: Quantifying the Relative Impact of User- and Marketer-Generated Content - Goh et al. (2013)

the way we communicate. When the Law Commission first published a report[73] in 1985 on laws relating to communications - which led to the creation of the Malicious Communications Act 1988 - the internet had not been invented. In 2020, over 1.15 million WhatsApp messages are sent per second and today nearly 500 million tweets are posted every day[74]. Criminal law therefore needs updating for these changes, and so in a 2021 report[75] the Law Commission recommended a suite of new offences to either create new laws or update existing legislation, including offences relating to false and threatening communications. The Law Commission has also identified areas where the current law is insufficient to protect people with epilepsy from the harm that viewing flashing images can cause them; people who may be susceptible to deliberate encouragement to cause themselves serious harm; and people whose intimate images are shared without their consent from the distress and humiliation that the sharing of such images can cause them. Strengthening the law in these areas will provide greater public protection, increase public confidence in the criminal justice system, deter crime, and ensure that offenders can be punished appropriately.

31. The Digital Regulation Cooperation Forum (DRCF) brings together 4 UK regulators – Ofcom, the Financial Conduct Authority (FCA), the Information Commissioner's Office (ICO) and the Competition and Markets Authority (CMA). The DRCF supports cooperation and coordination between members and enables coherent, informed and responsive regulation of the UK digital economy. This effective collaboration greatly benefits UK citizens and consumers online. In their 2024/25 workplan,[76] it was noted that on fraud their planned activities included:

    - Continued engagement between Ofcom and FCA as Ofcom develops its code of practice on fraudulent advertising.
    - A workshop following publication of Ofcom's Call for Evidence on duties impacting Categorised Services.
    - Exploring interaction of this work with other regulator interventions, including ICO work to scope what materials could support data sharing for scams and fraud prevention.

32. DRCF's 2025/26 workplan is currently being consulted on.

Wider international and regulatory context

Domestic context

33. **e-Commerce Directive**: In the UK, platforms that host illegal user-generated content cannot be held legally responsible or 'liable' for any of this type of content that they host, unless they have knowledge of its existence and do not expeditiously remove the material. This creates a 'notice and take-down' regime, which is known as the UK's intermediary liability (IL) regime. T. These protections for platforms derive from the EU's eCommerce Directive (2000/31) and were initially implemented into UK law by the Electronic Commerce (EC Directive) Regulations 2002. Further implementing regulations were made by the government in 2018.

34. **Audiovisual Media Services Directive (AVMSD)**: In 2010, AVMSD expanded in scope to include Video on Demand services (such as Netflix). UK video on demand services (such as TikTok, OnlyFans or Vimeo) are required to take proportionate measures to ensure children are not normally able to access pornographic content. AVMSD 2020 (Directive (EU) 2018/1808) introduced rules for video sharing platforms (VSPs) for the first time. In 2021, the then government announced

[73] Law Commission, Poison-Pen Letters (1985) Law Com No 147.

[74] New Tweets per second record, and how! X Engineering (2021)

[75] Modernising Communications Offences - Law Commission (2021)

[76] DRCF 2024/25 work plan

Ofcom as the national regulator for UK-established VSPs. The UK transposed the revised Directive through the AVMSD 2020 regulations which came into force on the 1st of November 2020. The revised AVMSD 2020 regulations place requirements on UK-established VSPs to protect all users from illegal content through taking appropriate measures. UK-established VSPs are also required to take measures to protect minors from harmful content. The regulations share broadly similar objectives to the OSA and will be superseded once the latter comes into force.

35. **e-Privacy Directive**: The ePrivacy Directive (Directive 2002/58/EC) was agreed at EU level in 2002, and transposed in the UK as the Privacy and Electronic Communications Regulations 2003 (SI 2003/2426) (PECR). The Directive, which has been amended several times since, aims to protect the privacy of electronic communications, reduce the incidence of nuisance calls, and restrict website and app developers' use of 'cookies' to track user activity.

International context

36. Many countries are considering how to make the internet safer for users and some governments are acting by introducing legislative measures to tackle harmful online content. Internet safety is also being discussed in a range of multilateral and multi-stakeholder fora. The government is working closely with many international partners to address this shared challenge in order to build consensus around shared approaches to internet safety and to learn from other nations' experiences of tackling online harm.

37. **Ireland**: Ireland passed their Online Safety and Media Regulation Act 2022 in December 2022. There are three main features of the Act:
    1. it establishes a new multi-person Media Commission (Coimisiún na Meán) to regulate broadcasting and online services. This replaces the existing regulator, the Broadcasting Authority of Ireland
    2. it updates the law on how broadcasting services and video-on-demand services are regulated
    3. it creates a new regulatory framework for online safety to tackle the spread of harmful online content. This will be overseen by an Online Safety Commissioner as part of the wider Media Commission

38. **Germany**: The German Act to Improve Enforcement of the Law in Social Networks (NetzDG), which came into full force in January 2018, requires social media platforms with more than 2 million registered users in Germany to remove 'manifestly unlawful' content within 24 hours of receiving a notification or complaint, and remove all other 'unlawful' content within seven days of notification or risk receiving a fine of up to €50 million. However, because the EU's Digital Services Act (DSA) will supersede several national laws about online safety and prevent any national legislature maintaining or adopting more stringent laws in the future unless provided for in the DSA, Germany's NetzDG will need to be repealed once the DSA is implemented.

39. **France:** As with other EU Member States, France must implement the EU's Digital Services Act. To support this, France has put in place its Digital Regulation Act (SREN Act), which came into force in May 2024. This Act designates ARCOM (France's audio-visual and communication regulator) to enforce the DSA. The Act gives ARCOM powers to create binding recommendations for digital services, including technical requirements for age verification systems on sites offering pornographic content. ARCOM can block sites that fail to comply with age verification obligations, without needing clearance from a judge, and instruct search engines to remove non-compliant sites from their lists. The SREN Act also requires social media to remove child pornography within 24 hours, as is already the case for terrorist content. Failure to comply would incur a one-year prison

sentence and a €250,000 fine, rising to 4% of global turnover. In July 2023, France also enacted regulation to strengthen parental control over minors' access to the internet. The regulation primarily applies to manufacturers of devices that enable minors to access online services and content. Such devices must put in place easily accessible parental controls that can prevent the download of material that minors are prohibited from viewing.

40. **European Union:** The EU's Digital Services Act (DSA) came into force in November 2022 and has been fully operational since February 2024, The DSA sets out new requirements on a wide variety of services, with more stringent obligations on Very Large Online Platforms and Very Large Search Engines (VLOPs and VLOSEs - defined as services with 45 million or more EU users). The DSA puts in place obligations on all services in scope to tackle illegal content, and some harmful content, particularly if a service is accessed by children. Services designated as VLOPS and VLOSEs have additional obligations, including the requirement to carry out risk assessments and put in place mitigating measures for risks other than those related to the presence of illegal and some harmful content.

41. **Australia**: Australia's Online Safety Act, which came into force in January 2022, aims to promote the online safety of Australians, and grants enhanced powers to the eSafety Commissioner (Australia's online content regulator) to administer complaints related to cyber bullying of children, serious online abuse of adults, and to order the take down of harmful online content. The Act contains a set of core online safety expectations for social media services, relevant electronic services and designated internet services, clearly stating community expectations, with mandatory reporting requirements. It also includes new abhorrent violent material blocking arrangements that allow the eSafety Commissioner to respond rapidly to an online crisis event by requesting internet service providers (ISPs) block access to sites hosting seriously harmful content.

42. **Multilateral collaboration:** In April 2021, the G7 agreed on a set of Internet Safety Principles. This is significant as it is the first time that an approach to internet safety has been agreed in the G7. The principles are broad in scope, allowing for both regulatory and non-regulatory approaches to increasing internet safety. The agreed text includes four underpinning principles that will inform approaches as well as four operational principles on safety technology, media literacy, child protection and youth participation where consensus for concrete action has existed. In addition to this, in March 2020, in collaboration with the Five Country Ministerial representatives of the US, Canada, Australia and New Zealand, the UK formally launched the Voluntary Principles to Counter Online Child Sexual Exploitation and Abuse. These range from pledges to stop existing and new child sexual abuse material appearing on platforms, taking steps to stop the livestreaming of abuse, identify and stop grooming and predatory behaviour.

Policy objectives

43. The policy objectives are:

- **to increase user safety online**: this will be achieved through reduced risk and incidence of specific online harms, especially with respect to children
- **to preserve and enhance freedom of speech online:** this will be achieved through strong safeguards in the OSA to ensure that the proposals do not result in 'over-blocking' and unjustified content removal, and also through the imposition of specific freedom of expression and transparency-related duties on the largest social media companies, in light of the influence these services can wield over public discourse
- **to improve law enforcement's ability to tackle illegal content online**: this will be achieved through both a general reduction in illegal harm online and by making it easier for law enforcement to tackle identified illegal harm through updating the criminal law

- **to improve users' ability to keep themselves safe online:** this will be achieved through expanded media literacy duties for Ofcom, duties for platforms to provide user empowerment tools, greater platform transparency, and a combination of non-regulatory support measures which focus on empowering users, such as government-led media literacy initiatives
- **to improve society's understanding of the online harm landscape**: this will be achieved through enhancing the amount and quality of information in relation to online harm that is available to government, industry and civil society

44. These objectives will form the basis of monitoring and future evaluation.

Options considered

Summary of options

45. This impact assessment (IA) considers only one regulatory option in addition to a *do nothing* baseline; however, a range of options were considered as part of the consultation[77] and earlier policy development.

Option 0 - do nothing: The *do nothing* option would mean no clear regulatory framework to tackle illegal content and content harmful to children

Option 1 - online safety framework: This option introduces a new regulatory framework establishing duties on companies to improve the safety of their users online, overseen and enforced by an independent regulator. Duties are set out in primary legislation with details of how in-scope companies can fulfil their duties set in codes of practice.

46. **Option 1 is the government's preferred option** as it is likely to achieve reductions in online harm while maintaining a proportionate and risk-based approach. It will also aim to deliver a vibrant and competitive digital economy with high levels of user trust and confidence. Doing nothing has not provided sufficient incentive for platforms to reduce online harm.

Justification for the preferred option

Consideration of additional options over the development of the Bill and the parliamentary passage

47. As part of the early policy development process, the then government considered a range of options to address the problem of online harm. In addition to considering a wide range of options throughout policy development, the then government assessed three distinct policy options against a *do nothing* counterfactual in its published consultation stage IA. These differed on how content that was legal but may cause harm to adults was addressed, as well as the role of guidance and codes of practice compared to detail in the legislation. The options were:

- **limited risk based scope**: safety duties for UGC and activity addressing illegal harm, and safeguarding children from both illegal and harmful content and activity. Duties are set out in primary legislation and guidelines or codes of practice.
- **full risk based scope**: safety duties for UGC and activity addressing both illegal and legal but harmful content, and safeguarding children from illegal and harmful content. Duties are

[77] Consultation outcome - Online Harms White Paper (DCMS & Home Office, 2019)

set out in primary legislation (and subsequent secondary) and guidelines or codes of practice.

- **uniformly applied safety duties**: Detailed safety duties setting out organisations' responsibilities in addressing illegal harm and legal but harmful content, and the safeguarding of children from both illegal and legal but harmful content. These safety duties are detailed in primary legislation (and subsequent secondary) and are uniformly applied across all categories of harm and organisations in scope.

48. The consultation stage IA provided an indication of the likely scale of impacts stemming from the three options and estimated an illustrative break-even point for each:

Table 1: Estimates presented in the consultation stage IA

	Limited risk based scope	Full risk based scope (preferred)	Uniformly applied safety duties
Net present value	-£1,690m	-£2,120m	-£7,360m
Equivalent annual net direct cost to business (EANDCB)	£156m	£206m	£814m
Percentage reduction in [quantified] harms required to break even	3.1%	3.9%	13.5%

49. **Limited risk based scope** was discounted as it did not address the significant problem of legal but harmful content accessed by adults. While the break-even point for a **limited risk-based scope** was lower than the full risk-based scope, this was mainly due to the difficulty in quantifying legal harm in the baseline. There are many harms that fit into the legal but harmful category, and for which preliminary evidence suggests are significantly prevalent in the UK to have potentially large costs but are yet to be quantified.[78] This is largely due to many of these harms emerging relatively recently, and so data and evidence on the impact of these harms remains sparse. These include many forms of online abuse, disinformation, content related to self-harm and suicide, children's access to pornographic content, and advocation of risky and dangerous behaviour.

50. The option assessing **uniformly applied safety duties** was discounted as it was expected to result in significant impacts on business while not reflecting the variance of harm on different platforms and failing the proportionality test. **A full risk-based scope** was the then government's preferred option in the consultation stage IA and remained the preferred option at final stage. A full description of the three options previously assessed can be found at page 21 of the consultation stage IA.[79]

51. The final stage IA appraised the OSB as introduced to Parliament. The results of that appraisal are reported in the table below.

Table 2: Estimates presented in the final stage IA

	Best Estimate	Low Estimate	High Estimate
Net present value	-£2,510m	-£1,790m	-£3,290m
Equivalent annual net direct cost to business (EANDCB)	£251	£174m	£323m
Percentage reduction in [quantified] harms required to break even	2.1%	1.5%	2.7%

52. The OSB was introduced based on full risk-based scope, however, during passage through Parliament the then government moved to a focus on user empowerment tools as an approach to address legal content that adults may find harmful. Other changes were made to Option 1 in

[78] Indicative figures for prevalence of particular types of harm, including some of the legal but harmful harms listed can be found in Ofcom's

[79] The Online Safety Bill - impact assessment (HMG, 2020)

response to parliamentary scrutiny and extensive engagement during the OSB's passage. All significant changes are laid out in more detail in Break-out Box 1, below.

Non-regulatory approaches

53. This IA does not specifically consider a non-regulatory option as an alternative to legislation. Self-regulation and voluntary approaches to tackle harm were considered as part of the long list policy development process, but given the wide-ranging and significant societal impacts of online harm, inconsistent current voluntary actions, and competing market incentives (as evidenced in the rationale for intervention), the government does not consider non-regulatory approaches on their own to be appropriate.

54. Alongside online safety legislation, there have also been non-regulatory online safety measures. These measures aim to create the optimal conditions for the legislation to be effective. These include initiatives on media literacy, national research projects on child and adult online safety, age assurance technology and safety-by-design, as well as support for the UK online safety technology sector.

55. While this IA does not consider a separate non-regulatory option, the government is committed to supporting businesses and users outside of planned legislation.

Development of the preferred option

56. The then government engaged with platforms, users, Parliament and civil society throughout the development of the online harms policy. Starting in October 2017, the then government published the Internet Safety Strategy green paper. The strategy considered the responsibilities of organisations to their users, the use of technical solutions to prevent online harm and the government's role in supporting users.

57. The Online Harms White Paper (OHWP) was then published in April 2019 . It described a new regulatory framework establishing a duty of care on platforms to improve the safety of their users online, overseen and enforced by an independent regulator. The OHWP proposed that regulation should be focussed on platforms that allow users to share or discover UGC or interact with each other online. Focusing on the services provided by companies, rather than their business model or sector, limits the risk that online harm simply moves and proliferates outside of the ambit of the new regulatory framework.

58. The open public consultation process received over 2,400 responses ranging from companies in the technology industry including large tech corporations and small and medium-sized enterprises, academics, think tanks, children's charities, rights groups, publishers, governmental organisations and individuals. In response to the consultation, the then government gave an indication of its direction of travel in several key areas in the OHWP - Initial Government Response[80], published in February 2020. Here, the then government reconfirmed its commitment to the duty of care approach set out in the OHWP and announced several further measures to guarantee proportionality and protect freedom of expression. It also indicated that the then government was minded to appoint Ofcom as the regulator.

59. Following this, further work was undertaken to develop and refine policy with several important changes made. The full intended policy position was set out in the full government response[81] published in December 2020 along with confirmation that Ofcom would be named as the regulator. In May 2021, the draft OSB was presented to Parliament alongside a consultation stage IA and Parliament established a joint committee to conduct pre-legislative scrutiny.

[80] Initial Government Response to the Online Harms White Paper (HMG, 2020)

[81] Full Government Response to the Online Harms White Paper (HMG, 2020)

Break-out Box 1 - Since the final stage IA: In March 2022, the OSB was introduced to Parliament alongside a final stage IA. A list of the key changes made prior to introduction can be found on page 16 of the final stage IA.[82] Since introduction, and as a result of extensive engagement, a number of policy changes have been made. These include:

- the removal of the 'legal but harmful' adult safety duties (discussed in Break-out Boxes 5, 11, and 16);
- the introduction of new transparency, accountability and freedom of expression duties (Box 17);
- changes to the user empowerment duties, including a new list of content categories that have been set out on the face of the OSA, as well as a requirement for providers to proactively ask users how they would like to use the user empowerment tools (Boxes 14-15);
- a new user empowerment content assessment duty (Box 11);
- further detail to clarify that companies are required to tackle both illegal content and illegal activity (Box 12);
- requirements for platforms to judge whether content is illegal (Box 9);
- expanding Ofcom's powers to give more flexibility to require companies to find the best-fit method of tackling CSA content, including designing solutions in-house or sourcing via a third party (Box 20);
- an expansion of Ofcom's statutory duty to promote media literacy under Section 11 of the 2003 Communications Act, insofar as it relates to regulated services (applies to Ofcom, out of scope for business impacts);
- removal of the harmful communications offence (relates to personal communications, out of scope for business impacts);
- addition of criminal offences relating to sending flashing images ('epilepsy trolling'), sharing of intimate images and encouragement of self-harm (relates to personal communications, out of scope for business impacts).

The impact of these changes is discussed in break-out boxes in the Costs and Benefits section below.

Option 0 – do-nothing

60. **The do nothing option is not able to deal with the current policy problem.** Where legal frameworks exist for illegal content, a significant increase in resources for reporting and law enforcement would be needed to tackle the problem. There is no existing consolidated legal framework requiring providers to put in place safety measures to protect their users from a comprehensive range of harmful user-generated content, tackle harm being caused to children, and safeguard their users' freedom of expression and other rights.

61. **Alongside reporting to the platform, there are several other routes for individuals to report content they believe to be illegal.** For example, the IWF provides a mechanism for individuals to anonymously report online CSEA content. True Vision provides an online mechanism for the reporting of hate crimes and incidents online. There are also government website tools for the reporting of online material promoting terrorism or extremism.

62. **The current systems can rely on voluntary action by platforms.** Under existing regulations, there is very little a user can do in terms of seeking redress and there is no regulatory oversight of a platform's enforcement of their own terms of service.

[82] The Online Safety Bill: impact assessment (HMG, 2022)

63. **In principle, an individual can bring a claim for breach of contract (either in the local small claims court or the High Court) if they consider that a platform has breached any of the terms of service.** Broadly, the individual would need to demonstrate that: (i) a contract exists between the individual and the platform, (ii) the contract was breached as the platform failed to fulfil its obligations satisfactorily, (iii) directly because of the breach, the individual suffered a loss and, (iv) should be compensated.

64. **An individual - who would not need to be a user in a contractual relationship with a platform - could also bring an action in negligence** if they can demonstrate: (i) the internet platform owed them a duty of care, (ii) which it breached, (iii) which caused the individual to suffer loss or harm, and (iv) which was reasonably foreseeable.

65. **In the event of a contractual breach, an individual can seek to recover damages for consequential loss, including personal injury.** Damages for non-monetary loss which don't amount to personal injury (e.g. mental distress or loss of amenity) are awarded only in exceptional cases. Awards of damages for non-monetary loss are more common in negligence claims. Pain, suffering and loss of amenity, and mental distress, are recognised as separate heads on which to bring a claim for non-monetary losses in tort.

66. **Although an individual could bring a claim against an internet platform to seek redress, the government is not aware of any cases having been brought on contractual or negligence grounds (whether successful or otherwise).** This likely reflects the practical and evidential challenges of bringing such claims, the difficulty in showing loss of a sort for which damages can be claimed, and the inevitable costs involved in legal action.

67. **Alternatively, individuals can report harmful content and activity to the platform.** But it is entirely up to the platform as to how it will respond, and how effective that will be, as a means of redress.

68. **The legal incentive for firms to address harmful user-generated content is insufficient.** There are multiple barriers to consumers seeking redress, resulting in limited legal action taken against platforms that may have been in breach of contract when failing to address harmful content.

69. **Under the do-nothing option, platforms face perverse and competing incentives in relation to improving safety and making their sites safe-by-design.** Many service providers' business models are premised on increasing user-engagement and clicks, and this in-turn can reduce incentives to design their sites so they are safe, and so their users are suitably protected from harmful content and activity.

70. **In contrast, some platforms will face incentives to address potentially harmful content to maintain advertising revenue. However, demand for advertising spaces on the main social media platforms is relatively inelastic.** By 2024, internet advertising is expected to account for 70% of total UK advertising spend,[83] and figures for 2021 suggest that this spending is focussed heavily on a small number of large companies (with Google and Facebook alone accounting for 68.5% of total digital advertising spend).[84] It is unlikely that the volume and concentration of spend is significantly sensitive to a platform's moderation activities. Further to this, it is difficult for advertisers to move away from popular platforms, smaller platforms cannot offer advertisers such a large and engaged user base.

71. **Public pressure can act as a driver of content moderation processes but this could ultimately lead to a delayed and reactive approach to addressing harm.** A study of VSPs highlighted that public pressure (as it relates to brand integrity) is a driver of investment in user

[83] Advertising spending in the United Kingdom 2021-2024 (Statsita, 2021)
[84] UK Digital Ad Spending 2021 (eMarketer, 2021)

safety measures.[85] While it is right for platforms to react to user sentiment, this leaves open the possibility that approaches are delayed and only reactive to harm which attracts media attention. Public pressure and a desire to maintain brand integrity is insufficient in ensuring a transparent and proactive approach to addressing harm.

Option 1 - online safety framework

72. Option 1 introduces a new regulatory framework establishing legal duties for companies to improve and protect the safety of their users online, overseen and enforced by Ofcom, an independent regulator.

Platforms in scope

73. The new regulatory framework will apply to:

- any service which hosts UGC which can be accessed by users in the UK
- any service that facilitates private or public interaction between service users, one or more of whom is in the UK
- search services; and
- any service which publishes pornographic content which can be accessed by users in the UK

74. In response to stakeholder views expressed through the public consultation, the then government incorporated the following exemptions for specific types of services:

- **'low risk functionality' exemption**: the OSA exempts user comments on digital content provided they are in relation to content directly published by a platform/service. This will include reviews and comments on products and services directly delivered by a platform, as well as 'below the line comments' on articles and blogs
- **services used internally by businesses**: this is defined as a service (or distinct part of a service), managed by an organisation, whose primary purpose is to host members' UGC and enable interactions between members within that organisation. This encompasses online services which are used internally by organisations such as intranets, customer relationship management systems, enterprise cloud storage, productivity tools and enterprise conferencing software
- **network infrastructure**: any service which doesn't have direct control over the UGC on their platform. In practice, this takes out network infrastructure such as ISPs, Virtual Private Networks and content delivery services as they don't have any control over an individual piece of content. This also rules out business-to-business services e.g. white label or software as a service offered to businesses where again the business doesn't have control over specific pieces of content or activity
- **educational institutions**: online services managed by educational institutions, including early years, schools, and further and higher education providers. This includes platforms used by teachers, students, parents, and alumni to communicate and collaborate. It also includes platforms like intranets and cloud storage systems, but also "edtech" platforms
- **email and telephony**: email communication, voice-only calls and short messaging service (SMS)/multimedia messaging service (MMS) remain outside the scope of legislation

75. Furthermore, business-to-customer interactions are not considered UGC and will also be out of scope (for example video and email interactions between a user and a business). An example of

[85] Understanding how platforms with video-sharing capabilities protect users from harmful content online - (EY, 2021)

this would be a complaints box where users can interact with a business as well as patient-doctor virtual services where users can have a virtual appointment with a physician.

76. Based on analysis conducted by Revealing Reality (RR) and explained in more detail later in the IA, the government expects approximately 25,100 UK businesses[86] to fall within scope of the online safety framework. This figure is different from Ofcom's estimates of scope because DSIT's estimates only relate to UK businesses, and that the number of services in scope worldwide are likely to be materially higher.

Break-out Box 2 - Since the final stage IA: The estimate of 25,100 platforms in scope of the OSA is an increase on the final stage IA. The estimate presented in the previous IA was 24,000. This does not reflect any change in methodology but simply represents the inclusion of pornography providers and the first year of the appraisal period as 2024 with full compliance with the regime from 2025 (the number of businesses grows in line with average growth rate in firms - 3% between 2000-2020). Prior to the announcement of the exemptions above, 180,000 organisations were expected to fall within scope (this reflects a reduction of around 155,000 organisations).

Content in scope

77. The OSA seeks to address the following broad categories of online content:

- ○ **illegal UGC and activity** which is an offence under UK law - such as CSEA, terrorism, hate crime and sale of illegal drugs and weapons;
- ○ **children's exposure to UGC and activity** which gives rise to a foreseeable risk of psychological and physical harm to children - such as children's access to pornographic content, legal suicide content, content promoting self-harm, and content promoting eating disorders, and content which may not be appropriate for younger children such as online abuse, cyberbullying, harmful health content and content promoting or encouraging violence;
- ○ **children's exposure to pornographic provider content** which is published and not user generated; and
- ○ **'protected content'**, including journalistic content, content of democratic importance, and news publisher content.

78. The OSA does not seek to address UGC which gives rise to a foreseeable risk of harm to corporations and organisations and their interests (e.g. copyright offences, competition law). In addition, several categories of UGC and activity are specifically excluded from the scope of the OSA because there are existing legislative, regulatory and other governmental initiatives in place, for example, breaches of data protection legislation, breaches of consumer protection law, and cyber security breaches or hacking.

Categories of regulated services

79. To ensure proportionality, the online safety framework will establish differentiated expectations on companies in scope, with the largest and most influential services subject to additional requirements. The OSA creates different categories of regulated services, these are as follows:

- ○ user to user services meeting the Category 1 thresholds
- ○ search services meeting the Category 2A thresholds
- ○ user to user services meeting the Category 2B thresholds

[86] This figure is based on the Interdepartmental Business Register, which includes all businesses registered for VAT or PAYE tax in the UK, covering businesses that make significant UK sales or employ UK workers.

80. Thresholds for these categories will be set out in secondary legislation; however, they will relate to a platform's number of users and its functionalities and any other characteristics or factors deemed relevant. At a high level, Category 1 platforms are likely to be the highest reach and most influential user to user platforms. We anticipate this will include the largest social media sites and some other services. Category 2A relates to the highest risk and highest reach search services, such as a small group of the largest online search engines. Category 2B services are expected to be high-risk platforms that don't meet the Category 1 threshold. Based on current policy intention, between 30-40 platforms are expected to be designated as either Category 1, 2A, or 2B.

81. In addition, based on regulating pornographic provider content, pornography publishers that do not host UGC or enable P2P interaction will be in scope of the OSA. These platforms will only be required to comply with the duties on published pornography in Part 5 of the OSA ("pornography provisions") and will not be in scope of the core safety duties.

Break-out Box 3 - since the final-stage IA: In the OSB, as appraised in the final-stage IA, Category 1 services would be identified based on the likely impact of the number of users, and its functionalities, on the level of risk of harm to adults from the spread of priority content that is harmful to adults. After the removal of the duties related to priority content that is harmful to adults, the criteria for identifying Category 1 services were amended. The Secretary of State must now consider how easily, quickly and widely user-generated content is disseminated, as well as the user numbers on the user-to-user part of the service, the functionalities of the user-to-user part of the service, and any other factors or characteristics that are deemed relevant.

This amendment affects how the government defines Categories 1, 2A and 2B, and does not directly affect categorised platforms' duties. Ofcom has considered how to assess these criteria and made recommendations to the government.[87] A parallel impact assessment on categorisation appraises the impact of the DSIT Secretary of State's decision on the definition of the categories, considering this amendment.

Break-out Box 4 - Since the final stage IA: The OSA also provides the Secretary of State with a delegated power to bring app stores into the scope of regulation, following consideration of Ofcom's report which must be produced between January 2026 and January 2027. If the Secretary of State considers that there is a material risk of significant harm to an appreciable number of children on or by means of app stores, they will have the power to make regulations placing duties on app stores to reduce the risks of harm presented to children from harmful content on or via app stores. The regulations may make provision to exempt certain types of app stores or specify threshold conditions to narrow which app stores fall in scope of the duties. Indicative estimates for the impact of applying these duties to app stores is found in Break-out Box 20.

Core platform safety duties

82. The primary responsibility for each company in scope will be to implement measures to prevent their services facilitating illegal content and activity, and to protect children from being harmed on their services. The table below outlines which categories of regulated services are expected to comply with each of the core duties.

[87] Ofcom, 2024. Categorisation research and advice submitted to Secretary of State - www.ofcom.org.uk/__data/assets/pdf_file/0023/281354/Categorisation-research-and-advice.pdf

Table 3: Differentiated duties on in-scope companies

Duty	All UGC services	Category 1	Category 2A	Category 2B	Pornography publishers[88]
Risk assessment duties: to assess the level of risk on the platform.	✓	✓	✓	✓	✕
Illegal duty: to put in place systems and processes to prevent, minimise and remove priority illegal content and to remove non-priority illegal content when identified through user reporting.	✓	✓	✓	✓	✕
Child safety duty: If the platform is likely to be accessed by children, to put in place systems and processes to protect children from harmful content.	✓	✓	✓	✓	✕

83. To comply with these core duties in-scope companies will complete an assessment of the risks associated with their services and take reasonable steps to reduce the risks of harm they have identified occurring. The steps a company needs to take will depend, for example, on the risk and severity of harm occurring, the number, age and profile of their users and the company's size and resources. Companies will fulfil their duties by putting in place systems and processes that improve user safety on their services. These systems and processes could include, for example, user tools and content moderation procedures.

84. Robust protections for freedom of expression have been built into the design of duties on companies. Companies will be required to consider users' rights, including freedom of expression online, both as part of their risk assessments and when they make decisions on what safety systems and processes to put in place on their services.

85. The then government has set out priority categories of content that is legal but harmful to children and identified priority categories of offences in primary legislation. This will focus companies', and the regulator's, efforts on the most harmful issues. Companies will still be required to tackle other relevant non-priority material on their services.

Additional requirements on platforms

86. All companies in scope will have several additional requirements beyond the core duties. The table below outlines which categories of regulated service are expected to comply with each of the additional requirements

[88] That do not host UGC or enable P2P interaction and are therefore not in scope of the core duties.

Table 4: Additional requirements beyond the core duties

Duty	All UGC services	Category 1	Category 2A	Category 2B	Pornography publishers[89]
Pornography provision: to prevent children from accessing published pornographic content.	✕	✕	✕	✕	✓
User reporting: to provide mechanisms to allow users to report harmful content or activity and to appeal the takedown of their content.	✓	✓	✓	✓	✕
CSA content: If the platform is a UK platform or is a non-UK platform that does not already report, to report identified online CSA.	✓	✓	✓	✓	✕
Transparency: to publish reports containing information about the steps they are taking to tackle online harm on those services.	✕	✓	✓	✓	✕
Fraudulent advertising: to minimise the publication and/or hosting of fraudulent advertising.	✕	✓	✓	✕	✕
User empowerment: to offer user empowerment tools to give users more control over their online experience.	✕	✓	✕	✕	✕
Freedom of Expression and privacy: to produce FoE and privacy impact assessments	✕	✓	✕	✕	✕
Protected content: to protect journalistic content, news publisher content and content of democratic importance	✕	✓	✕	✕	✕
Transparency, accountability and freedom of expression duties: to have systems and processes to only remove or restrict access to content, or ban or suspend users, in accordance with their terms of service.	✕	✓	✕	✕	✕
Duties related to bereaved parents' requests for data	✕	✓	✓	✓	✕

[89] That do not host UGC or enable P2P interaction and are therefore not in scope of the core duties.

Duty	All UGC services	Category 1	Category 2A	Category 2B	Pornography publishers[89]
Record keeping – of risk assessments and measures taken to comply with codes of practice	✓	✓	✓	✓	✓
Enhanced risk assessments and record keeping - summarising the findings of the most recent illegal content or children's risk assessment in a platform's terms of service	✗	✓	✓	✗	✗

87. Option 1 includes a specific provision which requires pornography publishers to prevent children from accessing published pornographic content (i.e. non user-generated content) using age verification or age estimation. The pornography provision does not form part of the core child safety duties but will be enforced by Ofcom with providers being subject to the same enforcement measures as other in-scope services. The pornography provision does not capture user-to-user content or search results presented on a search service, as the draft OSA regulates these separately (under the Act's core duties). Platforms in scope of the Act's core duties which also carry published (i.e. non user-generated) pornographic content would be subject to both the wider provisions in the draft OSB for user-to-user services and the pornography provision. The pornography provision will deliver an equivalent outcome to the duties for user-generated pornography, in preventing children's access to pornographic content.

88. In addition, all platforms in scope of the core duties will have to provide mechanisms to allow users to report harmful content or activity and to appeal the takedown of their content. Users must be able to report harm when it does occur and seek redress. They must also be able to challenge wrongful takedown and raise concerns about companies' compliance with their duties. All companies in scope will have a specific legal duty to have effective and accessible reporting and redress mechanisms. This will cover harmful content and activity, infringement of rights (such as over-takedown), or broader concerns about a company's compliance with its regulatory duties. Expectations on companies will be risk-based and proportionate, and will correspond to the types of content and activity which different services are required to address. For example, the smallest and lowest risk companies might need to give only a contact email address, while larger companies offering higher-risk functionalities will be expected to provide a fuller suite of measures.

89. The OSA also introduces a legal requirement on technology platforms to report online CSA content. Introducing a CSEA content reporting requirement on UK (and some non-UK) platforms will ensure that they are meeting best practice, which will help protect their users and provide law enforcement with the information they need to identify as many offenders and victims as possible. This requirement will apply differently to platforms depending on where they are based, which is different from the approach being taken within the regulatory regime, where duties will apply to all in-scope services that have UK users. UK platforms (those that provide services from within the UK) will be required to report all detected and unreported CSEA content (all CSEA offences set out in the OSA) to the National Crime Agency (NCA). Platforms providing services from outside of the UK will only have to report UK linked CSEA which has not already been reported under an existing reporting regime outside of the UK.

90. The remaining additional requirements apply only to a small group of influential and higher-reach services. Category 1, 2A and 2B platforms will be required to publish transparency reports containing information about online safety issues. Option 1 does provide the Secretary of State

with the power to expand the providers in scope of transparency reporting requirements if necessary.

91. Category 1, 2A and 2B platforms will also be required to have clear policies for disclosing data regarding a deceased child, and mechanisms for responding to disclosure requests from parents or guardians.

92. Category 1 and 2A platforms will also be required to minimise the risk of fraudulent advertising appearing on their platform. While specific steps will be set out in future codes of practice (subject to consultation and IAs), this will likely include conducting some form of additional checks on advertisers, sharing information on known fraudulent advertisers, and removing fraudulent adverts when reported by users.

93. An even smaller group of the highest influence user-to-user platforms (Category 1) will have several further requirements with which to comply. These include assessing their services for certain kinds of (legal) content, and then implementing user empowerment tools which give users greater control over whether they encounter this content or are alerted to its nature. They will also be required to implement user identity verification methods. They will need to assess the impact of their compliance measures on freedom of expression and privacy and publish this in the form of an impact assessment. They will also need to implement policies which take the importance of the freedom of expression of journalistic content and content of democratic into account, including when making content moderation decisions in relation to these kinds of content. They will need to apply these policies consistently and transparently. They will also need to take steps to protect recognised news publishers' content during content moderation processes, giving these publishers specialised appeal processes.

Break-out Box 5 - Since the final stage IA: Category 1 providers will also have new duties to implement systems and processes to ensure that their own content moderation-related terms of service are clear and accessible to users, and consistently and transparently enforced. They will need to put in place additional systems and processes to protect journalistic content and content of democratic importance, offering a higher level of protection for that content when making content moderation decisions.

Preferred option and implementation plan

94. **Option 1** is the government's preferred option and is the result of extensive engagement with platforms, Parliament, civil society organisations, and wider society. The preferred option is risk-based and proportionate and is expected to achieve the stated policy objectives. Importantly, the preferred option does not place undue burdens on platforms where there is a low, or no, risk of harm.

95. The OSA received Royal Assent in October 2023. Implementation is expected to be carried out in three phases, set out below. During each phase Ofcom will publish draft guidance and codes of practice for consultation, then refine these carrying out further research and analysis as required before publishing final guidance and submitting codes to be laid in Parliament.
    - ○ Autumn 2023 to Winter 2024 - illegal content duties
    - ○ Spring 2024 to Spring 2025 - children's safety duties and child access assessment
    - ○ Spring 2025 to Spring 2026- duties for categorised services

    Further detail is set out in Ofcom's implementation roadmap update on their website.[90]

[90] https://www.ofcom.org.uk/online-safety/information-for-industry/roadmap-to-regulation/0623-update

Appraising the preferred option

Approach to appraisal

96. At this primary legislation stage, it is not possible to predict with certainty the actions of Ofcom or the steps platforms may take to ensure they are compliant with the regulation. While the OSA sets out duties, the specific requirements and the actions platforms can take to comply will be set out in codes of practice laid by Ofcom and where necessary secondary legislation. Given that specific requirements are unknown at this stage, costs and benefits included here are largely illustrative and aim to indicate the potential scale or nature of impacts of the whole policy (scenario 2 in the Regulatory Policy Committee's (RPC's) primary legislation guidance).[91] All future codes of practice will be subject to an IA, including an assessment of the impacts on small and micro-businesses (SMBs) and innovation (IA requirements on Ofcom under the OSA go beyond current regulatory requirements under the Small Business, Enterprise and Employment Act 2015).[92]

97. While it is not possible at this stage to provide a fully monetised appraisal of the policy or a verifiable assessment of the EANDCB, every effort is made to provide an indication of the likely scale of impact of the whole policy (including future codes) through presenting illustrative monetised costs, proxied impacts from similar policies, and comprehensive qualitative analysis.

98. For this IA's assessment of potential benefits, it is not possible to develop a precise estimate of the reduction in online harm that will be achieved by the preferred option. Instead, this IA attempts to quantify the economic cost of online harm under a *do nothing* counterfactual and conducts both break-even analysis and scenario analysis based on a range of illustrative harm reduction scenarios.

99. While timelines are dependent on external factors, for appraisal purposes, this IA uses a ten-year appraisal period running from 2024. Familiarisation costs are assumed to be incurred in the first year of the appraisal period with full compliance from 2025. This approach is an analytical simplification - in reality, codes of practice are likely to be staggered each allowing time for businesses to ensure compliance and are unlikely to fall in line with calendar years. All impacts are presented in 2019 prices and 2020 present value base year.

Main sources of evidence

100. This final stage IA supporting the OSB, draws on several evidence sources to attempt to provide an indication of the likely scale of impacts.

○ **Revealing Reality (RR) research**: In 2020, the then government commissioned consultancy firm Revealing Reality to estimate the number of organisations in scope of the framework and to determine the likely incremental costs of compliance.[93] More details on the methodology are included later in this IA, and the results form the basis of our current estimates.

[91] RPC case histories – primary legislation IAs, August 2019

[92] Small Business, Enterprise and Employment Act 2015

[93] In-Scope Organisations' Approaches to Preventing Online Harm (Revealing Reality, 2022)

Break-out Box 6 - Since the final stage IA: In 2023, the government commissioned RR to conduct further research to understand the impacts on business following changes to the policy since the previous research. Estimates have been updated to reflect the results of this research. This will be published alongside this impact assessment as Updating Analysis of the Potential Impact of the Online Safety Bill on In-scope Organisations.

- **AVMSD research:**[94] the government commissioned consultants from EY to research the measures that VSPs take to protect users online ahead of the implementation of the rules for VSPs under AVMSD. The Directive sets requirements on VSPs to protect users from harm. Ofcom is the regulator for UK-established VSPs and therefore, actions taken and costs incurred by in-scope businesses represent another reasonable proxy for the costs of the OSA. Note that qualitative and quantitative evidence was collected from platforms within and outside of the UK's jurisdiction.
- **rapid evidence assessment (REA) of NetzDG**: Despite numerous countries considering how to make the internet safer for users (see 'international context' section above), international policies addressing this issue are either planned and not yet implemented or have not been fully assessed. As such, comparisons between the OSA and similar international policies have been limited to Germany's NetzDG which is aimed at combating hate speech online which came into effect on 1 January 2018. NetzDG has been in force for a reasonable amount of time and while there are significant differences between NetzDG and the OSA, both address online harm to some extent, and it is a useful proxy. The government conducted a REA to provide an overview of the impact of NetzDG in Germany specifically in relation to compliance costs faced by businesses, the impact of the law upon market innovation and whether it has reduced online harm.
- **business engagement**: Since the publication of the OHWP, the government has engaged extensively with affected platforms. Engagement includes a series of bilateral meetings, roundtables, and a cost survey. In addition to several bilaterals with in-scope platforms, the government has held several roundtables with industry on the topic of compliance costs and issues relevant to small and medium sized enterprises. In addition, following the publication of the full government response, the government sent cost surveys to a sample of 36 platforms to understand in greater detail how they are preparing for regulation and any costs associated with the preparations. The sample consisted of 10 of the 16 largest social media platforms in the UK, review platforms, games organisations, retail sites, dating sites, and forums. The result of this engagement is discussed throughout this IA but is mainly qualitative, given that platforms were largely unable to provide cost information without knowing the content of future codes. A selection of UK-focussed AV providers was also engaged with a cost survey, the results of this can be found in the age assurance cost section.
- **Ofcom call for evidence on VSPs**:[95] Ofcom published their call for evidence on VSP regulation on 20 July 2020. There were 39 non-confidential responses which included social media platforms, platforms with video sharing capabilities, public sector institutions and individuals. The findings from the call for evidence were used to inform the development of the draft and final guidance on the VSP regime. The call for evidence was divided into two parts: i) Queries for industry which included questions on the services provided and on the mechanisms for keeping users safe online; ii) Questions for all stakeholders which included queries on how the design of the VSP regulatory regime can best keep users safe online.

[94] Understanding how platforms with video-sharing capabilities protect users from harmful content online (EY, 2021)

[95] Consultation on guidance for VSP providers on measures to protect users from harmful material (Ofcom, 2021)

Costs and benefits

Baseline

101. Evidence on the current level of harm mitigation under the baseline is limited. The systems and processes platforms have in place vary significantly across platforms, as does spending on user safety. For some platforms keeping users safe online is part of the organisation's ethos and for other platforms activities such as content moderation is much lower in their priorities.

102. RR research found that, in general, the mitigations an organisation had in place were proportionate to the organisation's risk of potential online harm, i.e. higher risk platforms had many more protections in place than low risk platforms. Human and automated moderation was present across all risk categories of platforms, whereas processes such as reporting functions, paying for access to databases, such as PhotoDNA, and publishing transparency reports, were only present in higher risk businesses. This was supported in engagement with stakeholders. The vast majority of platforms engaged already conduct risk assessments, set terms of service and acceptable use policies, conduct both human and automated moderation, allow users to report harm, and have systems to handle complaints.

103. RR research also found that different types of mitigation are implemented to varying degrees. For instance, while automated moderation is used throughout, the complexity and tailoring of this to the specific platform varies. For example, a low-risk organisation interviewed uses 'off the shelf' automated moderation to detect spam, whereas a high-risk organisation uses their own bespoke automated software tailored to detect specific harm present on their site. Findings from the RR research on high variability in current mitigations was corroborated by EY's AVMSD research of VSPs. They found that the measures employed by each platform depended on the nature of the risks, the level of resources of the platform, the type of content on the platform, the impact on the platform's brand, and competitive considerations.

104. Most organisations are already investing in protecting their users in the absence of regulation and platforms expect this investment to continue to increase over time. EY's AVMSD research found that total annual expenditure on measures to protect users from harmful content ranged from hundreds of pounds for the very smallest platforms to over £1.5bn for the largest platforms.

105. Organisations and online platforms have also dedicated significant resources to specifically tackling online child sexual exploitation and abuse (CSEA). In March 2020, following a Five Country ministerial meeting between the UK, US, Australia, Canada and New Zealand, the UK launched the Voluntary Principles to Counter Online Child Sexual Exploitation and Abuse. These principles set out a consistent and high-level framework for industry actors, aiming to co-ordinate the approach to tackling online CSEA globally, outside of formal regulation. The promotion of these principles has been supported by the WePROTECT Global Alliance across 97 governments, 25 technology companies and 30 civil society organisations.

106. While accurate evidence does not exist on the current UK-wide level of harm mitigation under the *do nothing* option, this IA - where at all possible - attempts to incorporate this in the costing of Option 1, i.e. only the incremental costs of regulation have been included.

107. As outlined in the above rationale for intervention, many categories of online harm have been increasing in prevalence and increased screen time resulting from COVID-19 has likely

exacerbated this. This IA estimates that under the *status quo* online harm will result in a societal cost to the UK of at least £254 billion (PV) across the appraisal period (the calculations underpinning this estimate can be found in the benefits section below) - this is based only on a subset of harm that this IA was able to quantify.

Summary of impacts

108. All impacts are assessed in aggregate for all in-scope businesses over a 10-year appraisal period starting from the date of implementation. For present value costs and benefits, a discount rate of 3.5% has been applied in line with Green Book guidance. All costs are presented in 2019 prices with 2020 as the present value base year. Given the uncertainty around future requirements, costs and benefits are illustrative and attempt to provide an indication of the likely scale of impact from primary legislation, related secondary, and future codes of practice. Three scenarios are set out in Table 5: a low, central and high scenario. Details and assumptions underpinning these scenarios are outlined in further detail in each of the accompanying cost sections that follow the table.

Table 5: Summary of impacts

Impact	Treatment	Low	Central	High
Reading and understanding the regulations	Cost to business (monetised)	£39.7 million	£53.6 million	£74.1 million
Ensuring users can report harm	Cost to business (monetised)	£17.0 million	£22.0 million	£31.9 million
Updating terms of service	Cost to business (monetised)	£14.6 million	£17.5 million	£23.4 million
Reflecting the illegal content judgement	Cost to business (monetised)	£16.1 million	£41.0 million	£86.4 million
Conducting risk assessments	Cost to business (monetised)	£12.9 million	£28.1 million	£43.3 million
Undertaking additional content moderation	Cost to business (monetised)	£1,340.0 million	£1,920.0 million	£2,500.0 million
Employing age assurance technology	Cost to society (monetised)	£18.4 million	£36.7 million	£91.8 million
Transparency reporting	Cost to business (monetised)	£0.8 million	£6.7 million	£10.1 million
Fraudulent advertising duty	Cost to business (monetised)	£52.7 million	£120.0 million	£187.0 million

Impact	Treatment	Low	Central	High
User verification and empowerment duties	Cost to business (partially monetised)	£9.0 million	£12.3 million	£14.1 million
FoE and privacy impact assessments	Cost to business (monetised)	£1.0 million	£2.5 million	£11.0 million
Reporting online CSEA to the NCA	Cost to business (partially monetised)	£0.0 million	£0.1 million	£0.2 million
App stores - age verification	Cost to business (non-monetised)	£4.4 million	£9.6 million	£19.8 million
App stores - app review	Cost to business (non-monetised)	£1.4 million	£7.5 million	£13.6 million
Industry fees to regulator	Cost to business (monetised)	£539.0 million	£539.0 million	£539.0 million
Enforcement action (fines, business disruption measures and senior management liability)	Cost to business (non-monetised)	n/a	n/a	n/a
Justice impacts	Cost to government (monetised)	£99.4 million	£170.0 million	£241.0 million
Requirement to report CSA content	Cost to government (monetised)	£7.0 million	£10.8 million	£15.1 million
Wider impacts (freedom of expression, privacy, competition, innovation, trade)	Cost to society (non-monetised)	n/a	n/a	n/a
Reduced prevalence of online harm	Benefit to society (non-monetised)	Break-even: 0.9%	Break-even: 1.3%	Break-even: 1.6%

Approach to business costs

Number of platforms in scope

109. Given the wide-ranging scope of Option 1 and the lack of granular data, it is difficult to determine with precision the number of affected organisations. However, this IA considers the sampling approach used in the previous IA and explained below to be the most robust existing evidence on in-scope firms and the methodology, therefore, remains unchanged. The number of affected UK-based organisations (including businesses and civil society organisations) within scope of the regulations is estimated to be around 25,100[96] in the first year of the appraisal period. This equates to between 0.3-0.4% of all UK businesses.

110. To determine the number of platforms in scope, RR[97] extracted a stratified sample of 500 organisations from the Inter-Departmental Business Register (IDBR).[98] The sample consisted of 100 randomly selected businesses in each of the following size categories (sole traders, micro (not including sole traders), small, medium and large).[99] A sample of 500 is considered to be large enough to provide robust estimates as it ensured a relatively small margin of error at the 95% confidence level (between ±2.6 to 4.4 percentage points). Additionally, every organisation within the sample was manually reviewed and categorised according to features that could be considered as hosting or allowing users to search for UGC or enabling P2P interaction. Findings were then extrapolated using the Department for Business, Energy and Industrial Strategy's (BEIS) Business Population Estimates (BPE)[100] to estimate the total number of in-scope platforms in the UK. This is in line with RPC guidance on defining a business by taking a 'GDP approach', i.e. the assessment of impacts on business are in terms of the location of the economic activity being in the UK. This initial sampling and extrapolation resulted in an estimate of approximately 18,000 in-scope platforms.

111. Option 1 will apply to CSOs as well as businesses. While the IDBR (from which the original sample was taken) does include CSOs, BEIS' BPE does not. To address this, findings from the original sample were further extrapolated using data on CSOs in the UK Civil Society Almanac[101] and around 550 CSOs were added to the estimates. This is a reasonably reliable methodology for determining the number of CSOs in scope; however, it does have limitations:

- ○ The same methodology used for all businesses is applied to CSOs. This is therefore an approximation as the actual size and risk-level of CSOs will be slightly different to that of all platforms in scope.
- ○ The estimate of CSOs in scope is likely an overestimate. When CSOs do utilise UGC this is mainly through third-party sites like Facebook, X (formerly Twitter), and YouTube who themselves would be in scope rather than the CSO.
- ○ For all organisations in scope, 'size' was quantified in terms of number of employees (as in the SBEE Act). This is not possible for CSOs, largely because a large amount of the

[96] The exact estimate is 25,051

[97] Revealing Reality, 2022, In-Scope Organisations' Approaches to Preventing Online Harm

[98] A comprehensive list of UK businesses used by the government for statistical purposes.

[99] The definition is in line with SBEE Act.

[100] An estimate of the total number of private sector businesses in the UK at the start of each year, with their associated employment and turnover.

[101] UK Civil Society Almanac 2020 - NCVO

workforce are volunteers. Instead, and in line with standard appraisal practice in this area, CSOs are ranked by annual revenue.[102]

112. In line with the consultation stage IA, specific actions resulting in transition costs and compliance costs are assumed to be the same for businesses and CSOs (differing only based on organisation size and the risk of harm occurring on the platform). There has been no evidence submitted as part of pre-legislative scrutiny or in response to the consultation stage IA to suggest that this assumption should be revised. In addition, Option 1 is functionality based and is sector agnostic. An in-scope CSO with the same risk profile as an in-scope business would incur the same costs.[103]

113. Acknowledging the potential for gaps in the random sample (for example the lack of in-scope small platforms), additional types of organisations were identified and included in the estimates. For example, crowdfunding or fundraising sites, dating sites and forums were added to the sample on the assumption that all (or at least most) of these would fall within scope. Approximately 3,000 platforms were added to the estimates for a total of around 21,600. It is important to note that these additions do not represent an exhaustive list of all types of organisation that could be in scope, but are an attempt to deal with some of the larger groups to provide a more realistic estimate. Estimates for the number of in-scope platforms provided by the RR research were based on 2019 data from the IDBR. These were uplifted by the average annual growth in the business population to account for an implementation date of 2024. For modelling purposes, this growth rate continues throughout the appraisal period. It is important to note that the number of businesses does not reflect the number of services, as some businesses operate multiple sites and therefore the number of services is likely to be higher.

114. Steps taken so far focus on determining the number of organisations in scope based on whether they operate services that host UGC, enable P2P interaction or are search engines - these platforms are in scope of the core duties and are the main regulated entities. The pornography provision is an additional requirement on pornography publishers to prevent children from accessing non-user-generated pornographic content, regardless of whether they are in scope of the core duties. As highlighted by a study of children’s access to pornography, explicit adult content can be found on a variety of platforms, including social media sites, VSPs, search engines, chat sites, and dedicated pornography sites.[104] The majority of these types of platforms are already in scope of the core duties as user-user services and will be captured in the estimates above.

115. Even though many of the most visited pornography sites and sources are in scope of the OSA’s core duties, an important proportion of dedicated pornography sites are not, as they do not host UGC or enable P2P interaction and so are only in scope of the pornography provision. To illustrate this, an assessment of the top 200 pornography sites found that 36% of sites (or 72 sites) and 16% of traffic was to sites outside of the scope of the OSA’s core duties.[105] The pornography provision will ensure these sites protect children from harmful content.

[102] CSOs are matched to the business size categories based on average revenue by business size as presented in BEIS’ BPE.

[103] Of course, requirements would be proportionate to both risk and resources available.

[104] Another potential source of online pornography are UK-based video-on-demand (VoD) sites. The impact assessment for Part 3 of the Digital Economy Act estimated that there were around 100 of these based in the UK. More recent estimates suggest around 150, with 40-50 of these being adult services. VoDs are not in scope of this OSA as they are already subject to existing regulation (the UK’s transposition of the Audiovisual Media Services Directive) which includes protecting children from pornography as well as wider duties related to product placement, sponsorship, and incitement to racial hatred.

[105] BBFC Assessment of Adult Sites’ Functionality (BBFC, 2020)

116. Given the nature of the industry, evidence on the number of pornography publishers and the location of their economic activity is limited, it is difficult to determine with certainty the number of UK-based pornography publishers that do not host UGC or enable P2P interaction. Based on an assessment of the top 200 pornography sites most popular with UK users, the BBFC found that only four were based in the UK (out of 126 for which this information was available).[106] The vast majority of pornography sites are based in the US and, even there, industry reports on the US market put the number of businesses operating pornographic websites at 103.[107] The same report estimates that a single organisation, namely MindGeek, holds an 82% market share and owns many of the most popular sites. One report - although conducted in 2013 - estimates that 60% of pornography sites are hosted in the US, compared with 7% in the UK.[108] Using sites as a proxy for the number of businesses based in each country and comparing the US to the UK, this would suggest that the number of UK-based pornography publishers in 2022 is likely to be around 12, broadly in line with low number of sites based in the UK from BBFC's research findings on country of origin.[109] Uplifted by the average annual growth in the business population to the first year of appraisal, this impact assessment conservatively estimates an additional 12 UK-based businesses in scope as a result of the pornography provision.[110] The number of businesses does not reflect the number of pornography sites, as each business is likely to operate multiple sites - as is the case in the US market.[111]

117. Following the above steps, the final estimate for the number of in-scope platforms is approximately 25,000 organisations.

Table 6: Steps to attain an estimate for the number of in-scope platforms

	Micro	Small	Medium	Large	Running total
Percentage in-scope within sample	0.3 %[112]	0 %	2 %	4 %	-
Number of in-scope platforms within UK economy (nearest hundred)	17,100	0	800	400	18,300
Number of in-scope CSOs within UK economy[113] (nearest hundred)	400	0	100	<100	18,900
Accounting for gaps in sample with known types of platform	-	~1,000	~2,000	-	21,600
Number of non-UGC pornography publishers[114]	11	1	0	0	12

[106] BBFC Further Research on Traffic to and Functionality of Adult Sites (BBFC, 2020)

[107] Adult & Pornographic Websites Industry in the US - Market Research Report (IBIS World, 2022)

[108] Top 10 adult website host countries (Metacert, 2013)

[109] BBFC Further Research on Traffic to and Functionality of Adult Sites (BBFC, 2020)

[110] Of course, some of these businesses are likely to have already been captured as user-user services but this impact assessment conservatively assumes that all are additional rather than a proportion. As these businesses are not in scope of the core duties, there is no risk of double counting costs based on this conservative approach.

[111] The number of businesses does not reflect the number of pornography sites, as each business is likely to operate multiple sites - as is the case in the US market.

[112] Weighted data combining 0 employee and 1-9 employee bands

[113] Note the size of CSOs is determined by annual revenue in line with appraisal practice in this area.

[114] The proportion of additional pornography publishers estimated to fall within each size category is based on business demographics within creative industries. DCMS Sectors Economic Estimates 2022: Business Demographics (DCMS, 2022)

Number of in-scope platforms uplifted to 2024[115] (nearest hundred)	20,200	1,200	2,900	700	**25,100**

118. The methodology described above was conducted both before and after the then government announced a list of exemptions for specific types of services (see Schedule 1, e.g. email services, SMS and MMS services, limited functionality services, internal business services etc) in December 2020. Before the exemptions, it was estimated that around 3% of all UK businesses would have been in scope, equating to approximately 180,000 platforms. The exemptions therefore removed approximately 155,000 platforms from the scope, mostly SMBs exempted by the low risk functionality exemption. This IA conducts sensitivity analysis on the number of businesses in scope in the risks and sensitivity section.

Risk categorisation of platforms

119. Option 1 is risk-based which means that there are differentiated expectations on companies in scope regarding different categories of harmful content and the additional requirements outside of the core duties (see Table 4). Also, even within the differentiated platform duties, expectations on platforms will differ depending on the risk of harm on their platform and the resources available to the platform. For example, while every platform will be required to assess the risk of illegal harm on their platform, the level of detail required and the steps they must take in producing these risk assessments will vary greatly. This approach ensures proportionality both in the differentiated duties and in the specific way in which platforms can comply with codes. Platforms which offer services with the lowest risk of online harm will face the lowest regulatory burdens and platforms offering high-risk services will be required to take the most action.

120. Given that the specific way in which platforms can comply will be set out in future codes of practice, it is not possible to know exactly what they will do. However, to reflect this proportionality in the analysis of businesses' impacts, the then government commissioned the production of an organisation categorization framework (OCF) to split platforms into three risk tiers (low, mid, and high) which helps with estimating the type of likely actions they would take in complying with Option 1. The OCF was developed using extensive desk research and interviews with experts, such as the IWF, Childnet, and Internet Matters.

121. The OCF first identified all factors that could define whether an organisation could fall in scope of the OSA and factors that could affect its ability to tackle online harm. The two primary categorisation criteria incorporated into the OCF were 'features' and platform size (as measured by the number of employees). There were 41 features that enabled users to share or discover UGC or enable peer-to-peer (P2P) interaction assessed as part of developing the OCF. These included features such as the ability to livestream, share content that exists on the platform, post reactions to content, group message, video call, post comments under content, geo-tag, search functionality, and display a feed of UGC. The OCF was used for research purposes only and is not directly related to the contents of the OSA or a checklist by which platforms can determine whether they are in scope. Instead, it is a set of criteria which enables manual assessment of sample platforms.

122. The categorisation of in-scope platforms in the analysis was done through a 'scoring' system where in-scope features add to the service's risk score as does an organisation's reach - this approach is likely to be broadly in line with how the legislation's thresholds will work in practice. In addition, services targeted at or used primarily by children are assigned a higher score (this reflects the additional requirements on services 'likely to be accessed by children'). Scoring based on the OCF indicated that the majority of the around 25,100 in-scope organisations (over 97%) fall into the

[115] Start of the appraisal period and expected date of implementation

low and mid risk categories (49% and 48% respectively). Less than 3% of in-scope organisations could be considered high risk platforms and less than 0.001% are estimated to meet the Category 1 and 2A thresholds (additional requirements on the largest and highest risk platforms, based on policy intention this is expected to be around 20 platforms).

123. Platforms in scope will vary greatly as will the way in which they offer functionality that allows UGC and P2P interaction. Table 7 provides some examples of the types of organisations that could fall within each risk category:

Table 7: Example types of organisations within the OCF risk categories

Risk tier	Example features within sample	Example organisations
Low risk	● comments sections (for UGC) ● ability to like content	● retail websites (that are not out of scope due to the limited functionality exemption) ● blogging platforms
Mid risk	● ability to post content ● message someone you know or have friended	● forums ● dating sites ● online gaming
High risk	● feed of UGC ● live Streaming ● ability to contact unknown users	● social media companies ● large search engines ● video sharing platforms

124. It should be noted that Table 7 is illustrative and used for analytical purposes only. It presents example types of organisations that may fall within the different risk categories based on current understanding of the types of functionality present on these platforms.

Break-out Box 7 – Since the final-stage IA: The Regulatory Policy Committee have raised a question about whether our assumption that the number of Category 1 platforms will be static over the appraisal period is reasonable.

This appraisal assumes that the number of Category 1 platforms, which are subject to higher requirements, is likely to be stable. The regulation defines Category 1 firms in terms of reach, the number of users on a platform (expressed as a proportion of the population), as well as functionality.

As we anticipate that the largest social media platforms will be in scope of Category 1, we considered recent data about social media platforms' user base to test the assumption that the user base, and the number of Category 1 platforms, are unlikely to increase. Theoretically, the user base of user-to-user platforms could change based on either a change in the number of people on social media generally, the number of social media sites each user visits, or a mass shift in users from one platform to another. Therefore, we considered the proportion of people with internet access and access to social media and the average number of social media sites used per person.

The number of people using the internet and social media is broadly at saturation. Though internet and social media usage has increased over the last eight years, in the most recent three years, the percentage of adults accessing the internet has remained relatively stable at 95%

which we take to be saturation (and doesn't leave scope for significant further growth over the appraisal period). Concurrently, social media usage reached 89% of internet users in the most recent period[116], with this likely to continue considering the high proportions in the 12 to 15 age categories over the last two years (90-93%).[117] Reach is expressed as a proportion of the population, so thresholds will adjust with population growth, rather than "dragging" platforms into static population thresholds over time.

Data from recent years indicates that the average number of social media sites used per UK user has also remained relatively constant, with 6.8 sites in 2020[118] and 6.4 sites in 2024[119]. These statistics suggest that social media has reached a saturation point, making significant changes in the number of Category 1 firms unlikely and supports the assumption of a stable number of Category 1 firms.

Finally, platforms hosting user-to-user content are subject to network effects, where their value to users is dependent on the number of other users/users they know on the platform. This makes them less likely to suffer from mass shifts in users, because an alternative platform must offer sufficient benefits to offset the lost value from moving to a smaller platform. In 2023 a series of alternative platforms were established in response to a major platform changing its terms of service, though none have succeeded in amassing a comparable number of users.

Though none of these definitively prove that the number of Category 1 platforms will be static, it does suggest that the assumption is reasonable, and there are not better assumptions.

Development of platform actions

125. It is difficult at this stage to estimate with certainty the steps platforms will take and the costs they will incur complying with the OSA. This is because:

○ Option 1 establishes differentiated expectations on companies in scope regarding different categories of harmful content and the additional requirements outside of the core duties. Thresholds for Category 1, 2A and 2B will be set out in secondary legislation and therefore, it is unclear at this stage which organisations they will apply to.

○ Option 1 is proportionate even within duties, and expectations will vary greatly between for example small low risk platforms and large high risk platforms. Specific steps in-scope platforms can take will be outlined in future codes of practice laid by Ofcom (themselves subject to IAs).

○ companies will need to comply with the codes; however, if preferred, they will also be able to demonstrate to the regulator that an alternative approach is equally effective.

○ even while some aspects of the OSA will clearly result in specific actions such as conducting risk assessments or transparency reporting (for Category 1, 2A and 2B), the steps platforms can take and the information required in these will not be set out until future codes of practice.

[116] Ofcom, 2024, Adult Media Use and Attitudes Report

[117] Ofcom Online Nation and Media Use and Attitudes reports

[118] Backlinko, via Internet Archive Wayback Machine, impression October 2020 of a page updated in August 2020

[119] Backlinko, accessed 30 July 2024

- the high-level duties related to illegal content, transparency accountability and freedom of expression, and protecting children set out at primary stage legislation are not prescriptive and therefore, any attempt to estimate the specific actions taken by platforms is speculative.

126. A common theme of the then government's engagement with in-scope platforms is that they are unable at this stage to provide reasonable estimates of costs or even actions likely to be taken to comply. This is to be expected at this stage and following introduction, Ofcom will begin a series of consultations with industry on codes of practice and will produce IAs to determine the costs to platforms.

127. Given the uncertainties at the primary stage, this IA develops a plausible set of actions platforms may take based on estimates of the size and risk of harm on the platform. These include:

- **reading and understanding the regulations (familiarisation costs) -** this includes both primary legislation and related secondary, and future statutory codes of practice
- **ensuring users can report harm -** this relates to the mechanism through which users can report harm and could be as simple as a visible email address (already a statutory requirement) or a system which can triage large volumes of reports.
- **updating terms of service** - evidence discussed below suggests that this is a business-as-usual activity for in-scope platforms. However, platforms may decide to assess and update their terms of service in response to future codes of practice.
- **reflecting the illegal content judgement in content moderation advice -** this relates to the requirement to assess whether content is illegal, e.g. considering intention as well as behaviour, where relevant, and ensuring content moderators are equipped to make this judgement.
- **conducting risk assessments** - this relates to the requirement to carry out an illegal content risk assessment and 'if likely to be accessed by children' to carry out a children's risk assessment.
- **undertaking additional content moderation** - the OSA does not require additional content moderation; however, it is likely that platforms will increase resources in this area to comply with the duties.
- **employing age assurance technology -** in complying with the child safety duties higher risk platforms are likely to adopt age assurance (and specifically age verification) technologies. In addition, certain platforms will be required to use age verification or age estimation to protect children from primary priority content that is harmful to children and provider pornography.
- **transparency reporting** - this relates to producing annual published reports on platform harm and related actions taken by the platform.
- **fraudulent advertising duty (customer due diligence) -** as part of complying with the fraudulent advertising duty, it is likely that in-scope platforms will conduct CDD (customer due diligence) on advertisers.
- **user verification and empowerment duties** - this relates to the requirement on large social media platforms to offer optional user verification and provide user empowerment tools.
- **assessing impacts on freedom of expression and privacy -** this relates to publishing an assessment of impacts on freedom of expression and privacy and keeping this updated.
- **reporting online CSA content to the NCA** - this refers to the cost of reporting identified CSEA content to the NCA.

128. Cost estimates for this plausible set of platform actions is based on evidence provided by platforms, proxied from similar regulations, or based on reasonable assumptions of time requirements and standard appraisal practice.

Costs to business

129. For appraisal purposes, it is assumed that legislation enters into force in 2024. The first year is assumed to be a transition year giving platforms time to prepare for compliance based on the specific details set out in codes of practice and secondary legislation. This IA assumes that platforms will incur familiarisation costs and transition costs in the first year but will not incur compliance costs until year two. This is a simplified assumption for analytical purposes only - in reality, the codes of practice will be staggered, and platforms will ensure compliance across several years.

Familiarisation costs

Requirements

130. In-scope platforms[120] will be expected to familiarise themselves with the regulations which includes understanding which aspects of the safety duties apply to them and what steps they must take to ensure compliance.

Cost estimates

131. Platforms are expected to incur the following costs associated with familiarisation:

- **initial familiarisation**: while only in-scope platforms are required to familiarise themselves some, who think they could potentially be in scope[121] under a broad interpretation of the regulations, may have to read the regulations - even if only to determine that they are out of scope.
- **potential legal advice for SMBs**: in-scope SMBs may require legal advice to clarify aspects of scope and which parts of the OSA apply to them.
- **secondary familiarisation**: Beyond the initial familiarisation, actual in-scope platforms are expected to spend more time reading the regulations.
- **dissemination of information**: medium and large in-scope platforms are expected to disseminate the information across a proportion of their staff.

Break-out Box 8 – since the final-stage IA: As of the time of writing, Ofcom has published several consultations on draft guidance and codes of practice. Though the codes of practice and guidance are not finalised and impact assessments for the codes of practice will be produced by Ofcom, these draft documents indicate that the final-stage IA assumptions for familiarisation significantly underestimated costs.

We have updated the assumptions below, based on the draft guidance and codes of practice, while retaining the original methodology from the final-stage IA.

Two key tranches of guidance have not been published, the ‘transparency duties’ and ‘duties on categorised services’. In consultation with Ofcom, we have used 90 pages as a reasonable

[120] Including pornography providers that are in scope of the pornography provision but not the core duties. These platforms are expected to incur full familiarisation costs.

[121] Those which offer online services with any features that could be considered in-scope such as posting, sharing, reacting to content, messaging, calling, commenting, tagging, discovering or seeing UGC.

estimate for the transparency duties. For categorised services duties, we have assumed 1641 pages of guidance, the length of the illegal content draft guidance, as a conservative estimate.

The effect of these revised assumptions appears in Table 8.

132. For initial familiarisation, based on RR research, there are approximately 180,000 platforms that could be considered potentially in scope. For initial familiarisation, it is estimated that between 20%-50% of all platforms potentially in-scope would complete Ofcom's guidance and quiz on scope (25% in the central scenario) - this is approximately 20,000 out-of-scope platforms incurring costs of familiarisation. As with other regulations, it is very difficult to predict with certainty how many firms outside of scope would incur costs of familiarisation - evidence for this within the context of online harms is extremely limited. The assumed range merely represents a conservative estimate to provide an indication of the likely scale of impact on out-of-scope platforms. These platforms are likely to be on the margin where it isn't instantly clear whether they would come under the regulations, unlike for example, email service providers (who are out of scope) where it would be immediately obvious. For the initial familiarisation, one regulatory professional at an hourly wage of £21.12 is expected to read the regulations within each business (all wages in this IA come from the Annual Survey of Hours and Earnings[122] and are uplifted by 22% to account for non-wage labour costs). The online guidance on scope is close to 4000 words, and takes just over 20 minutes to complete based on a reading speed of 200 words per minute.[123] This results in the cost of initial familiarisation of between £0.3 million and £0.7 million (central, £0.3 million).

133. In addition, the central estimate includes one hour of legal advice for every in-scope SMB. Legal advice is not included in the low estimate and rises to two hours for every in-scope SMB in the high estimate. The inclusion of legal advice represents the potential need to confirm whether a platform does fall within scope and to advise on which aspects of the OSA are likely to affect them. While many SMBs may not require this, some will likely seek more extensive legal advice. By assuming one hour for every firm, this IA attempts to capture the total cost rather than provide an accurate per platform estimate.[124] This IA estimates the cost of legal advice to be between £0 and £1.0 million (central, £0.5 million).

134. For secondary familiarisation, in-scope platforms are expected to spend more time reading the regulations. For these (around 25,000), another member of staff in micro-platforms (rising to 2, 5, and 10 for small, medium, and large platforms respectively) is expected to read Ofcom's guidance and Codes of Practice. Considering the page length of the published draft guidance and RPC advice on the time to read guidance[125], we estimate this to take between 30-55 hours per reader. For medium and large platforms, these staff are expected to be regulatory professionals whereas wage estimates for Chief Executives are used for in-scope SMBs. The unpublished guidance applies to categorised firms only (assumed to be 35) and represents a further 16-29 hours of reading for these large businesses. Secondary familiarisation is expected to result in costs of £38.7 million-£70.3 million (central £51.4 million).

135. Finally, for medium and large in-scope platforms, costs are expected to be incurred through disseminating the information across a proportion of their staff. While it is unclear what exact proportion of staff will need to be made aware of the regulations, this IA estimates that between 5%-20% of staff within in-scope medium and large platforms will spend 30 minutes familiarising themselves (10% in the central scenario). This could be through a staff meeting or engaging with a

[122] Annual Survey of Hours and Earnings (ONS)

[123] Business Impact Target Appraisal Guidance - BEIS

[124] The wage of a legal professional is used here.

[125] HSE, 2013, Estimated time to read guidance, reproduced in RPC, 2017, Business Impact Target – Appraisal of guidance: assessments for regulator-issued guidance

summary email. Dissemination is expected to result in costs of between £0.7 million and £2.1 million (central, £1.4 million).

136. Following the methodology noted above, familiarisation costs are estimated to total between £39.7 million and £74.1 million (central, £53.6 million).

137. Previous estimates for familiarisation costs were not specifically challenged in response to the consultation IA. However, based on the qualitative evidence from engagement with in-scope platforms, the above approach reflects the following changes to previous estimates:

- The individual(s) within SMBs expected to familiarise themselves with the regulation has been changed from regulatory professionals to Chief Executives. While use of regulatory professional wages was only a proxy, this now better reflects that owners of smaller platforms are likely to be the ones who conduct familiarisation, a point noted by SMBs engaged and advised by the RPC.
- The addition of potential legal advice for in-scope SMBs.

138. Table 8 outlines the range of expected costs associated with reading and understanding the regulations:

Table 8: Reading and understanding the regulations (2019 prices, 2020 base year – 10-year PV)

	Low	Central	High
Option 1: Reading and understanding the regulations	£39.7 million	£53.6 million	£74.1 million

Transition costs

139. Table 9 sets out the total transition costs across the policy options. Details on how these costs have been estimated is below.

Table 9: Transition costs (2019 prices, 2020 base year – 10-year PV)

	Low	Central	High
Option 1: Transition costs	£87.4 million	£134.0 million	£216.0 million

140. Platforms are expected to incur the following transition costs:

- **ensuring users can report harm** - this relates to the mechanism through which users can report harm and could be as simple as a visible email address (already a statutory requirement) or a system which can triage large volumes of reports.
- **updating terms of service** - platforms may decide to assess and update their terms of service in response to future codes of practice.

- **reflecting the illegal content judgement in content moderation advice -** this relates to the requirement to assess whether content is illegal, e.g. including mens rea, as well as actus reus, where relevant and ensuring content moderators are equipped to make this judgement.

Ensuring users can report harm

Requirements

141. Under the framework, platforms will be expected to accommodate user reporting of harm and provide an avenue for user redress (challenge of content removal). User reporting and redress mechanisms are expected to vary across platforms. For example, for the smallest lowest-risk platforms, they may only be required to have an email address visible on their service (already a legal requirement under the Electronic Commerce Regulations 2002[126]) while high risk platforms may require reporting mechanisms which can handle and triage larger volumes of reporting.

Baseline

142. All available evidence suggests that most in-scope platforms already allow users to report harm. All respondents to Ofcom's VSP consultation allowed users to report harmful content with mechanisms ranging from three-dot icons to flagging buttons near the content. Through interviews with a sample of in-scope platforms, RR research indicated that all high-risk platforms and most mid risk platforms in the sample already had reporting functions and procedures for users who experienced or witnessed harm on their platforms. Many of these also tailored the options in their reporting functions to represent the categories of harm commonly reported on their sites, and to enable them to better triage reports to ensure they dealt with the high priority categories of harm first. 100% of respondents to a government stakeholder survey (out of 8 that answered the question) had reporting mechanisms in place and similarly, the AVMSD research found that most platforms allowed users to flag content for review. All platforms interviewed in the 2023 RR research had some form of user reporting and complaints process.

Cost estimates

143. Based on all available evidence, this IA expects most platforms to already allow user reporting. In line with Ofcom's findings in the context of the VSP regime, any costs are expected to be minimal, incremental, and relate to staff time[127] and ensuring reporting mechanisms remain fit for purpose, for example, simply repositioning of the organisations' email address for low risk platforms or minimally revising the triage functionality for higher risk platforms. This IA does not expect platforms to have to undergo significant redesign of online services to comply with the reporting requirement.

144. While the costs will be considered further once the code of practice has been developed, to provide an indication of the likely scale of the impacts at primary this IA assumes varying degrees of programmer time to make changes to the internal reporting mechanism:

- low risk platforms: 1 hours of programmer time for micro (rising to 2, 4 and 6 for small, medium and large respectively).

[126] The Electronic Commerce (EC Directive) Regulations 2002

[127] While there will be additional costs related to operating the user reporting system, this is considered as a compliance cost under additional content moderation.

- mid risk platforms: 2 hours of programmer time for micro (rising to 4, 6 and 8 for small medium and large respectively)
- high risk platforms: 8 hours of programmer time for micro (rising to 12, 16 and 20 for small medium and large respectively)

145. In addition to programmer time, for each in-scope platform, one hour of Chief Executive/Senior Official time is estimated for sign-off of the changes.

146. Previous estimates for user reporting costs were not specifically challenged in response to the consultation IA and the approach remains broadly the same. However, this IA allows for the possibility that a small number of in-scope firms in the baseline may not currently allow user reporting in any form. Only one platform provided estimated costs of £1,000 per year for their user reporting function. This IA conservatively assumes that between 5% - 20% of in-scope firms across low and mid risk platforms[128] may have to develop a user reporting mechanism rather than just make incremental changes (10% in the central scenario). In the absence of evidence on cost differentials across risk and size categories, these platforms are expected to incur costs of £1,000. The above approach results in a total cost of implementing or revising user reporting mechanisms in the first year of between £3.3 million - £6.1 million (central, £4.2 million). To reflect the possibility that organisations may need to make changes throughout the appraisal period to reflect decisions from the independent regulator, these costs are assumed to be incurred in each year but reduce by 50% from the second year.

147. Table 10 outlines the range of expected costs associated with ensuring users can report harm:

Table 10: Ensuring users can report harm (2019 prices, 2020 base year – 10-year PV)

	Low	Central	High
Option 1: Ensuring users can report harm	£17.0 million	£22.0 million	£31.9 million

Updating terms of service

Requirements

148. Under Option 1, all companies will be required to set terms of service for illegal content and, if relevant, protecting children. Category 1 organisations will be required to set clear terms of service in relation to the restriction or removal of user-generated content, and the suspension or banning of users on grounds related to UGC.[129]

Baseline

149. Available evidence from AVMSD research, platform engagement and the 2023 RR research indicates that terms of service are already widespread under the baseline. AVMSD research found

[128] Evidence suggests coverage in high risk platforms is universal and costs are expected to be incremental only.

[129] For Category 1 services, it should be noted that the legislation will not set what legal content is acceptable, or how journalistic and democratic content should be treated, only that these platforms set clear terms of service and enforce them.

that the most implemented user-safety measure was ‘acceptable use policies’ which large and medium sized platforms in the sample[130] considered to be fully functional at addressing critical risks. In addition, all respondents to a survey of stakeholders already had terms of service (out of 8 that responded to the question). In addition, nearly all respondents to Ofcom’s VSP call for evidence[131] had terms and conditions which prohibited the specific categories of harmful material under the VSP framework. All organisations interviewed in the 2023 RR research already had published terms of service and no organisation anticipated having to make major changes to their terms of service.

150. Also, changes to terms of service are a business-as-usual activity undertaken by platforms. AVMSD research indicated that platforms regularly update these policies in response to their users. This was supported by respondents to Ofcom’s VSP call for evidence with many platforms indicating that they regularly review and update their terms and conditions. While most platforms will already have some form of terms of service which outline acceptable use, and these are potentially business-as-usual activities, all in-scope platforms are illustratively expected to incur some incremental costs associated with assessing their own terms of service and revising them to reflect the regulator’s code of practice.

151.

Cost estimates

152. Based on an assessment of 14 of the most popular online services’ terms of service,[132] they range in length from 2,451 words to 15,260[133] with an average length of 5,976. It is estimated that 1.5 hours will be spent initially on reading, assessing, and making the changes based on an average reading speed of 200 words per minute[134], plus twice as much time to assess and re-write. One member of staff (one senior official at a wage of £38.97 for SMBs and one regulatory professional at a wage of £21.12 for medium and large platforms) is expected to read and assess the current terms of service and make the necessary changes. In addition, businesses are expected to potentially require between 1-4 hours of legal advice[135] (2 hours in the central scenario). Finally, this IA estimates one hour of Chief Executive / Senior Official time for sign-off of the changes is included in the estimates. To reflect the potential need for ongoing updates, this cost is expected to be incurred each year but reduced by 50% from the second year onwards.

153. The table below outlines the range of expected costs associated with updating terms of service:

Table 11: Updating terms of service (2019 prices, 2020 base year – 10-year PV)

	Low	Central	High
Option 1: Updating terms of service	£14.6 million	£17.5 million	£23.4 million

[130] In the AVMSD research platform size was based on the number of unique users as opposed to employees; however, with the exception of two platforms, this mapped to size definitions based on employees.

[131] Consultation on guidance for VSP providers on measures to protect users from harmful material (Ofcom, 2021)

[132] These include some of the most popular services such as Facebook, Instagram, X and TikTok.

[133] Visualizing the Length of the Fine Print, for 14 Popular Apps - visual capitalist (April 2020)

[134] Business Impact Target, Appraisal of Guidance: Assessments for Regulator Issued Guidance – Department for Business and Industrial Strategy, 2017

[135] Assumed to be given here by a legal professional at a wage of £39.23

154. Previous estimates for updating terms of service were not specifically challenged in response to the consultation IA and the approach remains broadly the same. However, based on the qualitative evidence from engagement with in-scope platforms related to the need for legal advice, these estimates include legal advice for all platforms rather than just medium and large as estimated previously.

Break-out Box 9 – since the final-stage IA: The OSA was amended to set out principles about how providers' systems and processes should approach judgements about content, including 'illegal content'. The term 'illegal content' refers to online content that is in-scope of providers' duties and that amounts to an offence that is in-scope of providers' duties. It set out a requirement on Ofcom to produce guidance on how providers can make judgements about whether content is 'illegal content' under the Act.

As of the time of writing, Ofcom's draft guidance has already been published and familiarisation with this guidance is incorporated into the estimate above. The OSA allows platforms to make this judgment in the context of automated as well as human moderation. We assume that these judgements, adapted for platforms' moderation models, will form part of the additional content moderation estimated as part of compliance costs below. However, we recognise a transition step between understanding the guidance and the judgement becoming part of existing moderation models that isn't captured elsewhere.

In the absence of better information about the process that platforms will use to meet these new criteria, we have assumed that 1-2 employees who are familiar with the Ofcom guidance will need to consider and document changes to the platform's moderation guidance. We have assumed that this will take 16-80 hours (central estimate: 40) apiece. This assumption was tested by confirming that it sits well within the time taken by an average person to type the corresponding Ofcom guidance in full. The time taken to consider and document changes is costed at the same rate as reading the guidance (i.e. CEO or regulatory officer, depending on business size).

Moderation staff will need to read revisions to their internal moderation guidance. We assume these revisions will be 10%-50% (central estimate: 20%) of the length of the original guidance. This is because the judgement will be simplified, and/or because some potentially illegal content may not be permitted on that platform – for instance, platforms that don't allow nudity will not be concerned with judging whether intimate images constitute intimate image abuse.

We assume moderation staff constitute 5%-15% of staff at medium and large platforms and use wage assumptions which match the costing of disseminating other guidance above.

The estimated costs for business to comply are reported in the table below –

Table 12: Illegal content judgement (2019 prices, 2020 base year – 10-year PV)

	Low	Central	High
Option 1: Considering and documenting changes to moderation	£15.8 million	£39.6 million	£79.1 million
Option 1: Disseminating changes to moderation	£0.3 million	£1.4 million	£7.2 million
Total	£16.1 million	£41.0 million	£86.4 million

Illegal content judgement

Requirements

Compliance costs

155. For appraisal purposes, it is assumed that legislation enters into force in 2024 and platforms are expected to comply with the codes from 2025. This IA therefore assumes that compliance costs will begin from the second year of the appraisal period. The table below sets out the total compliance costs across the policy options. Details on how these costs have been estimated is below.

Table 13: Total compliance costs (2019 prices, 2020 base year – 10-year PV)

	Low	Central	High
Option 1: Compliance costs	£1,430.0 million	£2,130.0 million	£2,860.0 million

156. Platforms are expected to incur the following costs associated with compliance:

- ○ **conducting risk assessments** - this relates to the requirement to carry out an illegal content risk assessment and 'if likely to be accessed by children' to carry out a children's risk assessment.
- ○ **undertaking additional content moderation** - the OSA does not require additional content moderation; however, it is likely that platforms will increase resources in this area to comply with the duties.
- ○ **employing age assurance technology -** in complying with the child safety duties, some higher risk platforms are likely to adopt age assurance (and specifically age verification) technologies.

Break-out Box 10 - since the final-stage IA: Certain platforms will be required to use highly effective age verification (AV) or age estimation (AE) to protect children from primary priority content that is harmful to children and provider published pornography. This is a change from the original drafting of the Online Safety Bill which would have required services to take proportionate measures (set out in Ofcom's codes and guidance) to prevent children from seeing pornography and other content that is harmful to children. The OSB originally listed age verification and age estimation as examples of measures that *could* be taken.

The approach taken in the final-stage IA does not fully monetise the impact of potential deployment of AV or AE. The reasons for this are covered in more detail in paragraph 178 of that document, but broadly it is due to the lack of evidence on approach and solutions.

The final stage IA provided an indication of the likely scale based on two approaches.

1. Presenting individual platform costs based on an industry pricing survey from January 2022, with different regulatory requirements.

2. Top-down user-modelling, which captures users not based in the UK.

The first approach was to provide an indication of the likely scale of costs. Platforms will face differences to total monthly visits and the amount that would have to be verified, the illustrative scenarios, show how these costs differ during negotiations with third parties.

The second approach does not monetise the direct cost to UK-based businesses but all economic impact, using the total amount of people across the appraisal period that would need to be verified.

The approach taken in the final stage IA already assumed that all relevant platforms (i.e. not a percentage) would employ age assurance by taking a user-based approach, rather than considering other, less-well understood approaches to achieve these requirements. The modelling already assumes that such platforms would need to use age verification. Age estimation was not modelled, as data was not available, but either option is available to use for compliance, so this specific requirement is reasonably captured (or marginally over-estimated if AE is generally less costly) within the existing estimates.

The types of AV or AE that will meet the bar of effective compliance still need to be set out by Ofcom in codes/guidance.

- **transparency reporting** - this relates to Category 1 platforms producing annual published reports on platform harm and related actions taken.
- **fraudulent advertising duty** - as part of complying with the fraudulent advertising duty, it is likely that in-scope platforms will conduct customer due diligence (CDD) on advertisers.
- **user verification and empowerment duties** - this relates to the requirement on Category 1 platforms to offer optional user verification and provide user empowerment tools.
- **assessing impacts on freedom of expression and privacy** - this relates to publishing an assessment of impacts on freedom of expression and privacy and keeping this updated.
- **reporting online CSEA to the NCA** - this refers to the cost of reporting identified CSEA content to the NCA.

Conducting risk assessments

Requirements

157. All platforms in scope will be required to produce and publish a risk assessment. Platforms will be expected to assess risks corresponding to the type of content and activity a platform is required to address. In practice, this means the vast majority will only be required to assess risks related to illegal content and activity and - if likely to be accessed by children - content and activity which is harmful to children.

Break-out Box 11 - Since the final stage IA: Category 1 services no longer need to assess risks related to content deemed ‘harmful to adults’, nor notify Ofcom of any ‘harmful’ content to adults present on the service, other than the previous priority content list. However, these services will be required to undertake an assessment of the incidence of relevant content on the service, and have a duty to supply Ofcom with their assessment of the likelihood of adult users encountering content relevant to the user empowerment tools, and a duty to summarise in their terms of service the findings of their most recent assessment.

The OSA specifies content where these tools would apply (regulated content, but also content that encourages, promotes, or provides instructions on suicide/self-harm, eating disorders, and abuse or hate speech based on protected characteristics). It also specifies the scope of the assessment.

Both the OSB and the OSA risk assessments consider the likelihood of users encountering this content, though the “likely to encounter" assessment (OSA) doesn’t require an assessment of “the nature, and severity, of the harm” (one of eight elements of assessment specified in the OSB). This will mean that OSA risk assessments may be more limited than the equivalent OSB assessments in terms of the content assessed. However, the two assessments will be very similar in terms of the scope and length (and therefore cost of compliance), and given the small number of platforms involved, these changes will not lead to a significant difference in costs.

Baseline

158. From engagement with industry, under the baseline, many (especially higher risk platforms) already conduct internal risks assessments. Platforms use these risk assessments to prioritise user safety resources and to ensure emerging risks are identified. In addition, while not an explicit requirement of the VSP regime, Ofcom already strongly encourages platforms in its guidance[136] to assess the level of risk on the platform.

Cost estimates

159. Risk assessment cost information provided by platforms is limited. Only two platforms provided the cost of producing a risk assessment but these both related to risk assessments they already produced and ranged from £2,500 to £10,000. Given that many organisations already produce these, this figure would overestimate the incremental cost. Based on qualitative evidence provided by platforms, the cost to business and effectiveness of risk assessments are likely to depend on:

- ○ **expectations on platforms**: that is the need to minimise the administrative burden on platforms required to assess risk across multiple duties.
- ○ **focus of risk assessment**: risk assessments need to consider the range of measures a platform has in place in relation to its specific risks. Some measures will be more important to some platforms than others, depending on the type of content they host and whether they are likely to be accessed by children.
- ○ **alignment with international and domestic regulations**: the need to ensure expectations on platforms align with current risk assessment practices which are conducted in compliance with other relevant regulations, for example AVMSD, Digital Services Act, and others.

160. Ofcom will set out the steps platforms can take to comply with the requirement to produce risk assessments in future codes of practice. Given that cost information is limited in the context of risk assessment, estimates presented in the previous IA are retained to provide an indication of the likely scale of impact at this stage. As discussed in the break-out box above, the introduction of the “likely to encounter” assessment is expected to impose a similar cost to the adult risk assessment requirements that have been removed from the OSB, and therefore these estimates are unchanged from the previous IA. These are based on estimates from the Networks and Information Systems Regulations 2018 (NIS)[137] used as a proxy for the cost of producing an online harm risk assessment (or revising an existing one).

161. To estimate the expected incremental costs associated with producing risk assessments, the NIS assumed that reports are produced by IT professionals and that evidence and reports are reviewed and discussed by senior management and legal professionals. Estimates proxied here include 1.5 hours of time for a legal professional (at a wage of £23.31) and 2 hours for a senior manager (at a wage of £22.08) for micro and small platforms, rising to 5 and 7 for medium sized platforms and 10 and 14 for large platforms respectively. In addition, in line with the possibility that some - while rare - may not currently assess risks on their platforms, this IA illustratively estimates

[136] Video-sharing platform guidance (Ofcom, 2021)
[137] The Network and Information Systems Regulation 2018 - DCMS (April 2018)

that between 0%-5% of in-scope mid risk and high risk platforms (2.5% in the central scenario) may incur costs as large as those provided by platforms in the context of risk assessments they already produce. While comprehensive evidence of baseline risk assessment practices is not available (given the scope of the regulation), all available evidence suggests some element of assessing risks on platforms is widespread across platforms. Based on the limited cost information available, this IA estimates these platforms could incur costs of £6,250 (or the midpoint of the range provided above). As the cost of producing a risk assessment is likely to reduce once a platform has reported for the first time, the cost is incurred each year for all businesses but expected to reduce by 50% from the second year of compliance onwards.

162. The table below outlines the range of expected costs associated with assessing the risk of harm on the platform:

Table 14: Risk assessments

	Low	Central	High
Option 1: Risk assessments	£12.9 million	£28.1 million	£43.3 million

Undertaking additional content moderation

Requirements

163. The core duties require all in-scope platforms to put in place systems and processes to address illegal content and - if likely to be accessed by children - to protect children from content which is harmful to them. There is an additional duty on Category 1 organisations to consistently enforce their terms of service relating to the restriction or removal of user-generated content, and the suspension or banning of users on grounds related to UGC. To protect freedom of expression and privacy, in fulfilling their safety duties, Category 1 platforms will have to put in place clear policies to protect journalistic content and content of democratic importance.

164. While platforms will fulfil their safety duties in many ways, Option 1 is expected to result in some platforms requiring additional content moderation. This could be through hiring additional human content moderators, employing automated content moderation systems, or a combination of both. Category 1 services must have the appropriate systems and processes in place to ensure their terms of service are enforced consistently, which may include additional training for their moderators and/or investment in more effective moderation technology. As with other aspects of Option 1, requirements on in-scope platforms will be proportionate and risk-based with the largest highest risk platforms expected to do more than the smallest lowest risk platforms.

In determining what is proportionate relating to duties to protect journalistic content and content democratic importance, the size and capacity of the provider of a service, in particular, is relevant. We expect that the costs associated this duty for smaller services, or those with lower capacity, will be on the lower end. Where services may be Category 1 and therefore required to comply with this duty, but do not have journalistic content on their service or content of democratic importance, we expect that ongoing compliance costs will be minimal. Ofcom will produce a code of practice on these duties.

165. For duties to protect news publisher content, we recognise that there may be instances where a service is designated as Category 1 but does not have any, or has a minimal amount of, news publisher content. In such instances, we expect that the ongoing compliance costs to services will be minimal. Ofcom will also produce guidance on the duties to protect news publisher content.

166. In addition, the terms of service duties for Category 1 services are subject to proportionality, which, in particular, requires that the size and capacity of the service provider must be taken into account (among other things). The existing breadth and extent of a service's existing rules may also affect the initial cost. We therefore expect costs to providers to vary on this basis. Ofcom will produce guidance for providers of Category 1 services to assist them with complying with these duties.

Baseline

167. The vast majority of organisations in scope will already be taking some action to reduce the risk of online harm on their services. Many of the largest platforms already employ large teams of content moderators and operate sophisticated automated moderation systems - Meta for example, employs over 15,000 human moderators across the globe to review potential violations on Facebook and Instagram[138] and Facebook reportedly employed an additional 186 moderators in response to Germany's NetzDG.[139] Both RR's research and EY's assessment of VSPs demonstrated that resources spent on moderation activities in the baseline vary greatly from less than £1,000 for the smallest lowest risk platforms to over £1 billion for the largest platforms.

168. RR's research found that some organisations consider it unlikely that the regulation will result in significant incremental costs. This is because of increasing user expectations over the safety of online communities, requirements set by advertisers and third-party suppliers, and to remain competitive in the industry. In support of this, EY's research on VSPs also found that most platforms in their study indicated that compliance with AVMSD (which includes many similar principles to the OSA) was not expected to result in very material investment.

Cost estimates

169. Compliance costs related to potential additional content moderation will depend in full on the specific requirements set out in future codes of practice. At this primary stage, this IA is only able to provide an indication of the likely scale of impacts. Ofcom will consult on future codes and produce IAs once the specific requirements are set. The vast majority of platforms engaged are unable at this stage to provide estimated costs associated with potential additional content moderation until they know what they will be required to do. On this basis, given that previous estimates for content moderation were not challenged in response to the previous IA, the approach remains broadly the same and is considered a reasonable indication of scale of potential future costs.

170. RR interviewed a sample of in-scope platforms[140] to determine: their current practices and processes to mitigate the risks of online harm occurring; where available, quantification of the associated resources and costs of practices and processes to identify and prevent harm; and how these costs and resources would change if duties were enforced.

171. A strategic sample of 118 organisations were contacted for interviews, and 25% (or 30 organisations) agreed to and completed an interview. This sample included: social media (13 of the 16 most used social media sites in the UK); forums; review sites; blogs; gaming; retail; P2P marketplaces; volunteering; official fan sites; job searching; fan fiction; search engines; accommodation searching; adult entertainment; and dating sites.

[138] The people behind Meta's review teams (Meta, 2022)

[139] NetzDG Transparency Report (Facebook, 2022)

[140] Under the policy position as set out in the OHWP and not the subsequent exemptions.

172. To estimate the incremental cost of compliance, the analysis discounts organisations that already have sufficient content-moderating systems and processes in place and organisations that - due to being very low risk or smaller mid-risk platforms - would likely not be expected to take additional actions in moderating content. Based on findings from the interviews, the percentage of in-scope platforms requiring extra spend on content moderation is conservatively estimated to be between 20%-30% of high risk in-scope organisations (25% in the central scenario) and between 5%-15% of medium and large mid-risk organisations (10% in the central scenario).

173. Among interviewed organisations in the RR research that expected to require additional moderation, estimates for the incremental cost of regulation ranged from 1% of turnover[141] (the lowest estimate) to 15%[142] (the highest). These estimates were provided in the context of platforms' interpretation of the OHWP, that is, the cost of additional content moderation for platforms required to address all categories of harm in the OHWP including extra protections for children. This IA therefore takes the midpoint of this range (7.5% of turnover) to represent the cost of additional content moderation for Category 1 organisations (those expected to address all categories of harm). Turnover estimates used come from average turnover by business size band in BEIS' BPE. Sensitivity analysis is conducted on the full range of estimates provided by businesses.

174. For non-Category 1 platforms - those not required to deal with content or users in accordance with terms of service - costs are expected to be lower than those incurred by Category 1 platforms. To calculate the cost to these, data from several large social media platforms' transparency reports on the volume of actioned content/accounts (content/accounts which were removed or minimised due to breaking the terms of service) are used as a proxy[143]. In the transparency reports, actioned content/accounts is split into a number of broad harm categories which were assessed as either:

- ○ likely to be considered illegal, or
- ○ likely to be considered legal but harmful to children or against platforms' terms of service.
- ○ other (categories such as 'spam' or 'fake accounts' which were not considered as an online harm in the RR research based on categories of harm within the OHWP).

175. Using the volume of actioned content/accounts in each category, an approximate percentage split of illegal vs legal actioned content/accounts was estimated. Four social media platforms' reports were assessed[144] and 2021 data was used. This approach has the following limitations:

- ○ it assumes that the cost of content moderation is linearly related with the volume of content. For organisations that use automated moderation this may not be the case;
- ○ it is difficult to determine whether content actioned under the broad categories in the transparency reports would be considered illegal or harmful - most of the reports do not break the data down in this way. For example, X uses a 'hateful conduct' category which - referring to X 's policy on the topic - is likely to contain both illegal and harmful content. For these categories, the volume of content actioned was split

[141] The lowest estimate was actually 1% of operating costs which would likely be lower than 1% of turnover; however, for ease and given data availability, turnover is used as a proxy.

[142] The exact figure given in the interview was 14% of revenue which was rounded and due to data availability, turnover was used as a proxy.

[143] In the sample of transparency reports, accounts actioned is used for X as it is more representative for the category "terrorism or violent extremism". This is due to the way X deals with content in this category.

[144] Facebook, Instagram, X, and Snapchat

equally between illegal and harmful.[145] However, some transparency reports provide additional clarity into each category, for example Snapchat provides a breakdown of the amount of “child sexual exploitation and abuse imagery (CSEAI)” in their “sexually explicit content” category. This allowed the split to be rebalanced to an 83-17% legal/illegal split.[146] Using this approach assumes that the only illegal content in this split was contained in the 17%, which due to broad categories may not be the case;

- ○ it is not clear that the four social media platforms’ transparency reports are representative of the wider sample of in-scope businesses. It may be the case that illegal content represents a smaller proportion of overall content on platforms or vice versa;
- ○ it is assumed that the cost of moderation from the RR research remains unchanged with the inclusion of ‘other’ harm categories;
- ○ the ‘other’ category makes up between 4-95% of total content across assessed platforms and varies the most across platforms, both in terms of what content is included in its category and its representation of total content. Therefore, due to this variation and the assumption that the cost of additional moderation will not change with its inclusion, it has been excluded from the calculation.

176. The percentage of actioned content in categories assessed as being likely illegal ranged from 18%-29%.[147] To reflect the costs to platforms not designated as Category 1 (those which are not required to deal with content or users in accordance with terms of service), given the ranges above, this IA estimates that the relative costs to these platforms would be approximately 23.5% of the relative costs to Category 1 platforms or 1.8% of turnover. Sensitivity analysis is conducted on the full range of estimates in the risks and sensitivity section.

177. Table below outlines the range of expected costs associated with additional content moderation:

Table 15: Additional content moderation (2019 prices, 2020 base year - 10-year PV)

	Low	Central	High
Option 1: Additional content moderation	£1,340.0 million	£1,920.0 million	£2,500.0 million

178. Three platforms provided cost information as part of the government’s stakeholder survey, relating to the annual cost of user safety measures they currently undertake (i.e. not because of regulation). Costs provided came from the top two size categories (medium or large) and top two risk categories (mid or high) and were all below £1m per year. Estimates provided by platforms in the AVMSD research varied widely from hundreds of pounds for the smallest platforms to £1.5 billion for the largest VSP. Except for a handful of the largest and highest risk businesses, for those expected to undertake additional content moderation, the per platform costs under the central scenario above would represent a doubling or more of current content moderation costs which is likely to be significantly conservative and potentially an overestimate. While the range of estimates only reflect the percentage of platforms expected to incur costs, the per platform cost - in terms of percentage of revenue - is also tested in the sensitivity section.

[145] Split categories include Facebook, Instagram, and Snapchat’s hate speech category. X, Facebook and Instagrams violence category which includes both threats of violence and glorification of violence. Snapchat’s sexually explicit content category.

[146] This comes from the Snapchat transparency report January 1, 2022 – June 30, 2022.

[147] The previous IAs used the same methodology with 2019 and 2020 data which resulted in comparable findings with content categories assessed as likely illegal ranging from 15-33% and 14-36%.

Employing age assurance technology

Requirements

179. While specific steps platforms can take to comply will be laid out in future codes of practice and regulator guidance, it is clear that some platforms will need to implement age assurance technologies to comply with the core child safety duties. There are also specific duties which require user-to-user providers to use age verification or age estimation to prevent children from encountering primary priority content that is harmful to children, where this content is not prohibited under its terms of service. In addition, pornography publishers must use age verification or age estimation to prevent children from encountering provider published pornography on their service.

180. Age assurance refers to any method to establish the age of a user online. Age verification is one type of age assurance method, which provides the highest level of confidence in the age of a user. It commonly relies on officially provided data or hard identifiers, such as a credit card or passport. For this reason, it is best suited to 18 years+ services and content, rather than providing access for children who often do not have suitable documents. Age estimation technologies are commonly AI based approaches that use biometric or behavioural data to estimate the age of the user. They are commonly more effective when the 'challenge age' is set higher than the target age e.g. 'Challenge 25' for a target age of 18 for the purchasing of alcohol. Commonly where the age estimation solution identifies that the user is close to the target age, they are referred for further checks to determine their actual age. Both age verification and age estimation technologies can be highly privacy preserving, and third party age estimation providers have the ability to provide a yes/no token to the service without sharing any personally identifying information. Age assurance technologies are important tools that enable companies to take steps to protect children from online harm, including both legal but harmful and illegal content and activity, for example, protecting children from grooming.

181. The specific child safety duties in the OSA apply to platforms which are "likely to be accessed by children". This approach has been established by the legislation underpinning the Information Commissioner's Office's (ICO) 'Age Appropriate Design Code' (section 123 of the Data Protection Act 2018) with regards to protecting children's data. Consistency across regulations reduces additional burdens on businesses, many of whom will already have taken steps to comply with the Age Appropriate Design Code. Some high risk services which are likely to be accessed by children will be required to know the age of their users to provide them with appropriate protections, and therefore may choose to implement age assurance technologies[148] to do this. As above, providers which allow primary priority content on their service will need to use age verification or age estimation to prevent children from accessing this content on their service. In addition, as part of the pornography provision, services that publish pornography will have to use age verification or age estimation to ensure that children are not able to access this type of content.

182. Without the pornography provision, platforms in scope of the core child safety duties would only be required to protect children from user-generated pornographic content. With the addition of the pornography provision, the intention is to minimise potential gaps in regulatory coverage and bring into scope children's access to non-user-generated pornographic content. This impact assessment focuses on the outcome, namely the implementation of age assurance technologies resulting from the OSA in its entirety, regardless of whether it is a result of the core duties or the published pornography provision. In its development of future codes of practice and regulator

[148] It should be noted that the codes of practice are unknown; however, at primary stage, it is reasonable to believe that some platforms may be required to introduce age assurance systems which could include age verification (if they do not operate them already) under this part of the duties.

guidance, Ofcom will further consider the separate but related impact on businesses in scope of both the core child safety duties and published pornography provision.

183. Pornography is hosted on a range of platforms. Some of these platforms are user-to-user services and in scope of the core duties. Others will only be in scope of the OSA because of the pornography provision. While there is no definitive study, the BBFC estimates that there are between 4-5 million dedicated pornographic websites accessible in the UK. However, the number of businesses that this represents is much lower as companies often operate multiple sites. Further, the number of UK-based businesses that this represents is even lower, with most sites operated by companies outside of the UK - the vast majority being in the US. The table below outlines the main types of platforms with the potential to currently host or publish pornography. It is important to note that not all platforms within each category will host or publish pornography and many will explicitly prohibit it as part of their terms of services. The table below should be viewed as a conservative upper bound estimate of platforms with the potential to host pornography (and therefore, the potential to implement age assurance technologies in response to the OSA):

Table 16: UK based platforms with the potential to incur age assurance costs

Type of platform	**Number of UK-based businesses 2023**	
	Total	**Potential to incur age assurance costs**
Social media platforms: The vast majority, if not all pornographic content on social media platforms is user-generated and in scope of the core duties.	225[149]	225 This is conservative as only those that host pornography and/or are likely to be accessed by children have the potential to incur costs. Many prohibit this content as part of their terms of service.
Search engines: These platforms are in scope of the core duties.	1,394[150]	1,394 This is conservative as only those that host pornography and/or are likely to be accessed by children have the potential to incur costs.
VSPs: Subject to a combination of the core duties and the pornography provision depending on the type of pornographic content (user-generated vs non-user generated)	20[151]	0 VSPs that host pornography are already required to prevent children from accessing sexually explicit content under AVMSD.
Dedicated pornography providers: Subject to a combination of the core duties and the pornography provision depending on the type of pornographic content (user-generated vs	12	12 See 'Platforms in scope' for further details on estimating the number of

[149] Social Media Platforms in the UK - Market Research Report (IBIS World, 2022)

[150] Search Engines in the UK - Market Research Report (IBIS World, 2022)

[151] Notified video-sharing platforms (Ofcom, 2023)

Type of platform	Number of UK-based businesses 2023	
non-user generated)		UK-based pornography providers.
Image sharing platforms: The vast majority if not all pornographic content on image sharing platforms is user-generated and in scope of the core duties.	Unknown	Unknown There is no definitive data on the number of UK-based image sharing sites. It is reasonable to assume this number is low when considering only UK-based businesses.
VoD platforms: These are **not** in scope of the OSA, and pornographic content on these sites will continue to be regulated under the video on demand regime.	c.150	0 VoDs are not in scope of the OSA.
Total number of in-scope platforms that could potentially incur some amount of age assurance costs in 2024[152]		1,680

Baseline

184. In relation to user-user platforms in scope of the child safety duties, of those self-declared as likely to be accessed by children in a survey of stakeholders, three out of four that answered the question said that they already employed age verification. It should be noted that for this question the term ‘age verification’ was not defined and therefore it is possible that platforms selected ‘age verification’ when in reality they currently employ weaker forms of age assurance, such as self-declared age, which on its own is unlikely to be considered an appropriate child safety measure. While most platforms designated as Category 1 services are expected to already employ some type of process to attempt to determine the age or age range of their users, this could range from robust age verification controls to a simple self-declaration (which on its own would not be considered age assurance). In addition, the AVMSD research highlighted that coverage and perceived effectiveness of current age assurance measures among small and medium sized platforms was lower than larger platforms.

185. When it comes to dedicated pornography sites, evidence suggests that age assurance technologies (and in particular age verification) are rare and that children can easily access pornographic content on these sites. BBFC research in 2020 indicated that of the top 200 pornography sites (which together account for over three quarters of UK traffic to adult sites), only 4.5% have existing mechanisms in place that may prevent, deter, or delay children accessing the site before displaying any pornographic content. Even these measures, which included having to sign up or register payment details, are not significantly robust given that children as young as eleven may have their own debit card - no sites in the top 200 required credit card-only payments. 14.5% of the top 200 sites have a pop up warning indicating that access is reserved for over-18s but this can easily be ignored by children that are intentional viewers of pornography. In addition, Ofcom’s report following the first year of the VSP regulation found that smaller adult video-sharing

[152] In line with the rest of this impact assessment, the number of potential platforms is uplifted by the average growth of UK businesses (3%) for an implementation date of 2024. These figures are lower than Ofcom estimates, at least in part because IAs are concerned with the impact on UK firms, while Ofcom considers purely overseas firms serving the UK market.

sites based in the UK do not have sufficiently robust access control measures in place to stop children accessing pornography.[153] Based on BBFC engagement with the adult industry, the current lack of age assurance - even when the industry has stated its willingness to adopt these technologies - appears to be the result of competitive concerns and the potential commercial impact if this requirement is not mandatory across all services. This was also raised by platforms in the 2023 RR study. It is therefore important that the child safety duties and pornography provision together apply to all pornographic content accessible to UK users.

Cost estimates

186. It is not possible at this stage to fully monetise the impact of the potential employment of age assurance solutions by some platforms in scope of the child safety duties and pornography provisions under Option 1. This is because:

- the platforms required to employ age assurance controls under the core child safety duties, the type of controls required, and the types of age verification or age estimation that will meet the bar of effectiveness to meet the specific duties will be set out in future codes of practice and regulator guidance (themselves subject to IAs).
- types of age assurance solutions, their accuracy and their availability are rapidly evolving. The government and industry expect technology to greatly improve in this market and there are significant opportunities for cost reductions between now and implementation of the OSA.
- different platforms are expected to take different approaches to meeting their duties under the OSA. For example, even within those likely to implement age assurance, some larger platforms, in particular the largest social media platforms, may develop in-house solutions while smaller platforms could employ off-the-shelf solutions which are cost effective and readily available. In addition, it may be the case that costs instead fall on the user. Evidence from the BBFC's engagement with industry suggests that most pornography sites were expected to use certified third-party solutions to minimise the risk of privacy concerns. Some of the larger pornography platforms have developed their own solutions but these are run as separate businesses (and still considered third-party).
- there are also solutions offered to both companies and users at no price but may contain advertisements[154] to create revenue for the age assurance provider or include a number of monthly free checks before paying a monthly subscription.[155]

187. While it is difficult at this stage to provide an accurate assessment of direct business costs, this IA presents a comprehensive indication of the likely scale based on two separate approaches, namely presenting individual platform costs based on an industry pricing survey conducted in January 2022 and top-down user-modelling scenarios.

188. To better understand individual platform costs, a selection of UK-facing providers of third-party age verification solutions were engaged through a survey distributed by the Age Verification Providers Association (the UK's industry body for age assurance providers). Illustrative costs were provided based on several example platform scenarios. These costs should only be considered as providing an indication of the likely scale of costs and the actual price paid will be the result of standard business negotiations between platforms and third-party services. In reality, costs will

[153] Ofcom's first year of video-sharing platform regulation (Ofcom, 2022)

[154] https://ageverify.com/

[155] https://www.1account.net/business-demo

depend on a number of factors and nuances not captured in the below example scenarios. The table below sets out the findings of this engagement with industry.

Table 17: Illustrative platform scenarios

Scenario	Description
Platform A	● 25,000 unique monthly UK users. ● 1% of users are assumed to be new each month and have not verified their age previously. ● 180,000 total monthly visits.
Platform B	● 100,000 unique monthly UK users. ● 1% of users are assumed to be new each month and have not verified their age previously. ● 730,000 total monthly visits.
Platform C	● 1 million unique monthly UK users. ● 1% of users are assumed to be new each month and have not verified their age previously. ● 7.3 million total monthly visits.
Platform D	● 4 million unique monthly UK users. ● 1% of users are assumed to be new each month and have not verified their age previously. ● 29.2 million total monthly visits.

189. The above platform scenarios are illustrative only and range from what would be considered a relatively small platform to a relatively large platform.

190. **Per check costs**: Costs per check ranged from less than 1p to more than £1. The large range reflects the variety of approaches and methods available to platforms. The only criterion given within the illustrative platform scenarios was that the approach should be able to determine whether a user is over 18 and meet standards defined by the British Standard Institute (PAS 1296:2018). Even within this criterion, AV providers offer an extensive range of approaches depending on regulatory requirements. While there is a large range, most of the per check costs provided were 10p or lower per age check with the upper bound of the range reflecting a suite of different approaches. We have used 10p (in 2021 prices) as an estimate for the cost of each check. Some estimates did reduce as volumes increased, with some providers' per-check cost lower for Platform D than for Platform A for example; however, others remained consistent throughout.

191. **Monthly costs**: AV providers were also asked to estimate the monthly costs for each illustrative platform based on per check costs or any other monthly pricing option. Information provided here was even more dependent on regulatory requirements and approaches taken by platforms. For example, costs depend on whether the platform would verify users each time they access the site or only new users. Some platforms provided monthly costs based on the per check costs outlined above, i.e. first month costs would include verifying the existing user base and from month two onwards, only new users are verified. Across a 12 month period, monthly costs provided for Platform A averaged just over £600, rising to just over £1,800 for Platform B. Monthly costs were estimated to be between £10,000 and £40,000 for Platform C and between £30,000 and £90,000 for Platform D.

192. There may be additional costs (not captured above) of integrating age verification solutions within each in-scope platform. Many third-party providers offer support packages to businesses with step by step instructions and developer support. As part of Yoti's submission to Ofcom's call for evidence on the VSP regime, it noted that it takes approximately half a day for a digital platform to integrate with Yoti's backend system.[156] Assuming between one and three developers are required and based on median developer wages, this could result in additional platform costs of between £108 and £324.

193. Costs provided by AV providers should be treated with caution as:

- technology in the age assurance market is moving quickly and the industry expects significant improvements in accuracy and reductions in cost in the short to medium term.
- there are significant movements towards interoperability with solutions that can work across several platforms. While this is not an established approach yet it is something that the government is supporting through its work on standards, including the Digital Identity and Attributes Trust Framework, which will support interoperable solutions to function. As such it is possible that platforms would not need to establish the age of every user as many will have had their age verified previously.
- the platform scenarios presented to industry are by nature static and artificial. Actual costs will reflect the outcomes of standard business negotiations between platforms and third-party providers.
- the AVPA noted that some AV providers would likely offer heavily discounted fees for smaller clients and start-ups.

194. Given the uncertainties and limited data, it is not possible at this stage to monetise the direct cost to UK-based businesses. However, it is possible to demonstrate the totality of economic impacts by taking a top-down user-based approach.

195. To estimate the economic costs of age assurance duties, this IA considers the population intentionally accessing pornography and the number of age assurance checks each person may generate, considering the number of sites someone may visit and the degree of interoperability of age assurance between sites. Together these give an estimate of total checks each year. This estimate was multiplied by the per-check cost to derive an indicative estimate for aggregate cost for age assurance checks each year.

196. Estimates from Ofcom and Revealing Reality for the proportion of adults[157] and children aged between 11-17 years old[158] that intentionally access pornography are applied to 2020 ONS population data.[159] Year-on-year, population is assumed to grow in line with average population growth between 2000 and 2020 (growth rate of 0.65% per year).[160] This modelling estimates that with an implementation date of 2024, there will be on average approximately 27.4 million unique adults and 1.7 million unique children intentionally accessing pornography each year across a ten-year appraisal period.

[156] Yoti response - Ofcom's call for evidence
[157] Online Nation 2021 report (Ofcom)
[158] Young People, Pornography and Age Verification (Revealing Reality, 2020)
[159] ONS Population Estimates (ONS, 2020)
[160] Population growth - United Kingdom (World Bank, 2020)

197. While there is no data on the average number of pornography sites visited by each unique user, data from BBFC indicates users visiting more than one pornographic site,[161] and evidence from VSPs more generally suggests that people tend to use a limited number of platforms to view videos.[162] On this basis, this impact assessment conservatively estimates that individuals accessing pornographic sites do so on average on five separate sites. Verification of age is assumed to last for 12 months before a user is asked to complete the process again. AV providers that were engaged as part of this impact assessment noted that this would be a decision for the platform and depend on the regulations at the time. Users may be provided with a 'token or credential' and only be verified once across the period, whereas it is also possible to verify a user every time they access the site.

198. As noted earlier, there are movements in the age assurance industry towards interoperability where, once age assured on one site, a user would not need to be assured again even when accessing a different site. The possible levels of interoperability in the age verification market are represented in the low, mid and high cost estimates. The low estimate assumes complete interoperability (where verification is required only once), the mid estimate assumes moderate interoperability with each unique user undergoing verification twice across the five sites, and the high estimate assumes no interoperability, i.e. five sites visited result in five checks.[163] This modelling estimates that there will be on average 29 million - 145 million (mid estimate: 58 million) verifications each year across the appraisal period. These estimates change slightly from year to year, based on population change.

199. Based on per check costs provided by AV providers, a cost of 10p per check (in 2021 prices) is applied to the total number of verifications (starting from year 2) resulting in total costs of between **£18.4 million and £91.8 million (central estimate = £36.7 million)** in present value terms across the ten year appraisal period. It is important to note that this user-based modelling represents total costs to online platforms, including platforms based outside the UK and platforms operated by individuals as opposed to businesses. While it is not possible to estimate the direct cost to UK-based businesses only, it will be much lower than estimated here given the geographic distribution of pornography providers. Further work will be done to refine business costs as part of Ofcom's development of future codes of practice and regulator guidance, including consultation with industry.

Table 18: Employing age assurance technology (2019 prices, 2020 base year - 10-year PV)

	Low	Central	High
Option 1: Age assurance	£18.4 million	£36.7 million	£91.8 million

Transparency reporting

Requirements

200. Option 1 requires platforms to produce annual transparency reports if they are designated as Category 1 (major user to user platforms), Category 2A (major search services) or Category 2B (other categorised user-to-user services). Thresholds will be set out in secondary legislation and will

[161] BBFC Further Research on Traffic to and Functionality of Adult Sites (BBFC, 2020)
[162] Understanding how platforms with video sharing capabilities protect users from harmful content online (EY, 2021)
[163] Interoperability in the model is varied by decreasing the average number of sites visited by 5 in the high estimate to 2 in the mid estimate and 1 in the low estimate. This reflects the number of times a user requires verification across the five separate sites they use to access pornography.

be based on factors including a platform's number of users and its functionalities. This IA estimates that between 30-40 platforms will be required to produce transparency reports.[164] In line with the wider requirement placed on the regulator to act in a proportionate and risk-based manner, transparency reporting requirements will differ between the different types of platforms who are required to report. The specific information that they will need to include, will be left to the regulator and will differ between platforms.

Baseline

201. Based on available baseline evidence, many large high-risk platforms already produce transparency reports. Three out of four large high-risk platforms that responded to the government's stakeholder survey already produced these, and it is clear from subsequent engagement that many do (through NetzDG requirements for example or just best practice). The vast majority of major social media companies already produce these, including granular data on harm, content removal, and content reinstated following challenges (see Facebook, YouTube, Instagram, X and others).

Cost estimates

202. Estimates presented in the previous IA were not challenged with cost evidence supplied by platforms, and they still represent a reasonable estimate for the incremental cost of transparency reporting - that is the cost of potential revisions to existing reporting practices. However, qualitative evidence from recent engagement with in-scope platforms highlights some of the key cost drivers that will influence the scale of the regulatory burden, these are:

- **alignment with international regulations**: the more that reporting requirements align with other international regulations (both current and planned) the less burdensome this will be for platforms.
- **alignment with current reporting practises**: as above, the more these requirements align with current transparency reports produced by platforms the lower the cost.
- **flexibility in terms of metrics presented**: one platform noted that the cost of reporting is trivial compared to the cost of collecting data not currently collected. The right balance between flexibility for platforms and ensuring important metrics are presented (potentially in different ways by different platforms) is key to minimising costs.
- **engagement between Ofcom and platforms**: year-on-year changes in key metrics presented in transparency reports could be due to external factors rather than solely changes in the level of harm. It will be important for Ofcom to work closely with platforms to understand the information presented and external trends.

203. To indicate the likely scale of the cost of this activity, this IA uses estimated costs from the transparency reporting requirements under Germany's NetzDG which were expected to be €50,000 (approximately £45,000).[165] Estimates provided for NetzDG are a reasonable proxy for the transparency reporting requirements under the OSA. The cost of this activity is likely to be front-loaded, especially for platforms without appropriate systems already in place - to reflect this, the cost of transparency reports is expected to reduce by 50% from year 2 onwards.[166]

[164] For costs, the midpoint of the range is taken.

[165] Act improving law enforcement on social networks [Netzdurchführungsgesetz – NetzDG] - European Commission (2017)

[166] If the information required from platforms under the reporting requirements is changed frequently throughout the appraisal period, it is possible that costs could increase back to year 1 estimates.

204. While the central estimate remains unchanged since the previous IA, the table below outlines some additional reporting costs gathered from other UK reporting requirements. Estimates below related to corporate governance reform and climate-related financial disclosures by publicly quoted companies form the low and high estimates as - outside of Germany’s Network Enforcement Act which is the most analogous - they are most like reporting requirements under the online safety framework in terms of the focus on data and metrics.

Table 19: Comparison of reporting costs

Reporting requirement	Estimated costs per business
Germany's Network Enforcement Act (2017) *Requirement to report quarterly in German on their efforts to tackle illegal harm, including complaints and performance data*	£45,000 annual cost of reporting
Minimum implementation of the EU Non-Financial Reporting Directive for public interest entities with over 500 employees (2016) *Costs of reporting on anti-bribery and corruption matters.*	£951 first year costs with ongoing costs of £455
Mandating climate-related financial disclosures by publicly quoted companies, large private companies and Limited Liability Partnerships (2021) *Requires in-scope companies to report on metrics and targets used to assess and manage climate related risks and includes publishing as part of their annual report.*	£73,700 first year costs with ongoing costs of £56,800[167]
Climate Change Risk – Governance and Disclosure (TCFD) Requirements (2021) *Requirement on pension schemes in scope to publish a Task Force on Climate-related Financial Disclosures (TCFD) report.*	£3,750 first year costs with ongoing costs of £3,375
Corporate Governance Reform (2018) *Requirement in-scope companies to report on pay ratio.*	£5,688 annually[168]
Payment Reporting Requirement (2016) *Requirement to report on payment information, including late payments*	£1,270 first year costs and £1,012 each year after[169]

205. Costs presented in the previous IA (proxied from NetzDG) remain the most reasonable and analogous. The table below outlines the range of expected costs associated with transparency reporting:

[167] This includes both the metrics and targets aspect and signposting which are analogous to the kind of information required under the Option 1.

[168] This includes data collection, presentation and board discussion, and sign-off at the committee level.

[169] These include reporting costs minus familiarisation costs

Table 20: Transparency reporting (2019 prices, 2020 base year – 10-year PV)

	Low	Central	High
Option 1: Transparency reporting	£0.8 million	£6.7 million	£10.1 million

206. Ultimately Ofcom will consider a range of factors in determining what information providers' will need to produce in their transparency reports. This includes (but is not limited to) the kind of service provided, its functionalities, its user base. It is likely that the information requested will vary between different services. In every case, however, Ofcom must take account of the capacity of the provider of the service. Therefore, while relevant compliance costs will vary, disproportionate outcomes should not arise.

Fraudulent advertising duty

Requirements

207. Option 1 places an additional advertising duty on Category 1 and 2A platforms to implement systems and processes to minimise the risk that they publish and/or host fraudulent advertisements (paid-for advertisements that amount to a fraud offence). While the exact steps businesses can take will be set out in future codes of practice (subject to consultation and impact assessments), this duty will result in these platforms being required to implement more comprehensive fraud prevention measures. In line with the rest of Option 1, the small number of platforms in scope of this duty (c.20) are likely to ensure compliance in a variety of ways depending on the risk of fraudulent advertising on their platform and any anti-fraud measures currently in place. Potential processes that these platforms could take include some form of increased customer due diligence (CDD), such as know your client (KYC) checks, credit checks, and sharing information on known fraudulent advertisers. They will also need to ensure that users can easily report fraudulent adverts and take appropriate action on receiving these reports.

Baseline

208. The digital advertising market is largely controlled by two platforms, namely Facebook and Google together accounting for 80% of all spending on search and display advertising. Based on desk research of large social media sites and search services, current baseline coverage of anti-fraud measures and advertiser verification is mixed. Some platforms do not verify advertisers and instead focus on advertisement curation and ensuring that they are not in breach of the sites' terms of service. Other platforms have very light touch signup requirements, such as verifying an advertiser's email address or website and potentially payment details. Many platforms operate optional verification for businesses wanting to advertise, where businesses are encouraged to undergo some form of due diligence to appeal to customers. Where platforms currently mandate advertiser verification, this is largely focussed on advertising related to social issues, elections, and politics. Advertisers wanting to post content on these issues are required to provide valid identification and comply with several rules, including adding disclaimers to adverts and the sources of funding.

209. Facebook - the second largest player in the online advertising market - verifies political advertisers but has not announced plans to extend this to all advertisers. Facebook along with X and Microsoft recently announced that they would only host advertisements for financial products from companies that are authorised by the FCA.[170] This covers some of the types of measures that

[170] Tech giants agree to only publish ads of FCA-authorised firms (International advisor, 2021)

Ofcom will expect from in-scope platforms. These measures were likely introduced due to pressure from government and consumers and in anticipation of likely upcoming legislation. In 2019, Facebook also took several fraudulent advertisers to court for violating advertising policies and for defrauding individuals and tricking them into installing malware. In addition, Facebook - like the vast majority of social media sites and search services - allows users to report fraudulent adverts.[171]

210. In 2018, Google announced a new identity verification policy for political advertisers requiring them to provide government-issued identification and source of funds. In 2020, Google announced that it would extend this programme to all advertisers on its platform[172]. Advertisers will need to submit to Google personal identification, business incorporation documents or other information that proves who they are and the country in which they operate. Additionally, in line with plans from other large platforms, Google verifies all UK advertisers that wish to post financial services related adverts of any kind and requires that they are authorised by the FCA. This is important as Google alone represents 90% of the search advertising market and is by far the single largest platform in the online advertising space.

211. It is not possible at this primary stage to discount platforms mentioned above from incurring potential business costs. While many of the current and planned measures are in line with actions businesses are likely to take to comply with Option 1's advertising duty, it is not clear how effective they are and companies will likely be required to go further by, for example, tackling broader categories of fraud beyond financial. Ofcom will ultimately consult platforms, assess current baseline measures, and determine the steps businesses can take to comply.

212. The Advertising Standards Authority (ASA) - the UK's independent advertising regulator[173] - has partnered with major online platforms to address fraudulent advertising. The ASA has introduced the Scam Ad Alert system which allows users to report fraudulent adverts. Once reported, the ASA works with online platforms to take fraudulent adverts down and to stop similar adverts appearing. In the first six months of the Scam Ad Alert system, the ASA received 1,274 reports resulting in 121 alerts being sent to online platforms. Given the lack of robust data, it is difficult to determine long term trends and therefore it is not possible to fully evaluate the current self-regulatory system. However, fraudulent adverts are still widespread online and result in significant financial (and non-financial) loss to victims. Platforms are currently taking voluntary measures in this space but it is not clear how effective these are or whether further action is necessary. On this basis, the government has determined that a specific advertising duty on Category 1 and 2A platforms to ensure regulatory oversight of anti-fraud measures is necessary to mitigate wide scale economic losses.

Cost estimates

213. It is not clear at this primary stage what platforms will be required to do in response to the advertising duty. Option 1 sets out necessarily high-level duties on platforms and Ofcom will work with industry to assess the impact of measures it deems appropriate for compliance, including a full assessment of the impact on small and micro businesses (who themselves will not be in scope of the advertising duty but may be affected by it). At this stage, this impact assessment draws on a range of evidence sources to provide an indication of the likely scale of impact.

[171] But it is unclear how effectively they act on user reports.

[172] Google, 2020, Increasing transparency through advertiser identity verification

[173] The ASA is an example of self-regulation and co-regulation and is funded by industry.

214. There is likely to be a range of potential measures that platforms could introduce to comply with this duty. For example, it could include verifying advertisers, credit checks, sharing information on known bad advertisers or a range of other anti-fraud measures. Specific steps platforms can take will be set out in future codes of practice but it is plausible at this stage to assume that the advertising duty will result in a requirement on Category 1 and 2A platforms to conduct more stringent CDD on advertisers. Based on policy intention, approximately 20 platforms are expected to be designated as Category 1 or 2A and in scope of the advertising duty.

215. To calculate direct business costs, this impact assessment takes a top-down approach. Evidence from a representative survey of SMEs conducted by the Interactive Advertising Bureau (IAB) indicates that, on average, 60% of SMEs take part in paid-for advertising online through placement of advertising.[174] Broken down by business size, this is 52% of micro businesses, 81% of small businesses, and 96% of medium-sized businesses. The IAB's findings - while representative - only included registered micro businesses. It is reasonable to assume that the proportion of unregistered micro businesses - that is, businesses too small to be registered for VAT - is likely lower than those that are registered. However, in the absence of specific evidence on this section of the economy, estimates for registered micro businesses are applied to unregistered businesses - this represents a conservative approach. The IABs survey also did not include large businesses. However, the proportion of businesses increases with firm size and, therefore, it is estimated that 99% of large businesses participate in paid-for advertising online. The proportion of each size category is then applied to BEIS' BPE[175] and UK Civil Society Almanac data[176] to determine the total number of UK businesses (or the total number of businesses likely to undergo CDD because of participating in paid-for advertising online).[177]

216. The proportion of businesses which advertise within each size category is expected to increase across the appraisal period. However, IAB survey data used to estimate the percentage of advertising businesses is only available for a single year. To reflect this potential growth in advertising businesses, this impact assessment uses the average growth in the proportion of UK businesses with websites between 2007 and 2019 as a proxy (or +1.5% per year).[178] Digital advertising spend was considered but evidence suggested that large players would lead to overestimates. The proportion of businesses who advertise can't grow indefinitely, so within each firm size band, growth in the proportion of businesses advertising online was capped at 99% in the modelling. This reflects a potential saturation point at which point all potential advertisers are already placing advertisements - both medium sized and large businesses reach the saturation point within the time-period. The proportion of micro businesses advertising online grows from 52% to 59% across the period (an increase of 1.5 million businesses) and the proportion of small businesses grows from 81% to 93% (an increase of 0.1 million businesses).

217. By the first year of the appraisal period, it is estimated that approximately 3.4 million UK businesses will advertise online, this figure grows to 5.0 million by year ten. It should be noted that this approach is conservative, as some of these businesses may participate in paid-for advertising on platforms outside the scope of the advertising duty only. However, given the high levels of market concentration in this space, it is reasonable to assume that many advertise on or through a platform likely to be designated as Category 1 or 2A - Google and Facebook alone have 1.2 million

[174] Powering Up: Helping UK SMEs unlock the value of digital advertising (IAB, 2020)

[175] Business population estimates 2021 (BEIS, 2021)

[176] UK Civil Society Almanac (NCVO, 2021)

[177] In line with the rest of this IA, the number of businesses grows in line with annual business growth across the period.

[178] This impact assessment conducts sensitivity analysis on a range of growth rates from 0% (no growth) to 6.4% annual growth (the largest annual increase in the proportion of businesses with websites which occurred between 2007 and 2008).

UK advertisers on their platforms.[179] This approach does not account for advertisers based outside the UK that target adverts towards UK users. While any costs on those advertisers would not normally be considered in an IA, the cost on UK-based Category 1 and 2A platforms of conducting CDD would be in scope. There is no existing evidence or data on how many non-UK based businesses advertise to UK consumers using Category 1 and 2A platforms and, therefore, it has not been possible to monetise these potential costs at this stage. In future codes, Ofcom will consider the full range of impacts through comprehensive consultation with affected platforms, including the cost of anti-fraud measures as they apply to non-UK based advertisers which target UK consumers.

218. As IAB estimates are based on active advertisers (having advertised in the last 12 months), this impact assumes that 100% undergo CDD in the first year. From year two onwards, only new advertisers undergo CDD checks.[180] Across the appraisal period, there may be additional due diligence required on already authorised advertisers resulting from, for example, business changes or updates to identity documents. Given the uncertainty around specific requirements, it is not possible to reflect this possibility with any reasonable accuracy at this stage. In addition, the steps platforms will take will depend on the risk of fraud on their platform and the changing fraud landscape.

219. Of course, some businesses advertise on multiple channels and will be required to undergo CDD on more than one platform. Based on evidence from the IAB, on average, the number of channels used across SMEs overall is 1.2, 2.4, and 3.7 for micro, small and medium sized advertisers respectively. Large businesses are much more likely to advertise across a range of channels, for example by advertising on some combination of the large social media companies and Google. In the absence of specific evidence related to large businesses, this impact assessment assumes that these businesses advertise on average across 5 different in scope platforms. The number of advertising businesses within each size category is then uplifted by the average number of channels for the respective size category.

220. There are four main costs modelled using the above approach:

- **set up costs**: the cost of updating systems and processes to account for new requirements related to CDD
- **CDD costs (platforms)**: the cost of conducting a CDD on an advertiser
- **staff time (advertisers)**: the cost to advertisers of completing any forms associated with CDD requirements and providing appropriate information
- **staff time (advertising agencies)**: the cost to advertising agencies of facilitating CDD between platforms and advertisers

221. Set up costs are proxied from the impact assessment supporting the Transposition of the EU Fifth Anti-Money Laundering Directive which - in the context of the cryptoasset market - estimated set up costs for each firm of between £130,000 and £522,500 (central estimate = £326,000).[181] [182] The Money Laundering Regulations required a variety of different customer due diligence activities and are a reasonable but conservative proxy for unit costs in Options 1's advertising duty. Set up costs in the context of crypto providers was based on firms without current anti-money laundering

[179] Online platforms and digital advertising - Market study final report (CMA, 2020)

[180] New advertisers incorporate both the growth in the proportion of businesses that advertise online and the growth in the business population.

[181] All figures have been uplifted from 2017 prices to 2019 prices in the model.

[182] Transposition of the Fifth Anti-Money Laundering Directive (HMT, 2019)

frameworks in place. Many Category 1 and 2A platforms already have anti-fraud measures in place and, therefore, proxied set up costs are expected to be an overestimate. These costs are incurred in the first year only and cover updating systems and processes. Based on baseline evidence that some large platforms already conduct similar kinds of anti-fraud due diligence and advertiser verification, this figure is likely conservative.

222. The unit costs of conducting CDD are also proxied from HMT's Money Laundering Regulations. Standard CDD is estimated to cost between £3 and £16 (central estimate = £10) and enhanced CDD is estimated to cost between £5 and £32 (central estimate = £19). While the vast majority of CDD resulting from the advertising duty is expected to be automated (at least to some extent), the inclusion of estimates for enhanced CDD allows for the possibility that a small number of cases require additional scrutiny, such as for advertisers operating in industries known for high levels of fraud. This impact assessment conservatively estimates that 5% of advertisers will require enhanced CDD resulting in greater costs for in scope platforms. Enhanced CDD was expected to be conducted on only 0.23% of customers in the context of anti-money laundering. However, under the OSA platforms may decide to take a more risk averse approach with more stringent checks on risky industries or types of businesses (as opposed to individual customers as is the case for anti-money laundering).[183]

223. In addition to the cost of Category 1 and 2A platforms conducting CDD, advertisers themselves will also incur costs associated with completing necessary forms and providing appropriate information to in-scope platforms. This impact assessment estimates that this will take between 10 and 30 minutes (central estimate = 20 minutes) for standard CDD and between 30 and 60 minutes (central estimate = 45 minutes) for enhanced CDD. There is limited evidence on the time taken for an advertiser to complete a process like this, but it is in line with estimates for the time taken to open a bank account in the UK (itself subject to anti-money laundering checks).[184] This impact assessment assumes that this process will be conducted by a Chief Executive in small and micro businesses and by a marketing associate in medium and large businesses. Finally, a proportion of advertising businesses will use advertising agencies who may incur costs because of facilitating the CDD process. To account for this, 25% of CDD checks in the model include additional staff time of between 10 and 30 minutes (central estimate = 20 minutes) for advertising agencies. This is based on evidence presented to the CMA that a quarter of advertising revenue is channelled through media agencies. Given that most revenue comes from a small number of large advertisers, the actual proportion of advertisers using agencies is likely much lower and 25% is a conservative estimate.

224. Applying the methodology above, this impact assessment estimates that the fraudulent advertising duty will result in costs of **between £52.7 million and £187.0 million (central estimate = £120.0 million)** across the ten-year appraisal period.

Table 21: Fraudulent advertising duty (2019 prices, 2020 base year - 10-year PV)

	Low	Central	High
Option 1: Advertiser due diligence	£52.7 million	£120.0 million	£187.0 million

[183] Taking HMT's estimate of 0.23% instead of 10% reduces the cost of the fraudulent advertising duty by 9.2% with <0.1% change to total policy costs.

[184] How to open a bank account online (Which?, 2021)

225. Overall, we anticipate that compliance costs for Category 1 and 2A services will vary based on the nature, severity and potential harm to individuals presented by such content and the degree of control a provider has in relation to the placement of adverts. Ofcom will publish a code which sets out measures recommended for the purpose of compliance with this duty.

Indirect costs of fraudulent advertising duty

226. The extent to which Option 1's fraudulent advertising duty results in indirect impacts is dependent on several factors, all of which are at this stage unknown. While Ofcom will consider these further through consultation with industry and subsequent impacts assessments, the table below provides a qualitative assessment of the measures potential effect on supply, demand, and price in the market:

Table 22: Fraudulent advertising duty potential indirect impacts

Supply	The effect of Option 1 on the supply of advertising space is uncertain. Firstly, it will likely result in a short term increase in the supply of advertising space for non-fraudulent advertisers due to a reduction in fraudulent advertisers being able to advertise on platforms freeing up space for legitimate advertisers. However, the removal of fraudulent advertising would be beneficial to advertisers that don't want their products or services to appear next to harmful content or scam adverts, likely offsetting any potential increases in supply.
Demand	Demand side changes are complex and are expected to be influenced by the ability of advertisers to comply with additional checks and additional costs they may incur, because of completing forms and providing relevant information. If checks are too burdensome, or they can only be met by a subset of current advertisers, a fall in the demand for online advertising might be expected. However, online advertising has been performing strongly with rapid growth due to the ability to reach large audiences, the ability to engage users and drive direct sales, and the ability to target relevant audiences. These attributes mean it's unlikely further checks will result in any decrease in demand from established advertisers and agencies. Further, Category 1 and 2A platforms have a strong incentive to ensure their CDD processes are easy and user-friendly, given their reliance on advertising revenue. Finally, most of the largest companies in this space are already implementing anti-fraud measures and, therefore, it is reasonable to assume that they do not see a trade-off between checks and advertising demand. Demand for advertising could increase if advertisers - currently hesitant to advertise on social media due to harmful content and scam adverts - decide to purchase advertising space. Anti-fraud measures could also positively impact on consumer confidence which could lead to increased purchasing and increased demand for advertising space.
Price	Category 1 and 2A platforms may decide to pass costs on to advertisers who may ultimately pass costs on to consumers. There is no indication how, for example, Facebook and Google would adjust their pricing, whether a one-off joining fee, or a change in the fees charged for services for each advert purchased. This could be an increase in the cost per impression or cost per click, or a reduction in the revenue share a publisher receives. The ability of intermediaries to pass on costs to advertisers or publishers will depend on the level of market power they have. As mentioned, advertising platforms are highly concentrated in search and social display advertising. While advertisers can still go through other routes to reach audiences, they cannot access most internet users that access search and social media services.

	Given the scale of digital advertising spend and the relatively modest estimated cost of implementing anti-fraud measures, the extent of price increases is expected to be minimal and would be considered pass-through.[185]

227. The fraudulent advertising duty is proportionate and only applies to Category 1 and 2A platforms. While this is necessary to minimise business burdens, it does create a potential risk of fraudulent advertising being displaced to smaller less well-equipped platforms. Digital advertising is highly concentrated because platforms like Facebook and Google offer large and engaged user bases. While it is possible that some fraudulent advertisers may move to smaller platforms, given the advertising market share of large social media companies and search services, Option 1 is likely to capture a large proportion of advertising activity. If a fraudulent advertiser was to move to a smaller online platform (outside of Category 1 and 2A), it could not hope to attract the same number of advert impressions and, therefore, there would be less chance of users falling victim to the scam. Ofcom will further consider potential indirect impacts and risks associated with the fraudulent advertising duty. This will include consultation with affected businesses and subsequent impact assessments.

Break-out Box 12 - since the final stage IA: The OSB was amended to clarify platform's responsibilities with respect to dealing with illegal content and activity.

The illegal content duties require providers to proactively mitigate the risk that their services are used for illegal activity or to share illegal content ('preventative duties'). Services will also be required to address illegal content once it appears on their service ('content moderation duties'). The amendments make clear that platforms have duties to mitigate the risk of their service "facilitating" an offence, including where that offence may occur away from their site. This clarification addresses activities such as breadcrumbing, where child sexual abuse (CSA) offenders post links or have conversations on a particular site, preparatory to a CSA offence, that may occur on a different platform, or even offline.

This is not a widening of the scope and is consistent with earlier assessments of the cost of compliance that were based on a proxy regulation that didn't draw a distinction between on-site offences and facilitation of offences. The clarified policy is also consistent with the original Online Harms White Paper - "The primary purpose of the duty of care will be to improve safety for users of online services, and to prevent other people from being harmed as a direct consequence of content or activity on those services…in some cases the victims of harmful activity – victims of the sharing of non- consensual images, for example – may not themselves be users of the service where the harmful activity took place."[186]

Break-out Box 13 - since the final stage IA: Foreign interference, human trafficking, illegal immigration, coercive or controlling behaviour and unnecessary suffering of animals were added to the list of priority offences (Schedule 7 of the OSA). For clarity, these offences were all illegal content in the original OSB, but they have been "upgraded" to priority illegal content in the OSA.

The OSA lists providers' duties with respect to illegal content under s9 and s10. These include duties to –

- undertake and maintain an illegal content risk assessment,
- take proportionate design measures to minimise contact with priority illegal content and mitigate harm identified in the illegal content risk assessment,
- minimise the time priority illegal content is present, as well as take down illegal content,
- include and apply provisions on illegal content in their terms of service, and

[185] RPC case histories - direct and indirect impacts, March 2019 (RPC, 2019)

[186] Online Harms White Paper, 2020. https://www.gov.uk/government/consultations/online-harms-white-paper/online-harms-white-paper

- publish transparency reporting (for Category 1, 2A and 2B platforms).

The three areas where there are specific provisions for priority illegal content concern –

- illegal content risk assessments, where platforms are required to consider the risk of users encountering priority illegal content (as well as other illegal content),
- proportionate design measures to reduce the risk of encountering priority illegal content as well as mitigate the risk of other encountering other illegal content' and
- content moderation to minimise the time priority illegal content is present, as well as take down illegal content.

When these duties were appraised in the final stage IA, the appraisal for risk assessments was based on an offence-agnostic proxy, the Networks and Information Systems Regulations 2018, so it would not improve the robustness of the appraisal to adjust the estimate.

In the case of content moderation, Revealing Reality surveyed platforms about the expected cost of additional content moderation, using the framing in the Online Harms White Paper (the specific framing was "to address all categories of harm in OHWP" which covers these offences). Given this, we are confident that the content moderation estimates the Revealing Reality research reported included these offences. Though Ofcom has not published their final codes of practice, their consultation guidelines do not require proactive content moderation for any priority offences other than CSEA and terrorism material. Given this, it is likely that there will no distinction between platform's duties for content moderation related to these offences because of this amendment.

User verification and empowerment duties

228. Option 1 introduces a duty on Category 1 platforms to offer optional user verification and user empowerment content tools to their adult users. In terms of optional user verification, Category 1 services would be required to put in place a mechanism by which an adult user could verify their identity, should they wish to do so. Ofcom will set out recommended verification methods in guidance, however Category 1 services will have discretion on which verification method they offer, which may be of any kind and although it need not require documentation to be provided, they are able to request this should they choose The duty to provide optional user verification is separate from age assurance requirements under the child safety duties and pornography provision.

229. In addition, Option 1 places a duty on Category 1 platforms to assess the incidence of certain kinds of legal content on their services (discussed above in the break-out box on risk assessments), and then (where proportionate) provide user empowerment tools for these kinds of content so that users can have more control over their online experience. Category 1 platforms have discretion over the legal content they permit on their services. However, if a provider allows the relevant categories of content on its service and users are likely to encounter them, a provider would have to offer adult users easy to access tools to enable them to reduce their exposure to the relevant categories of content, or be alerted to the nature of it. The content that these tools will apply to are set out on the face of the OSA in Section 16.

Break-out Box 14 - since the final stage IA: The then government strengthened these duties so that providers will be required to ask their registered users, at the first possible opportunity and until they indicate their preferences, how they would like these tools to be applied. The expectation is that providers will then save this preference and therefore not have to ask the user again, however, it's possible that there could be some friction for users accessing the service until they indicate their preferences. This impact is discussed in a new section of the IA on user experiences.

Services will also be required to offer users the opportunity to filter out content from non-verified users, and prevent non-verified users from interacting with their content. The user empowerment steps that services can take will be set by Ofcom in codes of practice. users will use these tools to have greater control over the content they see online and determine who they interact with online.

Baseline

230. Many of the largest social media platforms offer user verification already. This varies by platform and includes verified users either having to have an active subscription to the platform's services, meeting eligibility requirements such as having a display name and profile photo and a phone number, paid monthly subscriptions or verified badges.

231. In terms of user empowerment tools, a review of the large social media platforms found that existing tools can broadly be broken down into the following categories:

Table 23: Examples of current tools available on Category 1 platforms

Form of user empowerment tool	**Description**
Chat functionality controls	Platform tools that give users the ability to block or restrict who can chat or direct message them.
Content controls from/to specific accounts/individuals	Controls that allow a user to either block a user, preventing them from seeing and interacting with a profile; mute a user, preventing content from that user from being seen; or in some cases, choosing what types of content other specified users can see from a given account.
Inappropriate content controls	This can vary significantly from platform to platform but covers: - Options to hide comments on posts, stories, and live videos deemed by platforms to be inappropriate or offensive. - Options to filter out profanity. - Options to filter out content containing words/phrases the user does not wish to see.
In-app purchase controls	Controls enabling users to make decisions around if and how much spending can occur on a given platform by the user/user's child.
Privacy controls	Tools to provide users with control over how they can be found, the level of visible personal information, and information around location.

232. All major social media platforms reviewed as part of this IA have some level of user empowerment tools in place already, though this tends to vary quite considerably from platform to platform, and by platform type and functionality. In terms of current coverage, the most commonplace tools available to users are tools to filter out inappropriate content, tools to enable user blocking/muting, and privacy controls. These were available across almost all large social media platforms assessed.

Cost estimates

233. In-scope platforms are likely to take a variety of approaches to optional user verification. Ofcom will set out appropriate measures that Category 1 platforms can take to comply with this part of the duty. At the primary stage, this impact assessment takes a user-based approach to model potential impacts and to provide an indication of the likely scale of costs.

234. The proportion of each adult age group that uses social media sites is taken from Ofcom's most recent adult and child media use and attitudes surveys.[187] [188] The percentage of individuals within each age group ranges from 45% (for 65 year olds and older) up to 90% (for 35-44 year olds). Across the ten year appraisal period, the proportion of each age group that uses social media is assumed to increase in line with average social media use growth rates within each age group between 2015 and 2021.[189] Within each group, growth in the proportion of adults using social media stops when it reaches 95%. This reflects a potential saturation point at which point all potential social media users within each age group are using social media. Given the relatively low growth rates in age groups with over 90% already using social media, a 95% saturation point is realistic. Only one age group – namely the over 65s – does not reach the saturation point in the ten year appraisal period, increasing from 45% to 78%.[190] Social media use estimates are then applied to ONS population data[191] to obtain the total number of people in the UK that use social media in each year of the appraisal period. In line with the rest of this IA, population is estimated to grow in line with average growth rates between 2000 and 2020 (0.65% per year).

235. As user verification is optional under Option 1, it is not clear how many social media users will decide to be verified. To estimate this, the model takes account of relevant polling data related to anonymity online and identity verification on social media sites. Based on five recent polls, it is estimated that between 50% and 78% (central estimate = 68% poll average) have either a negative view towards anonymity on social media or a positive view towards identity verification on social media. For example, in one poll, half of respondents opposed being able to create an anonymous account[192] and in another, 78% thought that users should have to verify when signing up and/or display their real name at all times.[193] It is possible that, when faced with a specific polling question, individuals are more likely to say they support verification or oppose banning when in reality they would not opt for optional verification themselves. However, this represents a very conservative upper bound estimate of the proportion of UK users who may make use of optional verification measures. The proportion of users willing to be verified will also depend on the type of verification required and the information they are willing to give to the platform. Polling data is then applied to estimates for the number of social media users within each group. Finally, to determine the total

[187] Adults' media use and attitudes report 2020/21 (Ofcom)

[188] Children and parents: media use and attitudes report 2020/21 (Ofcom)

[189] Ofcom adults media use and attitudes report 2015 - 2021 (Ofcom)

[190] Varying the saturation point has minimal effects on total costs. For example, a saturation point of 99% results in a 3% increase in total verification costs compared to 95% saturation point.

[191] ONS population data - source

[192] Polling, Left Foot Forward, 2021

[193] Daily Question 13/07/21, YouGov 2021

number of potential verifications, estimates for social media users likely to opt for verification are uplifted by the average number of social media sites used (6.4).[194] The total number of potential verifications is estimated to be between 173 million and 270 million (central estimate = 235 million).

236. Verifying willing social media users is likely to be spread across several years as users become aware of the option and decide to be verified. This impact assessment assumes that willing users are verified equally across the first five years. From year 6 onwards, only new willing users are verified each year.[195] [196]The unit cost of verifying a user is highly dependent on the method chosen by the platform. While specific methods will be set out by Ofcom and determined by the individual platform, this impact assessment uses the lowest cost provided by third-party age verification providers in the context of verifying a user's age (£0.07p per verification). While verifying identity and verifying age are related, they do represent separate processes. This cost is considered a reasonable proxy to indicate the likely scale of impact. Age verification checks generally rely on official ID and therefore, this is likely an overestimate, especially for platforms that opt to verify email addresses only. Applying the methodology described above, total costs of offering optional user verification to adults are estimated to be between **£9.0 million and £14.1 million (central estimate = £12.3 million)** across the appraisal period. As platforms already verify a proportion of users under the *status quo* and social media sites in general collect a large amount of user data already, Category 1 platforms are not expected to incur significant costs associated with changing systems or extending user databases. However, estimates will be further refined by Ofcom as the regulator determines specific requirements on platforms.

Break-out Box 15 –since the final stage IA: The OSA was amended to specify a list of content categories that platforms would be required to provide user empowerment tools for.

In the initial version of the OSB section 14 set out the user empowerment duties. This applied to Category 1 services and mentions "where proportionate, features which adult users may use or apply to increase control over harmful content." This was to "reduce the likelihood of the user encountering priority content that is harmful to adults" and "alert the user to the harmful nature of priority content that is harmful to adults". Priority content to adults was not set out in this version of the OSB as this was intended to be set out in secondary legislation.

This amendment does not represent a widening of scope. Under the previous user empowerment duties, Category 1 services would need to provide user empowerment tools for categories of content that would have previously been set out in secondary legislation ('priority harms to adults'). This list of content categories has been replaced by a user empowerment content list which is now set out in primary legislation and covers the same types of content. In both cases, we are unable to estimate the cost of content-specific user empowerment tools, due to a lack of applicable data, but considering the intentions of the policy, this amendment has no substantial impact.

Break-out Box 16 - since the final stage IA: As noted above, duties related to "legal but harmful" were amended to require platforms to supply user empowerment tools to offer adults the choice to manage certain categories of legal content that they were exposed to. In terms of the categories of content included in the White Paper (the basis of cost estimates in the final-stage IA) compared to the OSA, 'misinformation and disinformation' and 'the intimidation of public

[194] Backlinko, 2024, Social Media Usage & Growth

[195] New users reflect both population growth and any growth in social media use.

[196] In an extreme scenario where 100% of willing users are verified in the first year, total verification costs increase by 14% with less than 0.1% effect on total policy costs. It is far more likely that verifications occur over a number of years.

figures' have been removed. Misinformation and disinformation will still be covered where it is illegal or harmful to children or covered by terms of service of Category 1.

The user-empowerment duties offer the same tools to adult users to manage content as the original duties, without proscribing particular content or telling platforms what legal content they can or cannot allow. Informal engagement with platforms on these changes indicates that services would see user empowerment tools as a significant cost to implement and maintain, while disallowing harmful content being seen as the cheaper option. Based on an exploratory review of large platforms' terms of service, many platforms do not allow this content on their platforms in the baseline. As a result, we expect that companies will restrict or remove this content being posted on their platforms, rather than providing specific user empowerment tools.

While services are expected to behave this way, and we conservatively assume the costs will be equivalent, the changes create more uncertainty about benefits as there is a range of potential platform responses. Benefits have been reduced as 'mis/disinformation' and 'intimidation of public figures', which were included in the July 2022 indicative list of priority harmful content for adults[197], have been removed from the categories covered by the user empowerment duties.

We have considered how we could refine these assumptions; however, it is difficult to update this costing as the benefits associated with the costs are not connected. This is because the final-stage IA estimated the costs of compliance based on a survey conducted by Revealing Reality based on the White Paper's initial list of harms in scope , and the percentage of revenue it would cost to moderate those harms in aggregate. This survey provided a lack of detail about what split of revenue would be spent on additional content moderation for each individual harm. This makes it difficult to judge the cost reduction associated with the removal of specific types of content.

We are unable to use the contents of terms of service of large platforms to assess what was already excluded, however this approach would pre-empt Category 1 classifications from Ofcom and was difficult to say with certainty that it provided a better understanding of costs. We also considered assuming lower shares of turnover from the Revealing Reality study (currently 1-15%), however these revisions would be arbitrary, without evidence to judge their reliability. Therefore, keeping the original costing assumptions represents the best, if conservative, judgment of cost available.

Table 24: Optional user verification (2019 prices, 2020 base year – 10-year PV)

	Low	Central	High
Option 1: Optional user verification	£9.0 million	£12.3 million	£14.1 million

237. For the user identity verification duty, there may be instances where Category 1 providers already have identity verification schemes in place which meet the requirements set out in the Act. In such instances, we expect that set up and compliance costs to providers will be minimal. However, providers will need to assess whether current identity verification schemes offered are compliant with the relevant duty in the Act. Ofcom will produce guidance to support providers of categorised services in complying with these duties.

[197] Written Ministerial Statement: https://questions-statements.parliament.uk/written-statements/detail/2022-07-07/hcws194#:~:text=Adults%3A,not%20be%20covered%20by%20this.

238. It has not been possible at this stage to monetise the potential impact associated with user empowerment tools, such as giving users the ability to have greater control over certain categories of content. Firstly, thresholds for Category 1 platforms will be set out in secondary legislation. With such a high degree of variability in current coverage amongst major online platforms (potential Category 1), it is not possible to estimate with any reasonable accuracy what platforms are likely to do to comply. Further, any potential incremental costs are likely to relate to platform design. These types of changes are very difficult to monetise and evidence of historical costs is limited (or non-existent) given the sensitive nature of costs. In addition, many of the largest services already prohibit the kinds of content covered by these requirements, and therefore assuming these terms of service are applied consistently (as required by the transparency, accountability, and FoE duties), these services would not face significant additional burdens. Ofcom will work with platforms to determine specific requirements and assess the impacts of any potential platform changes.

239. In addition, in determining what is proportionate when offering user empowerment content tools, the Act sets out that the findings of the most recent user empowerment assessment including the incidence of user empowerment content on their service, and the size and capacity of the service, are relevant. Ofcom will set out in a code of practice how companies can comply with the user empowerment tools duty and in guidance how companies can comply with their user empowerment assessment duty. Therefore, in instances where smaller services have this duty, or a service assesses that they have (and can evidence) a low incidence of user empowerment content, we expect initial set up costs and ongoing compliance costs to be minimal.

Assessing platform impacts on freedom of expression and privacy

Requirements

240. The largest and greatest influence platforms (Category 1 services) will have additional duties to assess the impact of their safety policies and procedures on freedom of expression (FoE) and privacy and demonstrate they have taken steps to mitigate this. Service providers will be required to include a section in that impact assessment which considers the impact on availability and treatment of news publisher and journalistic content on their service. They will need to publish this impact assessment and keep it updated (referred to in this section as a FoE and privacy IA). This builds on existing duties to have regard to these factors when adopting measures to comply with the safety duties.

Baseline

241. The government is not aware of any platforms currently in compliance with this requirement and considers the full cost of producing an assessment to be incremental. Whilst organisations interviewed in RR's 2023 research felt that their community guidelines already considered freedom of expression, none mentioned carrying out specific impact assessments on freedom of expression.

Cost estimates

242. Realised costs are dependent on requirements set out in future codes of practice; however, to provide an indication of the likely scale of impact, estimates from impact assessment requirements under General Data Protection Regulations (GDPR) are used as a proxy. Under GDPR, businesses are expected to produce Data Protection Impact Assessments (DPIAs) for any processing that is likely to result in a high risk to individuals. Businesses are also encouraged as good practice to produce a DPIA for any other major project which requires the processing of personal data. A DPIA must: describe the nature, scope, context and purposes of the processing; assess necessity, proportionality and compliance measures; identify and assess risks to individuals;

and identify any additional measures to mitigate those risks. The cost of producing a DPIA is considered to be a reasonable proxy for the cost of producing an FoE and privacy IA under the OSA. Costs will be considered further through Ofcom's consultations with industry and future impact assessments.

243. The unit cost of producing a DPIA comes from the Ministry of Justice's Proposal for an EU data protection regulation - Impact Assessment.[198] [199] In 2019 prices, this equates to £12,700 for a small-scale DPIA, £31,300 for a medium-scale DPIA, and £135,000 for a large-scale DPIA. The estimate for large-scale DPIAs was considered in the Ministry of Justice's IA as an extreme example of a large project involving sensitive data and was not used in its calculations. Given the uncertainties on requirements related to FoE and privacy IAs, calculations presented here use the full range of potential costs to form the low, central, and high estimate. Given that Category 1 platforms will be required to ensure this assessment is updated, this IA assumes that this cost is incurred each year but reduces by 50% from year 2 onwards. The table below outlines the range of expected costs associated with FoE and privacy IAs:

Break-out Box 17 - since the final stage IA: As stated above, service providers will be required to include a section in their FoE and privacy impact assessment which considers the impact on availability and treatment of news publisher and journalistic content on their service.

As discussed above, the cost of producing FoE and privacy impact assessments has been proxied using the cost of producing Data Protection Impact Assessments (DPIA) required under GDPR. Given the policy uncertainties, a wide range of estimates is provided to reflect the potential scale of the requirements.

Even with this specific requirement, there remains significant uncertainties on requirements related to producing FoE and privacy impact assessments, and this remains a reasonable proxy at this stage.

Table 25: FoE and privacy IA (2019 prices, 2020 base year – 10-year PV)

	Low	Central	High
Option 1: FoE and privacy IA	£1.0 million	£2.5 million	£11.0 million

Reporting online CSEA to the NCA

244. Option 1 introduces a legal requirement on technology businesses to report online CSEA. This requirement will apply differently to platforms depending on where they are based, which is different from the approach being taken to the Online Safety regime generally, where duties will apply to all in-scope services that have UK users. UK platforms[200] will be required to report detected CSEA content to the NCA. Platforms providing services from outside of the UK will only have to report identified CSEA offences that are UK-linked, and only if they do not already report CSEA to a body outside of the UK. All services will be able to decide whether to report to the NCA as the UK designated body or an equivalent foreign agency. This will ensure that platforms do not have to

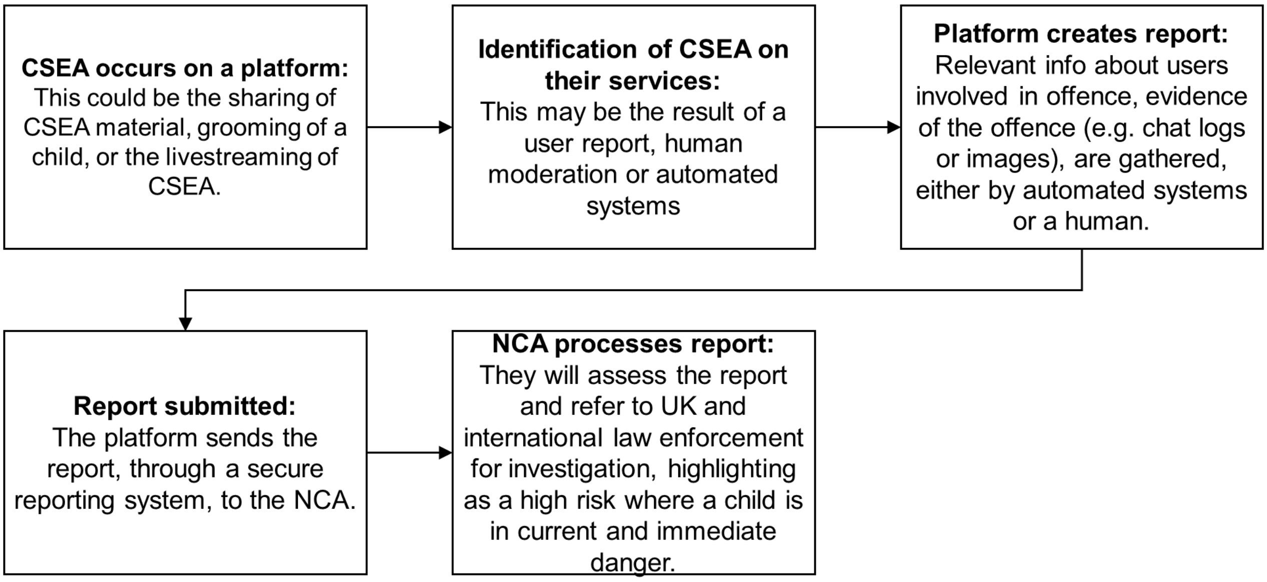
[198] Proposal for an EU data protection regulation - Impact Assessment (MoJ, 2012)

[199] To note, Ministry of Justice estimates are themselves taken from the EU commission's own estimates.

[200] This includes regulated user to user or search services which are either individual(s) who are habitually resident in the UK or an entity incorporated or formed under the law of the UK.

replicate their reporting efforts. For UK platforms, this will replace the current voluntary reporting regime within the UK. The below figure outlines the new process:

Figure 4: Mandatory reporting process for CSEA


CSEA occurs on a platform: This could be the sharing of CSEA material, grooming of a child, or the livestreaming of CSEA.
Identification of CSEA on their services: This may be the result of a user report, human moderation or automated systems
Platform creates report: Relevant info about users involved in offence, evidence of the offence (e.g. chat logs or images), are gathered, either by automated systems or a human.
Report submitted: The platform sends the report, through a secure reporting system, to the NCA.
NCA processes report: They will assess the report and refer to UK and international law enforcement for investigation, highlighting as a high risk where a child is in current and immediate danger.

245. Some countries, including the USA and Canada, have legal requirements on platforms to report online CSEA. Platforms based in the USA are required by law to report to the US National Centre for Missing and Exploited Children (NCMEC). In 2022, NCMEC received 32.1 million reports, of which 31.8 million were from platforms,[201] and 317,000 reports related to victims or offenders in the UK and were triaged and sent to the UK's NCA.

246. NCMEC data on reports by platforms in 2022 demonstrates that smaller technology platforms report less than the big technology platforms. The cost to smaller organisations to report is therefore likely to be low on average, as they are likely to identify fewer instances of CSEA to report to the NCA.

247. Most large technology platforms, where the majority of NCMEC CSEA reports come from, are based in the USA and already report to NCMEC. To avoid duplicate reports being made, these platforms will not be required to report again under the UK reporting regime. For example, Facebook UK will not report directly to the NCA as Facebook's headquarters are in the USA and the company already reports to NCMEC. In 2022, Facebook accounted for 67% of all platforms reports to NCMEC, with 21.2 million reports.[202]

248. Many platforms in scope of the online safety regime will have parent or subsidiary businesses based outside of the UK that already report CSEA. It is therefore highly uncertain how many additional platforms will report into the NCA as the UK's designated body because of the OSA, and how many reports of CSEA content the NCA will receive. Further, until a platform completes a risk assessment, as required under the OSA, and begins to proactively tackle CSEA (if it does not already), the volume of CSEA content that will be detected per platform is difficult to estimate.

[201] CyberTipline 2022 Report

[202] 2022 CyperTipline Reports by Electronic Service Provider

Break-out Box 18 - since the final stage IA: This IA updates the original illustrative estimates on reporting CSEA material to the NCA following more detail being added in the Act. This includes illustrative scenarios on the number of UK platforms estimated to report into the NCA and number of reports by each platform based on NCMEC data. The analysis is likely to represent a lower bound on number of business and reports expected for two main reasons: i) it does not account for reports from platforms providing services from outside of the UK (who do not currently report CSEA) who may report UK-linked content to the NCA (anecdotally there is a possibility of receiving as many additional reports from non-UK companies as from UK companies); ii) whilst the OSA applies to platforms offering user-to-user services and search services, this analysis focuses on user-to-user reports due to a lack of available data on potential reports expected from search services. Further work is needed to scope the operational details of this requirement and better understand the scale of CSEA content that may be reported by these services.

In 2022, over 1,500 platforms were registered with NCMEC to make reports of CSEA content. Of these 236 platforms sent reports to NCMEC in 2022, 17% of which were platforms not based in the USA voluntarily reporting to NCMEC. The top 5 platforms, namely Facebook, Instagram, WhatsApp, Google and Omegle, accounting for 94% of all reports (30 million) from platforms. Many platforms submitted much fewer reports of CSEA content, with the median number of reports by a platform at 35.

To understand how many platforms comparatively could report in the UK, two approaches are taken. Firstly, to assume the same proportion of businesses would report in the UK as in the USA. With 0.002% of businesses reporting to NCMEC in the USA, this equates to an estimated 35 additional businesses expected to report CSEA content to the NCA.[203] Secondly, to assume the ratio of the UK to the USA on a range of related metrics, including the number of internet users, number of businesses and population, is equivalent to the ratio of businesses reporting in each country. This suggests the UK could expect around a fifth of the businesses to report compared to the US, resulting in 39 businesses. These approaches result in similar estimates, with an average of 37 additional platforms estimated to report in the NCA due to the OSA.

As previously noted, the median number of reports each platform sent to NCMEC in 2022 is 35. Whilst platforms in the USA are subject to a legal requirement to report CSEA once they are made aware of it, NCMEC note that there are no legal requirements regarding proactive efforts to detect CSEA content, so the number of reports per platform can vary depending on a platform’s effort to detect CSEA content.[204] The OSA designates CSEA as a priority offence, such that platforms will have a duty to proactively identify, remove and prevent the content from being on their platforms as well as allowing user reporting. As a result, platforms may identify more CSEA content on their sites, and report more on average compared to platforms in the US. To account for this, the median reports per platform is used as a low scenario, with the central scenario assuming the number of reports per platform doubles (70) and the high scenario assumes the number of reports per platform triples (105).[205]

Previous data indicates that in specific instances, particular businesses may send in a high number of reports covering multiple years at one time, which can cause resourcing pressures for the reporting agency. This analysis is unable to account for this impact, as it is unknown whether

[203] 0.002% of businesses are estimated to report in the USA based on data from the United States Census Bureau.

[204] CyberTipline 2022 Report

[205] As this is based on 2022 NCMEC data, to get the number of reports expected per platform at implementation in 2025, these values are uplifted by the growth rate of reports per business, as outlined in paragraph 10.

new businesses reporting to the NCA will undertake reporting in this way, or how many reports they may be likely to generate. It is also likely that should businesses report in this way the reporting agency would proactively work with them to refine their reporting practices meaning this impact may only be short-term, albeit with additional resources required to rectify.

Over the appraisal period, the number of platforms reporting to the NCA is estimated to grow in line with the number of platforms in-scope of the OSA, at a rate of 2%.[206] The number of reports per platform is estimated to grow at 11% per year, following the growth in median reports per platform to NCMEC from 2019 to 2022.

On average over the appraisal period, 42 platforms per year are estimated to report to the NCA. Each platform is estimated to make 151 reports to the NCA on average per year, ranging from 75 - 226. As a result, platforms are estimated to make 6,300 reports per year to the NCA, ranging from 3,200 to 9,400.

The overall cost of reporting CSEA content for platforms is expected to be minimal, and to some extent is controllable by the organisation (for example, whether they use automated reporting). Some platforms reporting to NCMEC have indicated that they use a combination of manual and automatic processes to report CSEA. The cost for identifying CSEA is already part of the platforms' costs within their identification and moderation process. These processes will vary, with some proactively identifying CSEA using automation while others relying on user reports and human moderation. The undertaking additional content moderation section above estimates potential costs for businesses for additional content moderation due to the OSA.

To report CSEA content, platforms will need to set up an account on an NCA portal. As the design of the portal is currently ongoing, this impact assessment is not able to provide an estimate of potential costs associated with setting up an account. However, based on consultation with the NCA, the costs for businesses of setting up an account is expected to be minimal.

The cost of reporting CSEA content is the time it takes to send a report of the identified content or activity to the NCA, which translates to the cost of an employee's time, unless the process is automated. Based on engagement with stakeholders, the impact of sending a manual report is estimated to be 5-10 minutes of an employee's time per report. When applied to the hourly cost of a regulatory professional (including a non-wage cost uplift), this provides an estimated cost per report of £2.15 - £4.30. Applying this to the estimated number of reports per year results in a total cost to business across the appraisal period of £0.10 million, ranging from **£0.03 - £0.20 million**. Per platform reporting to the NCA, costs of reporting are expected to be minimal. However, these estimates are likely to reflect a lower bound, as they do not account for any costs relating to setting up an automated reporting system, given a lack of data on these costs, and uncertainty on how many businesses may undertake this following the OSA. Further costs may accrue to businesses if they undertake other processes as part of submitting a report, such as an investigation of a user's account or supervisory checks on a report. Due to a lack of data on the proportion of businesses who may undertake these processes and the costs associated with them, these are unquantified.

[206] This is the growth rate of UK businesses between 2000-2022. Business Population Estimates 2022.

Table 26: Reporting CSEA content to the NCA (2019 prices, 2020 base year – 10-year PV)

	Low	Central	High
Reporting CSEA content to NCA	£0.03 million	£0.10 million	£0.20 million

Once businesses submit a report of CSEA content to the NCA, they will be required to store the data from reports and associated account data of any user in the reports for 90 days. Engagement with stakeholders has indicated that this is unlikely to result in additional costs to businesses, noting that businesses would have also incurred storage costs had the content not been identified as containing CSEA.

Following a report of CSEA content, businesses may have further engagement with the NCA, for example on how to improve the information included in a report to ensure that the NCA can determine where the offence occurred and/or the appropriate law enforcement agency to receive the report (i.e., whether it should be referred to UK or international law enforcement agencies). Costs associated with this engagement are unquantified, due to a lack of information on the proportion of businesses that would require further engagement with the NCA, and the extent of that engagement.

Additional duties following the passage of the OSA

Break-out Box 19 - since the final stage IA: The OSA was amended to grant Ofcom powers to require platforms to source solutions and/or innovate to tackle online CSEA and terrorism content.

Under the Act, where necessary and proportionate, Ofcom can issue a notice requiring individual service providers to use accredited technology to identify, take down, or prevent users from encountering terrorism content and/or CSEA content.

“Whereas a notice relating to terrorism content can only apply to content communicated publicly, a notice relating to CSEA content could also apply to content communicated privately by means of the service. A notice could also require service providers to use accredited technologies more effectively, or (for CSEA content) require a service to use best endeavours to develop or source technology that works with their platform. This can be found in more detail in the Act section 121.

The costs to business will depend on the number of businesses issued a notice, and the potential cost of any measures required to identify and remove terrorism content and/or CSEA content following being issued a notice. Following Royal Assent of the OSA, Ofcom expects to provide advice to the Secretary of State regarding minimum standards of accuracy in the detection of terrorism content and CSEA content. The Secretary of State will then approve and publish minimum standards of accuracy, and only then will Ofcom (or a person appointed by Ofcom as relevant) be able to consider whether a particular technology can be accredited as meeting those minimum standards, and (if so) whether to issue a notice to a particular service provider. For this reason, this analysis is unable to estimate the costs to business in aggregate from this power.”

Break-out Box 20 - since the final stage IA: The OSA provides the Secretary of State with a delegated power to bring application (app) stores into the scope of regulation, following consideration of Ofcom’s report about use of app stores by children.

The Secretary of State will have the power to make regulations putting duties on app stores to reduce the risks of harm presented to children from harmful content on or via app stores. The regulations may make provisions to exempt certain types of app stores or specify threshold conditions to narrow which app stores fall in scope of the duties. Ofcom's report will assess the use and effectiveness of age assurance on app stores and consider if the greater use of age assurance or other measures could protect children further.

Baseline
The largest app stores already have app review processes and implement safety features, such as age-ratings for apps provided and age assurance.

Cost estimates
At this stage, it is not possible to monetise the impacts of bringing app stores into scope of the regulations. This is because:

- the conditions to determine which app stores fall in scope and the exact requirements they will have to meet will be set out in secondary legislation. This will be informed by Ofcom's report.
- the power is expected to be deferred until Ofcom publishes its report within two to three years of the child safety duties coming into force. This means that, if this power were exercised, these regulations would not be effective until at least 2027/28.

An accurate assessment of the direct costs to business will be provided at secondary stage, if the power is used. However, to illustrate the potential scale of these costs some estimates have been made below, assuming that app stores could be required to age-assure users in the UK with existing accounts, or at account creation, along similar lines to elsewhere in this impact assessment, as well as review apps on their stores to ensure that they are age-appropriate.

Age assurance
Requiring age assurance to create an account on apps stores may require those stores to have an age verification or age estimation system or solutions which are as effective. All users of app stores may need to have their age assured at account creation, and in some cases those with existing accounts may need to have their age assured retrospectively, depending on the approach taken should app stores be regulated. App stores already register the age of account holders (without technical verification), and limit access to apps. Apps are already self-assessed by developers for age appropriateness.

For this impact assessment, we estimate that in the first year of assurance (2027) approximately 118 million existing application store accounts may need to be age assured (low estimate 53 million, high 246 million). This will include accounts that are owned by adults, as well as children, though apps stores may not be certain of the age of the account holder until the account age assured. The total accounts assured is also based on the estimated number of mobile phones owned by groups at each age, as well as other (non-PC/laptop/netbook) devices that are used to go online, assuming a similar rate of child ownership to mobile phones.[207] This estimate assumes each user has two accounts in the central estimate (1 low, 4 high). Depending on the implementation option pursued, the number of accounts that need to be assured may be materially lower than the central estimate.

Following this first year of retrospective age assurance on existing accounts, app stores are expected to assure a much lower number of accounts, with age assurance only required on new

[207] Ofcom, 2023, Children and parents: media use and attitudes report 2023 – interactive data, https://www.ofcom.org.uk/research-and-data/media-literacy-research/childrens/children-and-parents-media-use-and-attitudes-report-2023/children-and-parents-media-use-and-attitudes-report-2023-interactive-data and Adults' media use and attitudes 2023: interactive report, https://www.ofcom.org.uk/research-and-data/media-literacy-research/adults/adults-media-use-and-attitudes/adults-media-use-and-attitudes-2023-interactive-report, as well as ONS population forecasts, https://www.nomisweb.co.uk/query/select/getdatasetbytheme.asp?opt=3&theme=&subgrp=

accounts. These new accounts are predominantly younger people growing up and owning a mobile phone/other device for the first time. The number of additional verifications is around 1.7m-1.8m per annum over the 10-year IA appraisal period, i.e. to 2033 (we estimate 0.9m in the low scenario, 3.5m in the high scenario).

Each age assurance is expected to cost £0.10.[208] This leads to a central estimated cost of compliance of £9.6 million (low estimate £4.4 million, high estimate £19.8 million) over the ten years of the appraisal period (2019 prices, 2020 present value).

App stores may also need to integrate with a provider of age-assurance services. Using the same approach as above, the cost would be less than 0.01 million.

No assessment has been made of the quantity of sales that would be lost from apps stores due to these restrictions, as data is not available on the value of apps that are age-inappropriate and are purchased by children. However, total app-related revenue in the UK was estimated to be £11.4 billion in 2022 and projected to be worth £18.8 billion in 2027.[209] The bulk of this comprises in-app advertising, followed by in-app purchases.

Reviewing apps for age-appropriateness
Apps stores may be required to review apps they sell to ensure that they are age appropriate for their specified age range. This already occurs in the case of one major app store and our estimates are based on available public data about that store as well as other market information.

We estimated the number of apps that need to be reviewed in any year by considering the number of apps available on the nine largest apps stores[210] and assuming that these represented a high proportion (97.5%[211]) of all available apps. The estimated number of apps available in 2024 was 6.7 million. We extrapolated out the number of apps to be reviewed in any particular year by estimating the compound annual growth rate considering public estimates of apps added per day, back to 2020.[212] We assumed that the backlog of apps to be reviewed (those already on the stores) would be reviewed over the first five years the policy was in effect. The number of apps reviewable in any particular year, starting from 2027, was 20% of the initial backlog of apps, as well as the new apps in each year. Three scenarios were considered –

- a low scenario where only high-risk apps were reviewed, this was proxied by apps rated 13+ on the two largest app stores, 10%[213]
- a high scenario where every app was reviewed, 100%
- a central scenario mid-way between the low and high

In the central scenario, 7.4 million apps will need to be reviewed, starting in 2027 out to the end of the 10-year IA appraisal period (1.4 million in low scenario, 13.5 million in high scenario).

[208] DCMS, 2022, Online Safety Bill Impact Assessment, https://assets.publishing.service.gov.uk/media/6231dc9be90e070ed8233a60/Online_Safety_Bill_impact_assessment.pdf, pp 44

[209] Statista market analysis, https://www.statista.com/outlook/amo/app/united-kingdom

[210] The number of apps on each store is estimated from a range of public sources including statista.com, 42matters.com, bankmycell.com, pixalate.com.

[211] This was informed by an estimate that the two major app stores control 95% of the market - https://www.businessofapps.com/data/app-stores/ - which did not seem credible, given the estimates of apps on other major stores, nevertheless implied the percentage of apps covered by major stores was high.

[212] Based on estimates on 42matters.com. Without knowing the nature of the growth rate in app publishing, CAGR has been used as a conservative assumption.

[213] 90% of mobile apps are for kids aged 12 and under, 2020, Pixalate. https://www.pixalate.com/blog/google-apple-mobile-apps-for-kids

To determine the cost of a review, we used an estimate of each reviewer considering 100 apps a day, based on a major app store[214], and the wages of an IT user support technician[215] including a non-wage uplift of 22%.

This led to an indicative estimate of £7.5 million (in 2019 prices and with a 2020 base year) to review applications for age appropriateness in the central scenario over the 10-years of the IA appraisal period, though beginning in 2027. This translates to £1.4 million in the low scenario and £13.6 million in the high scenario.

Overall, the illustrative estimated costs for bringing application stores in scope of the regulation is reflected in the table below. However, it should be recognised that these are extremely preliminary estimates, based on a limited number of available sources, some of which are speculative. If this policy was developed further analysis would be required and completed when more was known. As an illustrative cost, these figures are not reflected in the EANDCB.

Table 27: Bringing app stores in scope of the regulation (2019 prices, 2020 base year – 10-year PV)

	Low	Central	High
Option 1: Cost of age verification	£4.4 million	£9.6 million	£19.8 million
Option 1: Cost of reviewing apps	£1.4 million	£7.5 million	£13.6 million
Option 1: Total cost of bringing apps stores in scope	**£5.8 million**	**£17.1 million**	**£33.4 million**

Break-out Box 21 - since the final stage IA: The OSB was amended to allow the Secretary of State to amend by regulations the Act for or in connection with the imposition on providers of Category 1 services of an alternative dispute resolution (ADR) duty.

ADR refers to schemes that are available to help complainants resolve their disputes out of court. The most common forms are mediation, where an independent third party helps the disputing parties to come to a mutually acceptable outcome, and arbitration, offered e.g. by Ombudsman schemes, where an independent third party considers the facts and takes a decision, often binding on one or other of the parties. ADR can offer a low-cost and fast alternative for consumers and businesses seeking to resolve disputes, which they cannot resolve between themselves.

At the time of writing, we do not know the type of ADR that the Secretary of State might require Category 1 platforms to engage in, nor the volume of complaints that might be subject to an ADR. Therefore, we're unable to estimate the potential impact of the Secretary of Stats adopting ADR.

Indicative costs can be found in other impact assessments of ADR policies. In 2021, the then Department of Business Energy and Industrial Strategy (BEIS) found that, according to ADR

[214] Apple's App Review Fix Fails to Placate Developers, 2022, Wired

[215] Annual Survey of Hours and Earnings, 2023, Office for National Statistics

providers, costs can vary strongly depending on the length and complexity of a case. It can be as low as around £120 for early resolution within days to £800 - £4,000 for long cases where inspections or expert opinions are needed. The department used £350, £450, and £550 as low, central and high estimates for external cost of ADR cases.[216] The department also assumed that familiarisation would be an hour for a senior executive and a further half an hour for 10 customer service representatives, across fewer than 20 Category 1 firms this cost is negligible.

In terms of potential benefits of ADR, BEIS cited commissioned research that found that 38% of court cases could have gone to ADR previously, had mandatory provision existed.[217] Estimating a cost per court case for businesses between £1050 and £1860, BEIS estimated significant savings (£700-£1315 per case) from using ADR.

Aggregate costs and benefits have not been estimated for this policy due to the lack of data on disputes that could be subject to ADR.

Break-out Box 22 - Since the final stage IA: OSB was amended to require Category 1, 2A, and 2B platforms to have a clear policy in their terms of service what their policy is about dealing with requests from parents of a deceased child for information about the child's use of the service. Categorised services must also have a dedicated helpline by which parents can easily find out what they need to do to obtain information in those circumstances. These duties also include allowing coroners to request access via Ofcom as part of their investigation.

Baseline

Though some platforms have processes in place for bereaved family to interact with a deceased person's account, there is no statutory right to receive information about social media activity.

For the costs of meeting coronial requests, we used ONS data on the deaths of people aged 5-19 in England and Wales by cause in 2022[218] and made a judgement on the causes of deaths where a coroner is likely to request access to the child's social media via Ofcom (i.e. where a child has died as a result of suicide, suspected suicide, or potential social media "challenges"). Available data for older teens covered 15-19-year-olds, which was adjusted to the number of 15-17 year-olds deaths likely to result in a coronial request using population estimates for the relevant ages, as well as proportionally upscaling deaths in England and Wales to reflect the UK.[219] This resulted in an estimate of 207 deaths in a year that might lead to a coronial request.

To appraise the cost of the policy, we applied the population-weighted proportions of children aged 5-17 that use social media[220] to derive an estimate of 189 cases where the coroner might request access via Ofcom. We multiplied the resulting number by the average number of social media accounts per person in the UK, 6.4.[221] The resulting number was multiplied by the estimated cost to platforms of compliance:

- the £0.37 cost of verifying the identity of a deceased child's parents/carers, comprising £0.07 cost of identity verification as stated in the OSB IA and a £0.30 cost of document authenticity testing.

[216] Alternative Dispute Resolution Impact Assessment, 2021, BEIS.

[217] ICF Consulting on behalf of BEIS. RESOLVING CONSUMER DISPUTES: Alternative Dispute Resolution and the Court System. 2018.

[218] ONS, 2022, Deaths registered in England and Wales – 21st century mortality

[219] ONS, 2018, National population projections by year of age

[220] Ofcom, 2024, Children and parents: media use and attitudes report

[221] Backlinko, 2024, Social media users, accessed June 2024

- the £14.64 cost of enabling a coroner to access to the social media account(s), comprising the wage costs of a record clerk/assistant taking one hour to give access to the accounts to the coroner, uplifted by 22% for non-wage costs, per RPC guidance.

The cost of providing social media accounts access for bereaved parents was estimated by accounting for the number of children who may pass away from all causes, using age-specific mortality rates[222], and assuming a similar cost for social media firms to allow them access to their child's social media account data.

Social media firms already have automated processes to allow users to extract the data platforms hold on them. We assume that they will leverage this process for coroners and bereaved parents allowing an hour for processing by a relatively low-level employee.

Three scenarios were modelled to derive low, central, and high estimates. The low scenario covers only coroners and parents of children in cases where a coronial request is likely, the medium scenario also considers 50% of children who have died through other causes, the high scenario adds all children who have died through other causes. Including an annual growth rate of 10.9% in social media accounts, this results in the cost estimates below -

Table 28: Deceased child duties (2019 prices, 2020 base year – 10-year PV)

	Low	Central	High
Option 1: Deceased child duties	£0.4 million	£0.8 million	£1.3 million

249. We expect the set-up costs to be similar across the board as all categorised services must make clear what their policies are for dealing-with such requests and complying with the duties. We expect that ongoing compliance costs will vary depending on the size of the service, the policy they adopt, and the number of requests. Ofcom will produce guidance to support providers of categorised services in complying with these duties.

Industry fees

250. Ofcom's operating costs will be paid through an annual industry fee. The annual industry fee is expected to be proportionate, with a threshold at or above which a provider would be required to notify and pay an annual fee. Providers below the threshold will not pay the annual fee, though will still be subject to the regulatory regime. An appropriate threshold will ensure small and medium enterprises are exempt from the direct costs of paying a fee. In determining fees, the regulator will refer to qualifying worldwide revenue and other factors that Ofcom consider appropriate. Parliament will approve regulations determining qualifying worldwide revenue and the threshold.

251. While the industry fees will depend on the realised costs to the regulator of operating the online safety regime, DSIT has worked with Ofcom to estimate a reasonable and realistic ten-year profile of operating expenditure based on current profiling for 2024/25, which is the first full year of implementation. This assessment estimates that the annualised industry fee on average could equate to £62.6 million per year and total £539.0 million across the appraisal period (2019 prices, 2020 present value base year). This includes the recouping of the initial costs incurred by Ofcom before the first fee charging year, over fiver years following the initial charging year.[223] Under Section 22(4)(a) of the Small Business, Enterprise and Employment Act 2015, taxes, duties, levies, and other charges are excluded from the Business Impact Target. This cost therefore has not been

[222] ONS, 2022, Deaths registered in England and Wales: 2021

[223] For analytical purposes, it is assumed that recouping costs will be equally spread across the five years following the initial charging period (years 4-7 of the appraisal period).

included in the calculations of the illustrative EANDCB (although it is included in the illustrative net present social value (NPSV)).

252.

Enforcement

253. Ofcom will have a suite of enforcement powers to act against platforms that fail to meet their regulatory responsibilities. These are: issuing confirmation decisions, imposing fines, requiring companies to make improvements, business disruption measures (including blocking via the Courts, in the most serious cases) and, in specific, tightly defined circumstances, offences for both providers and senior managers. Such enforcement powers will apply across different types of platforms, e.g. size, revenue, activity, and be used proportionately to potential or actual damage caused, and size and revenue. These powers also extend extra-territorially. Ofcom can investigate, take action against, and apply the full range of sanctions on operators (and, in some instances, executives) who are based outside the UK, but who provide services to UK users. Where relevant, it can refer decisions about sanctions to overseas courts for enforcement, It can also impose pursue court ordered business disruption measures (BDMs) which will (e.g.) disrupt an operator's financial viability in the UK, or block (restrict access) for non-compliant operators to UK users. Application of such measures will not require action beyond UK borders. Ofcom will be required to consult and produce guidance setting out how it will use its enforcement powers. The table below sets out details of these enforcement measures:

Table 29: Regulatory enforcement powers

Confirmation decisions	**Description**: Where Ofcom identifies a breach, it can issue a confirmation decision confirming the breach. These can set out the steps required by the company to come into compliance with their duties and the financial penalty being imposed (if any).
	Costs to platforms: As is standard practice in regulatory appraisal, this IA assumes full compliance with the regime. Therefore, any costs to platforms from rectifying actions undertaken because of receiving a confirmation decision is already captured in this IA's assessment of compliance costs. Confirmation decisions are therefore unlikely to result in material costs.
Fines	**Description**: Under the new regulatory framework, investigations conducted by Ofcom can end with an in-scope organisation being issued a monetary penalty for failing to comply with their duties. The approach to enforcement will aim to encourage compliance and drive positive cultural change. The regulator will support platforms to help them understand the expectations placed on them, and how the regulator's use of its enforcement powers will be proportional. Civil fines can be issued of up to £18 million or 10% of qualifying annual global turnover, whichever is higher.
	Costs to platforms: Fines and penalties are excluded from the Better Regulation Framework[224], for illustrative purposes, details of fines issued by the ICO for non-compliance with the requirements it enforces can provide an indication of the likely scale of impact. In 2020/21, the ICO issued three fines for contraventions of GDPR totalling £39.7 million. In addition, for non-compliance related to nuisance calls, the ICO issued 35 fines totalling £2.3 million.[225] Fines issued by Ofcom under the online safety

[224] Department for Business and Trade (2023) Better Regulation Framework guidance, p20

[225] Information Commissioner's Annual Report and Financial Statement (ICO, 2021)

	framework are expected to be rare but may be large if issued to large social media companies for example.
Business disruption measures	**Description**: In the most serious instances of non-compliance, Ofcom will have the power to initiate business disruption measures, to be used as a last resort. These include requiring third parties to withdraw key ancillary services (like payment or advertising services) to make it less commercially viable for non-compliant businesses (service restriction orders) and in some cases, restricting access to a non-compliant platform's service (access restriction orders).
	Costs to business: The frequency with which these measures are used depend on future codes of practice, the level of compliance, and the effectiveness of preceding regulator action on in-scope organisations, e.g. confirmation decisions and fines - all of which are unknown at this stage. To provide an indication of the likely scale of potential impacts, estimates from the IA for 'Age verification for pornographic material online'[226] which similarly involved notifying payment service providers and internet infrastructure providers, enabling them to withdraw their services and/or initiate blocking are presented. It was estimated here that the cost to payment service providers of working with the regulator and processing requests would be approximately **£0.5 million** per year and the cost to ISPs - based on a domain name system approach to blocking - was estimated to be between **£0.1 million** to **£0.6 million** per year.[227] In the context of the OSA, the two largest payment service providers engaged suggested that the cost of a small number of business disruption measures (as expected under the existing provisions) would be negligible to zero and would be absorbed into existing processes for responding to regulatory requests.
Senior management liability for managers who fail to ensure their company complies with Ofcom's information requests	**Description**: Ofcom will have the power to prosecute criminal sanctions against senior managers who fail to ensure their company properly complies with Ofcom's information requests.
	Cost to platforms: As above, this IA assumes full compliance with the regulatory framework, including full compliance with information requests. These sanctions are therefore unlikely to impact on costs.
Senior management liability for managers who fail to ensure - compliance with specific steps in confirmation decisions	**Description**: Ofcom will have the power to prosecute criminal sanctions against senior managers who fail to ensure their company properly complies with certain steps in confirmation decisions - specifically, where those steps relate to specified child safety duties (11(2)(a), 11(3)(a) and 72(2)(b)), and where those steps relate, whether or not exclusively, to a CSEA requirement, i.e. a step that relates, whether or not exclusively, to a failure to comply with specific illegal safety duties in respect of CSEA content. Failure to comply with such steps will be an offence.
	Cost to platforms: As above, this IA assumes full compliance with the regulatory framework, including full compliance with confirmation decisions. These sanctions are therefore unlikely to impact on costs.

[226] Age Verification for Pornographic Material Online, Impact Assessment - DCMS (2018)

[227] Converted from 2016 price and present value base year to 2019 prices and 2020 present value base year.

254. While impacts associated with regulator enforcement action are not monetised at this stage, the above information provides an indication of the likely scale of impact. In addition, given the disincentivising effect of, for example, large fines and damage to reputation, the government expects enforcement action which results in platform fines or business disruption measures to be rare.

Cost to individuals

255. The new regulatory framework will apply to companies or individuals who provide services that host UGC or enable P2P interaction, search engines, and online service providers that publish pornographic content. This is to futureproof the regulations as technologies develop which lower the bar to entry; and to prevent a loophole under which bad actors could make individuals (rather than companies) the service provider to evade regulation.

256. Given the low risk functionality exemption, the consultation stage IA noted that the vast majority (if not all) individuals are likely to be out of scope and that the then government did not have any evidence of individuals who could be in scope of the regulation. While the current scope (especially the low risk functionality exemption) is likely to have removed the vast majority of individuals, evidence provided in response to the consultation stage IA did highlight that at least some individuals will fall in scope. One individual noted that they would shut down their service because of potential compliance costs. The main concerns highlighted in evidence submitted include:

- not being able to engage legal services to both consider whether the platform is in scope and compliant and to ensure continued compliance (especially as the platform is already run at personal expense); and
- the financial risk posed by Ofcom's power to fine platforms for failing to comply. It was also noted that the OSA could potentially result in trolls purposefully flooding the platform with illegal content to overwhelm current moderation systems.

257. Given the significant risks around creating a regulatory loophole, and the suspected rarity of individual cases, the government does not consider it proportionate to exempt non-businesses from scope. However, Option 1 does include several provisions to ensure impacts on individuals are minimised and to avoid individuals shuttering online services, these include:

- regulatory expectations on services will be reasonable and proportionate to the severity of the potential harm posed and the resources available to the service. If the risk of harm on a platform is low, and the platform in question has little capacity, then regulatory burdens should also be minimal.
- Ofcom will be under an obligation to create codes of practice which are feasible and which cater for all service providers, whatever their size and capacity. This would include non-business services.
- Ofcom will have a legal duty to assess the impact of its codes of practice and other significant proposals on businesses and wider society which would include individuals within the scope of the regime.
- Ofcom will produce easy to use and easy to understand guidance supporting its codes to avoid the need for individuals and smaller services to seek legal advice.

- Ofcom will take a proportionate and targeted approach to monitoring and enforcement. It will focus on the services where the risk of harm to users is highest. It will seek to engage collaboratively with companies and individuals to help them understand their new duties, and what improvements might be needed, before initiating enforcement action, where this is required. Only in the most egregious cases where regulator engagement has failed is it expected that an individual operating a site would be subject to any of the financial enforcement mechanisms.

258. The government and Ofcom will continue to engage with individuals (and smaller services) through implementation to ensure any costs are minimal.

Cost to government

Justice impacts

259. The following aspects of Option 1 are expected to result in impacts on the criminal justice system (CJS):

- a power to introduce a new criminal offence for named senior managers who fail to respond to Ofcom's information requests.
- court orders required to apply for business disruption measures.
- a new appeals process, via the Upper Tribunal Administrative Appeals Chamber, heard using judicial review principles. Appeals will be made against enforcement decisions, designation as Category 1/2A/2B provider and the designation of companies in scope of the additional transparency reporting threshold.
- an impact on the number of incidences of online illegal activity and content reported to law enforcement and/or other authorities.
- additional criminal offences in the OSA, under Ofcom’s information gathering powers, including recklessly submitting false information, providing false or misleading information in interview or failing to attend, destruction of relevant data or falsifying data in response to Ofcom exercising its powers, and obstructing Ofcom accessing premises, data and equipment.

Break-out Box 23 - since the final stage IA: The during the passage of the OSB, Parliament added new offences for failing to comply with steps set out in a confirmation decision, specifically where those duties relate to child safety duties set out in (12(2)(a), 123)(a) and 81(2)(b)), and where those steps relate, whether or not exclusively, to a CSA requirement. The impact of these offences has not been appraised in terms of the cost of compliance, because full compliance with a confirmation decision is assumed.

A revised justice impact test has been conducted which assesses the new regulatory framework as having a *de minimis* impact. Current estimates indicate that the only costs occurring will be for establishing and operating the appeals system, with an estimate from the Ministry of Justice of £42,000 for the first year made up of £7,000 start-up cost and £35,000 running cost (which equates to £3,500 per case and 10 cases expected per year). For the purposes of the IA, the appeals body cost estimates are rounded up to £50,000. Ongoing costs for future years may be lower or greater and would be dependent on the number of cases being heard. Given the uncertainty around the number of future cases, this IA assumes justice impacts estimated here are constant across the appraisal period.

260. In addition to the new regulatory framework, the OSA also introduces a range of new and/or amended communications offences. These are:

- false communications offence;
- threatening communications offence;
- offence of encouraging or assisting serious self-harm;
- offences related to the showing of flashing images (epilepsy trolling);
- cyberflashing offence;
- intimate image abuse offences.

261. The new false and threatening communications offences are not expected to have a significant impact on the criminal justice system. The new offences are designed to update aspects of the existing communication offences and address limitations with the previous malicious communications offence to raise the criminal threshold.

262. The creation of additional criminal offences related to cyberflashing, showing of flashing images, encouraging and assisting serious self-harm and intimate image abuse has the potential to introduce additional costs to law enforcement and the criminal justice system. Additional impact assessments for these offences have been produced by the Ministry of Justice, the analysis of which is summarised into this IA. The total cost to the criminal justice system is estimated to be between £99.1 million and £240.0 million (central estimate = £170.0 million) over the appraisal period. However, these impacts are not incorporated into the NPSV of the business regulation.

Table 30: Justice impacts (2019 prices, 2020 base year)

	Low	Central	High
Option 1: Justice impacts	£99.1 million	£170.0 million	£240.0 million

Requirement to report CSEA

263. This section sets out estimated costs to the government for the body that will be responsible for receiving and processing CSEA reports, confirmed as the NCA. The costs will vary significantly depending on the number of reports that are made.

Break-out Box 24 - since the final stage IA: The estimate of the costs to government of reporting CSEA to the NCA have been updated to reflect new intelligence.

As outlined in the costs to business section, the NCA is estimated to receive an additional 6,400 reports per year over the period, ranging from 3,200 to 9,600. This assumes that platforms already reporting to NCMEC continue to do so, and do not change to reporting directly to the NCA. These figures do not include potential reports from platforms outside of the UK relating to UK-linked CSEA content (where platforms do not already report) and reports from search services, as outlined in the *reporting online CSEA to the NCA* section.

Since the NCA was confirmed as the designated body responsible for receiving reports, the NCA has incurred setup costs, including setting up technological systems and infrastructure to enable the NCA to securely receive, process and store industry reports and recruiting staff to support the setup of the system and assess new reports. These costs are not included in the NPSV where they are not expected to continue into the appraisal period (2025 onwards).

There will be ongoing costs relating to the assessment, triage and casework relating to reports by the NCA. Costs will occur to the NCA from the initial assessment of a report (estimated to take between 20 and 60 minutes). [228] Following this, some reports will not be progressed, for example if they include viral content, or the content in the report does not amount to a CSEA offence. Reports that are progressed will be triaged by the NCA according to the content of the report, with costs associated with further casework on reports progressed. This analysis bases the proportion of reports progressed following initial assessment and triaged into different casework routes on reports received by the NCA from NCMEC from 2020-2022. Currently the NCA only receives UK-linked CSEA reports from NCMEC, whereas in the direct reporting process, the NCA will also receive reports that relate to non-UK users. These reports will undergo an initial assessment but will be disseminated to international law enforcement following either the initial assessment or further investigation. This analysis is unable to account for costs associated with referring reports to international law enforcement due to a lack of information on the proportion of reports that will not be UK-linked, though they will often incur more time to disseminate than UK referrals. Therefore, the monetised costs below include assessment and triage, but not referrals to international law enforcement.

The NCA will also incur IT costs and storage costs for files relating to reports. IT costs are associated with the hosting platform for the direct reporting system and the upload and hosting of the image on the Child Abuse Image Database (CAID). Storage costs depend on the number of images and/or videos in each report. In 2022, as part of their reports, platforms submitted 49.4 million images into NCMEC, and 37.7 million videos. On average from 2021-2022, there were 1.5 images and 1.4 videos per NCMEC report. Following engagement with industry experts, an average image is assumed to require 6MB of storage, and a video is assumed to require 63MB of storage, with a growth rate in storage required of 12% per year.

CAID upload costs are estimated based on an average cost per upload using CAID costs from 2021-22.[229] The proportion of reports that may require uploading to CAID is uncertain as the NCA develop new processes for assessing and progressing cases. Estimates of 58% of reports requiring upload to CAID has been included and assumed. The potential range of costs is large and uncertain but is estimated at £9.9m (ranging from £5 - £14.9m). There will be further IT costs relating to technical infrastructure development over the appraisal period. Due to a lack of information on the costs of any potential infrastructure developments, this analysis is unable to provide estimates. However, any technical developments costs may reduce the time taken to assess and triage reports of CSEA content.

Total costs to the NCA resulting from staffing for the assessment and triage process, IT and storage costs are estimated at £19 million over the appraisal period, ranging from £12 million to £27 million (2023 prices). As previously stated, this does not include: reporting of UK-linked CSEA from non-UK companies; reporting from search services; disseminations to foreign law enforcement partners; or ongoing IT support and development, nor potential adjustments to the assessment and triage process going forward. In addition, there are other costs for the NCA in running the reporting service, engaging with reporters, providing information to Ofcom, liaising

[228] The range for the initial assessment of the report reflects that this is a new assessment process for the NCA. Over the period following implementation, the time taken to assess certain categories of reports may decrease, in which case associated staffing costs would decrease.

[229] These figures are taken from an unpublished NCA report held by Home Office

with overseas law enforcement and reporting centres, running the IT, managing the reporting data, etc.

Table 31: Costs to NCA as designated body for CSEA reporting (2019 prices, 2020 base year – 10-year PV)

	Low	Central	High
Costs to NCA	£7.0 million	£10.1 million	£15.1 million

As noted in the costs to business section, following a report of CSEA content, the NCA may have further engagement with businesses, for example on how to improve the information included in a report to ensure that the NCA can determine where the offence occurred and/or the appropriate law enforcement agency to receive the report (i.e. whether it should be referred to UK or international law enforcement agencies). Costs to the NCA associated with this engagement are unquantified, due to a lack of information on the proportion of businesses that the NCA would engage with, and the extent of that engagement.

There may be additional costs to law enforcement from an increased number of referrals from the NCA. Data from 2020-2022 indicates that on average 16% of reports that the NCA received were disseminated to law enforcement. Applying this to the potential increases in reports from the OSA could result in an additional 900 reports on average per year to law enforcement, ranging from 450 to 1,300 reports.

It is difficult to estimate the potential cost associated with each additional report to law enforcement, as some reports may instigate new cases, whilst others may be added to existing cases or to wider intel data. For context, analysis from a small sample of police forces in 2017 indicated that the average cost of an indecent images of children (IIOC) case for law enforcement (where a case includes 1 offender and 1 victim) could range between £2,200 and £5,700.[230] Inflating the estimate to 2023/24 prices results in a range between £2,700 and £6,800 per case, this does not include the prosecution, legal aid, court and sentencing costs.

The total NCA costs represented in the IA are based on projections of potential reporting volumes which are modelled on current NCMEC and NCA data. There are significant gaps in modelling that have been presented in the IA such as potential reporting volumes from international companies reporting UK and non-linked content and reports from search services that are not able to be quantified. The IT infrastructure and triaging processes will also need to develop over the ten-year appraisal period. NCA projections of costs are significantly more than the costs reported here but without additional evidence to quantify these and how they would impact over the appraisal period it has to be recognised that costs may be higher than estimated.

Benefits

[230] This is based on a IIOC investigation with 1 offender, 1 victim and limited digital forensics. Costs can vary where the number of offenders, victims and digital forensics work required differs.

264. The calculation of benefits is challenging: it is not possible to develop a precise estimate of the reduction in online harm that will be achieved by the policy options. This is due to:

- limited longitudinal data on the impact of internet use given the way in which the nature of the internet and its uses have evolved over time;
- the novelty of the proposed policy measures, which means there is a lack of relevant precedent in other sectors or countries;
- the scale of the internet and the way in which it is used, which means that it is not possible to run trials or experiments in a way that can be robustly scaled up;
- the rate of change in the sector and the way people use technology; and,
- ultimately, the regime will be implemented and operated by Ofcom as the independent regulator there is also uncertainty as to how platforms will change their behaviour in response to new regulation

265. This was the case at the time of writing the previous IA and therefore, the methodological approach remains the same. However, this IA also presents costs based on the most recent data. Estimated benefits are illustrative only and have not been included in the NPSV of the policy. The estimates below are the total amount of harm caused online under several categories of online harm; no attempt has been made to estimate the proportion of harm avoided by the introduction of this legislation (outside of illustrative scenario analysis). As with other similar uncertain policies, to address the problems with benefit estimations, this IA presents break-even analysis: estimating the reduction in online harm[231] required to exactly match the economic costs.[232]

Methodology

266. There are a wide range of different categories of harm in scope of the OSA, both illegal and legal but harmful to children. Of these, a total of seven defined categories of harm have been quantified, at least partially, based on available evidence. These include six illegal categories of harm:

- contact CSEA;
- modern slavery;
- hate crime;
- illegal sales of drugs;
- cyberstalking; and
- fraud as facilitated by UGC

267. One further category of harm that is legal but harmful to children has also been costed:

- cyberbullying

268. Quantitative evidence is provided to demonstrate the scale of the problem as well as more qualitative assessments based on expert judgement. These calculations rely on several uncertain assumptions, proxies and experimental data. They do not reflect a government view of the impacts, rather, they represent simplified, indicative estimates designed to enable analysis of online harm. One of the challenges of estimating the online element of illegal harm is that the way in which harm

[231] Harms which have been quantified and are therefore included in the estimated benefits are cyberbullying, cyberstalking, contact CSA, modern slavery, hate crime, drugs facilitated online.
[232] These costs comprise all monetised costs within this IA.

occurs varies, with some harm being purely online (e.g. viewing indecent images of children) and others taking place offline but being facilitated through online activity (e.g. grooming children online prior to a physical offence).

269. Regarding the categories of illegal harm that have been quantified, this IA uses data on the prevalence of crime from the ONS which includes experimental data based on the online crime flag. In 2015, it became compulsory for police forces to return quarterly information on the number of crimes flagged as being committed online (in full or in part). This does not provide information on the extent of the online component, that is, whether it was a significant or a minor part of the offence. It also does not provide information on whether, in the absence of the online component, the offence would still have taken place via alternative means.

270. Data quality has varied between forces. The data quality issues concern under-recording and a difficulty in defining online crime in a way that would yield useful data.

271. The online crime flag is currently being tested in the National Data Quality Improvement Service (NDQIS). This system is an automated process for flagging offences directly within police force systems. It has already been rolled out for the knife crime and domestic violence and domestic abuse flags. NDQIS aims to improve recording, while also reducing both human error and the burdens on police force staff in having to manually check records and apply flags.

272. Testing is progressing well and NDQIS plan to complete the rollout for online crime later in 2023. Current estimates from Home Office Data Hub data suggest 4% of total police recorded crime is committed wholly or partly online and this is believed to be an underestimate.

273. In addition, many of these categories of harm can involve both online and offline elements, which are often closely linked (e.g. traditional bullying and cyberbullying). It can therefore be difficult to completely disaggregate the impacts. Many of the harms are likely to have more than one source and while the OSA may address the online element, the harm may still occur through other sources. However, even where a harm may have multiple sources, a reduction or removal of the online element of the harm is not in and of itself inconsequential, and so there is still expected to be some degree of benefit to this occurring in these specific cases.

274. As well as issues relating to the use of the flag by police forces,[233] an additional limitation is that not all crime is reported and recorded by the police. Therefore, the Crime Survey for England and Wales (CSEW) is generally preferred as a source of data to establish the prevalence of crime since it allows for the measuring of “hidden” crime (that is, crime that is not reported and therefore that law enforcement does not come across). Consistent with other Home Office analysis, this IA uses a multiplier approach to uplift the ONS data to take account of actual levels of crime rather than just reported crime.

275. All quantified categories of illegal harm below contain the cost to the CJS, aside from Modern Slavery[234]. The CJS costs for cyberstalking are set out explicitly in the related cost table. For all other quantified categories of harm, the full methodology, including the costs covered, is set out in the associated Home Office statistics publication, each of which can be found in the footnote

[233] This is defined as “An offence should be flagged where any element of the offence was committed online or through internet-based activities (e.g. through email, social media, websites, messaging platforms, gaming platforms or smart devices)”. Source: Counting Rules Crime Flags, Home Office, updated April 2020. While the data shows an increasing volume of offences with an online flag, this is likely largely due to increased use of the flag rather than an increasing online component of crime.

[234] It has not been possible to estimate the cost to the CJS for several reasons. Modern slavery offences that go through the CJS are long and complex and can often take up to two years to complete. This is reflected in the proceedings data for these offences. The cost model that the Ministry of Justice used to estimate the cost of other crime types relies on a full set of data to profile the cost through the courts for a given year. Because of the lags from a criminal proceeding being commenced to its disposal, the data for all modern slavery offences produces results that are not reliable.

for each harm. For some types of harm there is inadequate quantitative evidence to enable the government to develop a rough estimate. This is because the true prevalence of harmful content or activity may be unknown, and because of the shortcomings of data that is available (for example, screen time does not reflect what that time was used for). In some cases, it was not possible to establish a causal link between online activity and the harm.

276. There is evidence that online harm is growing and many survey based measures show increases in the percentage of internet users experiencing harm such as cyberbullying, misinformation, online abuse, and other key categories of online harm.[235] In addition to individuals' experience, there is also evidence that the volume of harmful content is also increasing.[236] While, there is no single indicator for the prevalence and growth of online harm, any estimate will by definition be only an attempt to mirror the real growth of online harm across a ten-year period. Potential growth rates include:

Table 32: Potential online harm growth rates

Measure	Average annual growth
Hours spent online (Ofcom)[237]	4.5% per year between 2015-2020
Total number of videos viewed online (EY)[238]	5% per year since 2017
Percentage of adult population that has recently used the internet (ONS)[239]	1.31% per year between 2015-2020
Adults that have had potentially harmful online experiences in the last 12 months (Ofcom)[240]	1.6% between 2019-2020
Percentage of adult population that have recently accessed social networking sites (ONS)[241]	5.1% between 2011-2020 or 4.2% between 2015-2020

277. Based on the available proxies above, this IA estimates that online harm will grow by 3% per year. The sensitivity section below uses the full range of proxy growth rates (1.3 - 5.1%) to illustrate the effect on the break-even point.

Quantified harm

Contact CSEA

278. The government is aware that placing a monetary value on abuse may seem reductive to those that have experienced CSEA and recognises the profound human costs of CSEA to victims and survivors. The impacts of CSEA and other harms are only monetised to put the impact to business in context, in compliance with overarching government standards, and are not intended to represent the full range of impacts for victims and society from these crimes.

[235] Internet users' concerns about and experience of potential online harms (Ofcom & ICO, 2018-2020); Leading Bullying Research (Ditch the Label).
[236] National Center for Missing and Exploited Children, by the numbers (NCMEC)
[237] Ofcom's Adults' Media use and attitudes reports (2015-2020)
[238] Understanding how platforms with videosharing capabilities protect users from harmful content online (EY, 2021)
[239] Internet users, UK statistical bulletins (ONS, 2015-2020)
[240] Internet users' concerns about and experience of potential online harms (Ofcom, 2019-2020)
[241] Internet access - households and individuals (ONS, 2011-2020)

279. Contact CSEA[242] was estimated to result in costs of at least £10 billion in the year ending 31 March 2019.[243] While online abuse can be a feature of both contact and online CSEA, the overlap is complex and difficult to disentangle. It is not possible to quantify the harm from online abuse alone nor for non-contact abuse. Estimates for contact CSEA are available and include the financial and non-financial (monetised) cost relating to all victims who continued to experience contact sexual abuse, or who began to experience contact sexual abuse, in England and Wales in the year ending 31 March 2019. As this cost is of victims whose abuse lasts multiple years it is not a true annual cost and as such it needs to be divided by the average length of a CSEA case to produce an annual estimate.

280. Using the CSEW, that looked at the time abuse lasted for adults who had been abused as children,[244] this IA can estimate the average time. Two assumptions are needed to calculate the average, first that abuse for adults who responded to the survey is representative of the average length of abuse for current victims. Second, it is assumed that for those that selected 'abuse lasted for less than a year' the abuse lasted one day. This is a conservative minimum that has been chosen due to the absence of evidence suggesting a longer period of abuse. This results in an estimated length of abuse of 2.17 years.

281. Dividing the estimate for the cost of contact CSEA by 2.17 gives an annual cost of contact CSEA of £4.93bn (2021/22 prices). This cost includes the lifetime impacts of contact CSEA victims. This cost is for all forms of contact CSEA and not just CSEA that has an online element, it does not include the cost of non-contact offending. The estimated proportion of contact CSEA that includes an online element is estimated using the online crime flag. This assumes that the proportion of recorded offences with online elements is similar to the proportion of victims that experience abuse with online elements. This assumption is used as the online flag is the best available proxy for estimating how much of the total cost of contact CSEA may be attributable to CSEA with an online element. Police flagging data from April 2022 to March 2023 estimated that 19% of recorded contact CSEA offences had an online element (this excludes Devon and Cornwall who were unable to supply data). In this analysis it is assumed that this proportion is constant across the appraisal period.

282. The true level of online offences may be higher than 19% due to issues with the flag. First, the flag is typically manually applied by officers, and therefore accurate use relies on officers being aware of the flag, remembering to apply it to specific cases, and recognising that an online element is present. This can be difficult in some cases, such as where online messaging services like WhatsApp or Kik are used, which officers may not recognise as 'internet enabled' or online. Second, the flag differs in usage between forces and is not evenly applied throughout England and Wales. In addition, online technology has proliferated over the last decade, which gives offenders more opportunities to target children, with 88% of children having a smartphone aged 12[245]. Trends like this may increase the opportunities for online facilitated contact but as we are unable to disentangle links between online and offline offending we cannot account for such changes.

283. The table below summarises the data and calculations used to estimate the impact of contact CSEA with an online element. This is calculated by multiplying the estimated annual cost of all contact CSEA (£5.41bn) (uplifted to 2023/24 prices) by the estimated proportion of contact CSEA that includes an online element (19%). This gives the estimated annual cost of contact CSEA with an online element (£1.03bn), or £7.63bn when considered over the 10-year appraisal period, with a 2020 present value and 2019 prices, for comparability with costs.

[242] For definition of contact CSA, see page 107 in Working Together to Safeguard Children (HMG, 2018)

[243] Tackling Child Sexual Abuse Strategy 2021 (HMG)

[244] Child sexual abuse in England and Wales: year ending March 2019 (ONS, 2020)

[245] How many children have their own tech? (YouGov, 2020)

Table 33: Online contact CSEA costs (2019 prices, 2020 base year – 10-year PV)

Harm	Estimated annual cost	Proportion online	Cost with online elements
Contact CSEA	£40,200 m	19%	£7,630 m

284. It is worth further emphasising that this cost (£1.03bn) is a likely underestimate for the true scale and impact of contact CSEA with an online element and does not account for costs for non-contact abuse that may have an online element. The impact may be greater because the true scale of contact CSEA facilitated online is based on the irregular use of the online flagging tool and the fact that estimating the full cost in all areas of abuse is difficult and sometimes unquantifiable.

Modern slavery

285. This section considers the economic and social cost of physical modern slavery offences with an online element.

286. The unit cost of modern slavery is £329,720.[246] It covers the costs of physical and emotional harm, the cost of lost output and time, costs to health services, costs to victim services and law enforcement costs. This unit cost is given in 2016/17 prices. Inflating the estimate to 2023/24 prices provides an estimate of £399,000. This cost relates to physical modern slavery offences. It is assumed, for the purposes of this analysis, that modern slavery offences do not take place solely online. There could theoretically be a scenario where the definition of modern slavery could be met with an entirely online situation, but that would be unusual and infrequent. The facilitator needs to somehow benefit which could be difficult virtually.

287. It is important to note that the cost of modern slavery is calculated on a victim basis. It is a cost of new cases of modern slavery identified between April 2016 and April 2017 that may continue into succeeding years and does not capture those ongoing identified prior to the identification year. It is difficult to ascertain whether this will mean this an over or underestimate for the cost that modern slavery has to society as the average length of modern slavery cases vary between exploitation types. The median durations of 'labour exploitation' and 'sexual exploitation' are both 274 days, whereas 'domestic servitude' lasts 730 days on average.[247]

288. This unit cost can then be applied to an estimate of modern slavery offences with an online component, to provide an estimate of the impact of these offences. This estimate does not involve any judgement as to the extent of the online component, or what would happen in the absence of the online component. It simply reflects an estimate of the cost associated with modern slavery offences flagged as having an online component.

289. The approach used above is to use the police recorded crime data, where there were 10,229 recorded modern slavery offences between April 2022 and March 2023.[248] This is then multiplied by the proportion of cases flagged by the police as having an online element (0.88%) to give 90 cases with an online element. Applying this to the unit cost of modern slavery indicates an annual cost of **£35.9 million** or £267 million when considered over the 10-year appraisal period, with a 2020 present value and 2019 prices, for comparability with costs. This figure is likely to underestimate the true prevalence of modern slavery with an online element given the limitations of the online crime flag.

[246] The economic and social costs of modern slavery (Home Office, 2018)

[247] The economic and social costs of modern slavery (Home Office, 2018)

[248] Police recorded crime and outcomes open data tables (Home Office, updated 2023). Note that this excludes figures from Devon & Cornwall as they were unable to submit a full year's data.

Table 34: Modern slavery costs (2019 prices, 2020 base year – 10-year PV)

Harm	Prevalence	Unit Cost	Total Cost
Modern slavery with an online element	90	£0.3 m	£267m

Hate crime

290. The law recognises five types of hate crime based on race, religion, disability, sexual orientation or transgender identity. Any crime can be prosecuted as a hate crime if the offender has either demonstrated hostility or been motivated by hostility based on any of these characteristics. Offence categories include violence against the person (VATP), public order offences, criminal damage and arson offences, and other notifiable offences.

291. To obtain a proxy measure for the number of online hate crimes, this IA looks at offences measured by police as being racially or religiously aggravated and also flagged as having an online component. The quantification of harm is focussed on VATP, as this represents the majority of racially or religiously aggravated offences that are flagged as online and cost data is not available for the other offences. As discussed above, hate crime could also be motivated by other factors such as sexual orientation, disability, and transgender identity. Therefore, this cost estimate is likely to underestimate the total cost of online hate crime. The table below summarises the data and calculations used to estimate the impact of hate crime with an online component.

Table 35: Hate crime online cost (2019 prices, 2020 base year – 10-year PV)

Harm	Prevalence	Unit cost	Total cost
Hate crime: racially or religiously aggravated offences with injury	87	£14,800	£11.1 m
Hate crime: racially or religiously aggravated offences without injury	1,641	£6,270	£90.5 m
Total			£101.7 m

Figures may not add up due to rounding

292. The prevalence of online racially or religiously aggravated offences was calculated using statistics from police recorded crime from April 2022 to the end of March 2023. Within this period, there were 75,128[249] racially or religiously aggravated offences recorded by police forces within England. Of these racially or religiously aggravated offences, 3,795 offences were 'with injury' offences, and 71,333 were 'without injury' offences.

293. To calculate the proportion of online racially or religiously aggravated offences within the overall category of racially or religiously aggravated offences, the police online crime flagging tool data is used. The estimate of the proportion of offences that were committed online or enabled by online devices is 2%[250]. This gives a prevalence of 87 racially or religiously aggravated with injury offences, and 1,641 racially or religiously aggravated without injury offences which have the online harm flag applied.

[249] Police recorded crime and outcomes open data tables - GOV.UK (www.gov.uk)

[250] The estimate of 2.3% (rounded 2%) is from the 2017/18 Hate crime in England and Wales statistical bulletin. Hate crime, England and Wales, 2017 to 2018 (publishing.service.gov.uk)

294. This figure is likely to underestimate the true prevalence of online racially or religiously aggravated offences in the UK given that these offences often go unreported and previously noted issues with the online crime flag. Additionally, the proportion of online racially or religiously aggravated offences is low (2%) due to a high proportion of racially or religiously aggravated offences being only able to be committed offline, with 69% of all racially or religiously aggravated offences being racially or religiously aggravated public fear, alarm or distress offences which are highly unlikely to be committed within the online sphere. This means the proportion of online racially or religiously aggravated offences recorded by police may be significantly lower than the amount of online hate crime committed.

Break-out Box 25 - since the final stage IA: New data has allowed alternative estimates of racially or religiously aggravated offences. Applying an alternative estimate of the proportion of offences that were committed online enables us to consider sensitivity analysis. The Hate crime bulletin (2017/18) also mentions that 6%[251] of online crimes were recorded within the VATP hate crime offence group. Using this proportion gives us a prevalence estimate of 213 racially or religiously aggravated with injury offences and 3995 racially or religiously aggravated without injury offences which have the online harm flag applied. This would give us an alternative annual cost of £3.7 million and £29.6 million respectively (and an alternative total cost of £33.3 million). This higher estimate has been used in the “low” scenario (i.e. high benefits from reduction, low break-even point).

Illegal sales of drugs

295. Combating the sale of drugs online is a key element of Option 1, with drugs increasingly being sold using a variety of online methods including via social media. Using police online crime flagging data, it was estimated that around 1.9% of all drug supply offences (using offence code 92A; trafficking in controlled drugs) are conducted using online or online enabled methods.

296. These police-recorded figures are likely to underestimate the prevalence of online selling, as other sources suggest that although the share of drugs sold online is relatively low, it is likely to be higher than the recorded figures suggest. The Global Drug Survey 2021 found that around 10% of people buy their drugs online, and 2.3% of adults in England and Wales who had taken drugs in 2021/22 reported their source as the internet[252] (excluding social media, due to methodology, and the dark web).

297. As unit cost data is unavailable for this harm, a top-down approach has been taken to estimate a proxy value for the impact instead. The total social and economic cost of organised drugs supply is estimated to be £20 billion[253]. Inflating this figure from 2015/16 prices to 2023/24 provides a total estimate of £25.0 billion. Applying the proportion of recorded drugs offences flagged as online (1.9%) and the share of adults that buy drugs online (2.3%) to this total cost provides an indicative estimate of the cost of online drugs offences of around £469 million - £567 million, or, using the central estimate, £3.85 billion when considered over the 10-year appraisal period, with a 2020 present value and 2019 prices, for comparability with costs. This is likely to underestimate the true impact, given that the sources used for both the lower and upper bound are likely to underestimate the share of drugs sold online, however, the central estimate has been used for simplicity.

[251] The estimate of 5.6% (rounded 6%) is from the 2017/18 Hate crime in England and Wales statistical bulletin. Hate crime, England and Wales, 2017 to 2018 (publishing.service.gov.uk)

[252] Drug misuse in England and Wales - Appendix table - Office for National Statistics (ons.gov.uk)

[253] Understanding organised crime 2015/16 second edition (publishing.service.gov.uk)

Table 36: Illegal sale of drugs online cost (2019 prices, 2020 base year – 10-year PV)

Harm	Share of drug offences	Unit cost	Total cost
Illegal sale of drugs online	1.9% - 2.3%	n/a	£3,850m

Cyberstalking

298. There is no single definition of cyberstalking, however it is widely used to refer to the repeated use of online communications tools to stalk, harass or frighten a victim. Currently there is no formal governmental definition of cyberstalking and this makes collecting data specifically on this area difficult. Within this analysis, a measure from the CSEW has been used to estimate the prevalence of cyberstalking. This is based on the answer to two questions included in the CSEW Survey asking;

1. if anyone "has sent more than one unwanted email or social network message that was obscene or threatening"
2. or "put personal, obscene or threatening information about you on the internet on more than one occasion and which caused you fear, alarm or distress?"

299. This represents a change in methodology since the previous IA, which proxied the prevalence of cyberstalking using the proportion of stalking offences with the online flag. The CSEW measure is used instead as it is currently the best metric available, however it is important to note it may not reflect the true prevalence of all aspects of this crime. The table below sets out the data and calculations used to estimate the impact of cyberstalking.

300. The unit cost of a cyberstalking incident is based on the cost to a victim of a stalking incident from a 2019 Home Office report. Using the work of Paladin, the national stalking advocacy service, cyberstalking inflicts the same amount of psychological damage as offline stalking and therefore it was deemed appropriate to use the costings relating to all stalking in this analysis. The unit cost comprises three elements (prices in 2016/17 prices): emotional cost to the victims (£21,920), cost to health services (£1,210) and cost in lost productivity (£6,560). The total has been uplifted to 2023/24 prices.

301. The latest CSEW includes published estimates of the number of cyberstalking victims annually. This estimates 807,000 cyberstalking victims in the year ending March 2022. Our modelling assumes the number of victims in 2023/24 remains in line with this. There has been a further publication of CSEW data to cover the year ending March 2023, but the estimates specifically related to cyberstalking have not yet been published and therefore the year ending March 2022 data has been used.

Table 37: Cyberstalking cost (2019 prices, 2020 base year – 10-year PV)

Harm	Prevalence	Unit cost	Total cost
Cyberstalking	807,000	£30,800	£218,000m

Fraud

302. There were 3.5 million instances of fraud in England and Wales in the year ending March 2022[254] and over 70% of these are estimated to have an online element.[255] However, the legislation does not cover every form of fraud which is cyber-enabled or cyber-dependant and therefore, further consideration has been taken to provide a better estimate of the proportion of fraud which Option 1 could potentially address.

303. The Economic and Social Cost of Crime estimates that the average cost per fraud incident is £1,611. This impact assessment draws on three main data sources to assess the impact of fraud, namely the National Fraud Intelligence Bureau (NFIB), Action Fraud (AF) and the Crime Survey for England and Wales (CSEW). Note the CSEW presents a prevalence estimate, whereas NFIB analysis uses fraud reports. Due to the limitations of the data available we had to engage with expert colleagues to generate an estimate for the proportion of fraud in scope of the legislation.

304. Fraud is highly underreported with only 13% of the estimated CSEW offences reported to Action Fraud in the year ending March 2022. As such, the approach taken has been to downscale the overall prevalence estimate the CSEW provides to remove frauds which are likely to have been out of scope of the OSA. Based on the CSEW and NFIB estimates for frauds having an online element, as well as engagement with stakeholders and colleagues, the scale of frauds in scope has been estimated to be 45% as a mid-estimate, allowing for the exclusion of email enabled fraud. Low and high estimates are also given to highlight the uncertainty around the mid-estimate. It should be noted that whilst the OSA covers 45% of fraud presently and although we expect the OSA to have a significant impact on introduction, it is reasonably likely that fraudsters will displace to alternative means of defrauding victims. This approach has the following limitations:

- The unit cost of fraud is taken from 2015/16 data and could be outdated. Additionally, this method applies the same unit cost to all fraud types.
- This method uses historical fraud prevalence estimates, and it is possible that fraudsters may divert to other methods to avoid detection because of the OSA, such as email scams which are out of scope.
- This method may underestimate benefits if companies work beyond the scope and have a greater impact.
- The estimate of scale is based on stakeholder and expert engagement, rather than a reliable data source due to limitations in what is available.

Table 38: Fraud cost (2019 prices, 2020 base year – 10-year PV)

	% potential in scope of Option 1	Number of offences	Total cost
Low	30%	1.0 million	£12,600 million
Mid	45%	1.6 million	£18,900 million
High	60%	2.1 million	£25,100 million

305. For the main calculations the mid estimate is taken; however, both the low and high estimates are tested in the sensitivity analysis.

[254] Crime Survey for England and Wales - year ending March 2022
[255] Internal NFIB estimate

Cyberbullying

306. Cyberbullying is defined as bullying which takes place over digital devices, such as mobile phones, tablets and computers. Cyberbullying can be both public and private, acting on public forums or through private messaging.[256] Cyberbullying can take the form of many behaviours including: harmful messages; impersonating another person online; sharing private messages; uploading photographs or videos of another person that leads to shame and embarrassment; creating hate websites/social media pages; and excluding people from online groups.

307. Whilst the lines between cyberbullying and traditional bullying can sometimes be blurred, online bullying does have several elements that make it different from traditional bullying. Cyberbullying can occur day and night and may be seen and shared by a much wider audience. A cyberbullying incident may also have a much longer lasting impact. Further, anonymity can make cyberbullying incidents more intimidating, and the degree of separation between bully and victim can make it hard for perpetrators to appreciate the impact of their behaviour.[257]

308. Given most academic research available focuses on the impact of cyberbullying on young people, the estimates used in this analysis focus on the impacts on those aged 10 to 15 years old (based on the age range typically used in cyberbullying studies). Therefore, the estimate will underestimate the impact of cyberbullying on the UK. The table below outlines the core unit costs for the central estimate of the economic costs of cyberbullying.

Table 39: Cyberbullying cost (2019 prices, 2020 base year – 10-year PV)

	Category	**Prevalence**	**Unit cost**	**Total cost**
Cyberbullying	Direct impact on victim	911,587 (19% of 10-15 year olds)[258]	£640	£5,600m
	Cost to health services of treating related depression	84,996 children accessing specialist mental health treatment	£354*	£289m
	Cost of treating cyberbullying related self-harm	*1,943 children*	*£838**	*£18m*
	Total	**911,587**	**£673**	**£5,880m[259]**

*Assumption of one incident each year

309. There were an estimated 4.8 million children in the UK aged 10-15 in 2020.[260] Estimates for the proportion of children who have experienced cyberbullying can vary depending on the study used. The most up-to-date studies from 2020 that address the question of prevalence are from the ONS and Ofcom. The ONS estimate that 19% of 10 to 15 year olds were cyberbullied in the year

[256] What is Cyberbullying? - (StopBullying, 2018)

[257] Bringing an end to online bullying: Whose job is it anyway? - (Anti-Bullying Alliance, 2019)

[258] Online bullying in England and Wales Online bullying in England and Wales: year ending March 2020 (ONS, 2020)

[259] This figure does not include the cost of treating cyberbullying related self-harm - this is shown illustratively only given the minimal evidence base. Other numbers may not sum to total due to rounding.

[260] ONS Population Estimates (ONS, 2020)

ending March 2020,[261] whilst the Ofcom figure is 26% and covers 12 to 15 year olds.[262] The ONS prevalence figure is used in this IA as it covers a broader age range of children. This central prevalence estimate of 19% equates to 911,587 child victims of cyberbullying in the UK in a given year.

310. Based on the range of prevalence estimates in the studies observed,[263] sensitivity analysis is conducted using 7% and 26% as upper and lower bounds which results in lower and upper bound estimates of 336,000 and 1,250,000 cyber bullied children. The costs to the victim of a cyberbullying incident include the impact on the victim's mental health and wellbeing, which may result in a depressive episode. This impact is estimated using quality-adjusted life years (QALYs), which enables quantification (in monetary terms) of the impact of various health conditions on a person's quality of life.

311. To estimate the cost of a minor/moderate depressive episode required information includes:

- ○ the likelihood of sustaining depression (LIKE);
- ○ the percentage reduction in quality of life (REDUCEQL);
- ○ the duration of the depressive episode (DUR) as a fraction of a total year; and
- ○ The value of a year of life at full health (VOLY).[264]

312. These are multiplied together to give an estimate of the average cost associated with the crime. On this basis, the depression associated with non-violent crime, which is used here as the closest available proxy for the impact of cyberbullying, has a QALY loss (REDUCEQL) of 14.5%.[265] The duration (DUR) is estimated at 0.167 years (or 2 months) and a value of a life year (VOLY) of £71,385 (uplifted to 2019 prices). Therefore, the unit cost is 0.145 * 0.167 * £71,385 = £1,728.[266] This £1,728 unit cost can then be multiplied by the probability of harm occurring (LIKE) – that is, what proportion of victims of cyberbullying suffer depression as a result. An annual bullying survey in 2017 found that 37% of those who were victims of cyberbullying went on to suffer from depression.[267] This can then be multiplied by the total number of cases to give an estimate of the personal cost (in terms of quality of life reduction) to the individual.

313. As outlined above, it is estimated that 37% of cyberbullying victims go on to suffer depression as a result based on Ditch the Label's Annual Bullying Survey. This gives a central estimate of 337,000 children per year suffering from depression because of cyberbullying. Currently, NHS digital research has found that only 1 in 4 children (25.2%) who report having mental health

[261] Online bullying in England and Wales Online bullying in England and Wales: year ending March 2020 (ONS, 2020)

[262] Internet users' experience of potential online harms: summary of survey research (Ofcom and ICO, 2020)

[263] A number of studies were reviewed as part of work to understand the likely prevalence of cyberbullying. As well as the ONS and Ofcom studies already mentioned the following studies were also reviewed: Annual Bullying Survey (Ditch the Label, 2017); Mental Health of Children and Young People in England (NHS Digital, 2017); The Suffolk Cybersurvey (2017) Bullying in England, April 2013 to March 2018 Analysis on 10 to 15 year olds from the Crime Survey for England & Wales (DfE, 2018). For sensitivity, an estimate was also produced looking at a wider range of studies between the years 2013 and 2017 which also produced an average prevalence of 17%.

[264] Valued at £60,000 by the Department of Health (DfE) and referenced in HMT Green Book (page 72) in 2012 prices. Uplifted to 2019 prices, giving a value of £71,385.

[265] This represents the estimated impact of a mild episode of a depressive disorder - see below footnote for further information.

[266] The Economic and Social Costs of Crime (Home Office, 2018). The estimate for 'REDUCEQL' originally comes from Disability weights for the Global Burden of Disease 2013 (Saloman et al., 2015). The duration (DUR) (0.167 years or two months) is an average originally derived from Impact of crime on victims (Wasserman and Ellis, 2007).

[267] Annual Bullying Survey (Ditch the Label, 2017)

problems access specialist mental health services.[268] It is assumed this is also the proportion of cyberbullied children who have developed depression that access mental health services. The National Institute for Health and Care Excellence (NICE) estimates the following costs for the treatment of depression:[269]

- a referral for psychological treatment: £14.50;
- of the referrals, 67% accept the psychological treatment;
- 60% of these are low-intensity interventions at a cost of £45; and
- 40% of these are high intensity at a cost of £1,125.

314. This gives an average cost from referral through to treatment for all patients (including those who are referred but don't subsequently take up full treatment) of £334.09 per person for a single treatment. Once uplifted to 2019 prices using a GDP deflator, this is £354.37 per patient on average. This IA also assumes each individual accesses an intervention once per year - it is quite likely that a proportion of those seeking treatment may be treated multiple times and so this assumption is conservative.

315. Assuming those children who have suffered depression due to cyberbullying access care in similar proportions to all children with mental health problems (25.2%), this would give an annual cost of cyberbullying to health services of **£30.1 million.**

316. The 2017 Annual Bullying Survey found that 25% of cyberbullying victims surveyed went on to self-harm. This implies that around 228,000 children per year self-harm because of cyberbullying. A large proportion of self-harm incidents will go unnoticed or treated (there is a three-fold difference in prevalence of self-harm as reported by young people and by their parents, suggesting that many acts of self-harm in the young do not come to the attention of their families). As such, information on how many of these children formally seek help or attend hospital as a result is uncertain. Based on a study in the Lancet, the average cost to UK hospitals of treatment of self-harm is £809 per incident.[270] Uplifting this to 2019 prices yields a cost per incident of £838. Given the uncertainty above, this IA assumes a conservative proportion of those self-harming due to cyberbullying require hospital treatment (1% or 1,943 cases), this would result in an annual cost to the NHS of £1,910,000. Given the difficulty in ascertaining exactly how many of those who self-harm due to cyberbullying would go on to require NHS treatment, this cost is only included as an illustrative upper estimate of cyberbullying, and is not included within the total estimated cost in the table above.

317. In the previous IA, an estimate for the lifelong impact of cyberbullying was included which explored the long-term economic impact associated with childhood bullying. As this estimate was based on a single academic study,[271] and was not included in the central cost estimate for cyberbullying, it has been decided not to include this estimate again until a more comprehensive and established evidence base around the long-term effects of cyberbullying becomes available.

Qualitative benefits

318. The approach taken in this IA is to attempt to quantify a subset of online harm which occurs under the baseline and conduct break-even and scenario analysis. Data on harm is limited and this IA is only able to monetise a small subset. This section presents evidence on additional harm for

[268] Mental Health of Children and Young People in England, 2017 (NHS Digital, 2018)

[269] Resource impact statement: Depression and anxiety disorder (NICE, 2015)

[270] General hospital costs in England of medical and psychiatric care for patients who self-harm: a retrospective analysis (Tsiachristas et al., 2017)

[271] Long Term Economic Impact Associated with Childhood Bullying Victimisation (Brimblecombe et al., 2018)

which evidence is insufficient to include within the break-even analysis, and several non-monetised benefits expected to accrue because of Option 1.

Break-out Box 26 - since the final stage IA: New data allows a revised assessment of the impact of certain online harms.

Terrorism/extremism

Terrorism has an impact on survivors, witnesses, and loved ones which the OSA seeks to mitigate. These individuals experience a wide range of impacts, including post-traumatic stress and anxiety, life changing injuries, and financial hardship.[272] Terrorist attacks can have profound impacts that extend across society.[273] The online space can be used to spread propaganda designed to radicalise, recruit and inspire people, and to incite, provide information to enable, and celebrate terrorist attacks.

The UK's counter-terrorism strategy, CONTEST 2023, assesses that the overall threat from terrorism is enduring and evolving and the risk from terrorism is once again rising.[274] The terrorist threat in the UK is dominated by individuals or small groups who may sometimes be inspired or encouraged by organised terrorist groups but are acting without their direction or material support. For more than a decade, every terrorist attack in the UK has been from such individuals. Terrorist groups use propaganda to encourage susceptible individuals to commit acts of terrorism on their own initiative.[275] This is exacerbated by online environments which bring together and facilitate individuals sharing and validating thoughts and ideas.[276]

The internet continues to make it simpler for individuals and groups to promote and to consume radicalising content. The barriers to entry of in-person terrorist group activity have been replaced with an online environment built for ease of access and unrestricted by geographical location. This has increased the accessibility for everyone to spaces populated by radicalisers and terrorist content, including groups such as minors or those with mental ill-health or neurodiversity conditions.[277]

Latest evidence suggests that the internet is playing an increasingly prominent role in radicalisation processes. Analysis of individuals convicted of TACT or TACT-related offences in England and Wales between 2005 and 2017 found that online radicalisation was the predominant pathway for an increasing proportion of TACT offenders. Over the period, the percentage of extremist offenders who were subject to some degree of online radicalisation increased from 35% in 2005-2009 to 64% in 2010-2014 and then to 83% in 2015-2017.[278] This includes those who were primarily radicalised online and those who were radicalised through a combination of online and offline influences

The internet has also continued to provide resources to enable attacks. This includes instructional material and access to the components needed to construct an improvised explosive device (IED).[279]

[272] García-Vera, M. P., Sanz, J., & Gutiérrez, S. (2016). A systematic review of the literature on posttraumatic stress disorder in victims of terrorist attacks. *Psychological Reports, 119*(1), 328-359. https://pubmed.ncbi.nlm.nih.gov/27388691/

[273] Counter-terrorism strategy (CONTEST) 2023. https://www.gov.uk/government/publications/counter-terrorism-strategy-contest-2023

[274] Counter-terrorism strategy (CONTEST) 2023. https://www.gov.uk/government/publications/counter-terrorism-strategy-contest-2023

[275] For example, provision of specific direction, training, money, weapons or expertise.

[276] Counter-terrorism strategy (CONTEST) 2023 - Home Office, 2023

[277] The Challenge of Understanding Terrorism in a New Era of Threat, The RUSI Journal (Vol.168, No. 4, 2023), pp. 1-9, p.5.

[278] Exploring the role of the Internet in radicalisation and offending in convicted extremists. Ministry of Justice Analytical Series, 2021

[279] Counter-terrorism strategy (CONTEST) 2023 - Home Office, 2023

The rapid proliferation of terrorist content on multiple online services continues to play a significant role in exposing individuals to online spaces and communities encouraging and glorifying violent acts. Live stream videos of attacks are especially potent and harmful.[280] An example of this is the 15 March 2019 terrorist attack in Christchurch, New Zealand which led to 51 fatalities and 50 injured. The perpetrator live streamed the attack which lasted 17 minutes and was viewed approximately 4,000 times before being removed.[281]

It has not been possible to quantify or monetise the impact of the OSA on terrorism. This is because it is not possible to say exactly how the OSA will affect the use of the internet in facilitating radicalisation, providing resources and instructional material to commit attacks. Therefore, the exact extent to which harms to society will be mitigated cannot be estimated. Additionally, although the internet is playing a growing role in facilitating radicalisation, it is not possible to estimate the extent to which this drives the number of attacks. As such, even if the OSA is impactful in mitigating this, it is not clear what impact this would have on the number of terrorist attacks.

Estimates of the economic cost of terrorist attacks indicate a broad sense of scale as to the potential benefits if the OSA led to the prevention of a terrorist attack. The five terrorist attacks which took place across the UK in 2017 are estimated to have cost £172 million in direct costs.[282] Separate analysis by RAND Europe estimates potential indirect impacts on GDP of up to £3.4 billion.[283]

Although it is hard to quantify the benefit of the removal of terrorist content and activity from the online sphere, its removal will almost certainly influence the level of terrorism in society and some of these costs could be reduced.

Assisting illegal immigration

It is likely that most of the illegal migration is facilitated online or involves the internet at some point in a migrant's journey. The Illegal Migration Bill Impact Assessment estimates the unit cost of relocating an illegal migrant to be £169,000. If the OSA were to reduce the facilitation of illegal migration then this cost would be saved for each migrant who does not enter the UK.

Coercive & controlling behaviour

Coercive or controlling behaviour (CCB) is a form of domestic abuse which captures patterns of abuse that occur over a prolonged period, or cause fear of violence on two or more occasions, which enables an individual to exert power, control or coercion over another[284].

CCB can be perpetrated via technology and the internet. This enables the abuse to be perpetrated both within the home and from a distance and can take place during an intimate relationship but also post-separation. Examples of online CCB include, hacking a current or former partner's phone or computer to obtain intimate images or other private information, obsessive and persistent texting or messaging and the use of mobile technology to check a partner's location in a way that feels controlling[285].

[280] Counter-terrorism strategy (CONTEST) 2023 - Home Office, 2023

[281] Christchurch Call (Ministry of Foreign Affairs and Trade NZ)

[282] 2021 prices. Home Office estimate – see CONTEST 2023 Annex D for further detail.

[283] European Parliamentary Research Service and RAND Europe (2018). The fight against terrorism: Cost of Non-Europe Report. Figures adjusted to 2021 prices in Pounds Sterling by the Home Office – please see CONTEST 2023 Annex D for further detail.

[284] Controlling or coercive behaviour statutory guidance (publishing.service.gov.uk)

[285] Measuring technology-facilitated gender-based violence. A discussion paper (unfpa.org)

There is currently no reliable and accurate measure of the prevalence of CCB or online CCB, although the ONS are testing new questions to measure domestic abuse and CCB in the Crime Survey for England and Wales (CSEW)[286]. The CSEW is considered the most reliable measure for the prevalence of domestic abuse in England and Wales. Estimates from the CSEW for the year ending March 2023 showed that 4.4% of people aged 16 years and over experienced domestic abuse in the last year. There were 44,568 offences of CCB recorded by the police in England and Wales in the year ending March 2023[287]. There were 566 convictions for CCB in the year to December 2022[288].

It is difficult to accurately estimate the proportion of CCB offences that contain an online component. Women's Aid[289] found that 11% of Women's Aid service users in 2021/22 reported surveillance or harassment online or through social media[290].
The Home Office[291] estimates that in the year ending 31 March 2017, domestic abuse is estimated to have cost over £66 billion in England and Wales. Inflated to 2023/24 prices this cost is estimated at £81 billion, or an average of £42,000 per victim using 2017 victim prevalence data. This figure incorporates the harms suffered by domestic abuse victims for the complete period of their abuse and victims' recovery time. CCB can lead to emotional and psychological harms because of manipulation and criticism, this may lead to social isolation, anxiety and depression in victims[292]. It is currently not possible to provide a monetised estimate of the harms of CCB, the above figure provides evidence that domestic abuse is a high harm crime.

Money laundering

Money laundering is a process that allows criminals to use the proceeds generated from illicit activities. It aims to conceal criminal funds and reduce the risk of detection by law enforcement. The UK has a very broad definition of money laundering, set out by the Proceeds of Crime Act (POCA): "A person commits an offence if he— (a) conceals criminal property; (b) disguises criminal property; (c) converts criminal property; (d) transfers criminal property; (e) removes criminal property from England and Wales or from Scotland or from Northern Ireland."[293] Therefore, all criminal proceeds generated in the UK should be presumed to be laundered.

There is no single figure for the scale and costs of money laundering in the UK, or an accurate estimate of the total value of money laundered in the UK each year. The International Monetary Fund (IMF) estimated that 2 to 5% of global GDP is laundered annually, however there is a lack of supporting material and methodology documenting this estimate.[294] Ferwerda, Saase, Unger & Getzner (2020) estimated that 4.3% of the UK GDP was laundered in or through the UK in 2014.[295]

Despite there not being a single accurate estimate, due to the total volume of financial transactions made within the country each year, it is a possibility that it is in the hundreds of billions range. "Accurately assessing the scale of money laundering impacting on the UK remains

[286] Developing a new measure of domestic abuse: April 2023 - Office for National Statistics (ons.gov.uk)

[287] Crime in England and Wales: Appendix tables - Office for National Statistics (ons.gov.uk)

[288] Criminal Justice System statistics quarterly: December 2022 - GOV.UK (www.gov.uk)

[289] The-Domestic-Abuse-Report-2023-The-Annual-Audit-FINAL.pdf (womensaid.org.uk)

[290] From a sub-sample of 31,291 service users within the overall sample of 38,045 service users for whom an abuse profile on current abuse is available. Of these 38,045 service users, 24,943 were seeking support for current abuse and 6,926 were seeking support for historic abuse.

[291] The economic and social costs of domestic abuse (publishing.service.gov.uk)

[292] Controlling or coercive behaviour statutory guidance (publishing.service.gov.uk)

[293] Proceeds of Crime Act (POCA): 2002, section 327

[294] United Nations Office on Drugs and Crime, 2011

[295] Ferwerda, J., Saase, A. v., Unger, B., & Getzner, M. (2020). Estimating Money Laundering Flows with a Gravity Model-based Simulation. Scientific Reports, 10, 18552.

difficult, but it is a realistic possibility that it is in the hundreds of billions of GBP annually" (National Crime Agency, 2021a). Due to the broad definition of money laundering, and the lack of accurate estimates of the total value of money laundered, it isn't possible to give an accurate estimate for the prevalence, unit cost or annual cost of money laundering.

Incitement to and threats of violence

Violence without injury in the Crime Survey for England and Wales captures assault without injury offences, and excludes incitement or threats of violence, but is the closest proxy available. The unit cost of this offence category is £7,704 (inflated to 2023/24 prices) and is mostly comprised of physical and emotional harms and costs to the criminal justice system.[296]

Police recorded crime does not specify incitement or threats of violence, except for threat to kill. There were 51,742 threat to kill offences recorded in year ending December 2023. Applying the violence without injury multiplier from the costs of crime estimates that around 78,000 offences will have taken place in 2023 (including those not recorded by police), with an estimated economic and social cost of £0.6 billion (2023/24 prices). It is worth noting that the estimated number of offences is likely an underestimate of the true figure given the extrapolation from the 'threats to kill' category which is less prevalent than general threats.

The extent to which these offences, and wider incitements and threats of violence, are conducted online is not known. It is therefore not possible to attribute a proportion of these costs to society to online activity. It is likely that online activity, through incitement and threats of violence / threats to kill, does contribute to the total cost of violence without injury. Any reduction in online activity is likely to partly displace incitement / threats made offline or via direct messaging.

Weapons and firearm offences

The closest available proxy for this category of harm is violence with injury. The unit cost of this offence category is £17,550 (inflated to 2023/24 prices), and is mostly comprised of emotional and physical harms, impact to criminal justice system, and loss of output (productivity and time off work).[297] Violence with injury captures assault with injury offences and assault with intention to cause serious harm. This proxy is likely to be an underestimate of the unit cost of some knife and firearm related crime due to the greater level of potential harm from these types of weapons, compared to unarmed violence. On the other hand, the separate offence of use of knives and firearms in threats of violence will likely have a lower unit cost than violence with injury and would be partially captured by the "Incitement to and threats of violence" category.

The unit cost of homicides enabled by knives and firearms can be established using costs of crime. The economic and social cost of a homicide is £4.0 million (inflated to 2023/24 prices).

Latest police recorded figures show that there were 50,849 knife-enabled offences and 6,365 firearm related offences in 2023/24. There were 3,856 hospital admissions for assault with sharp objects in the year ending September 2022. Due to challenges in developing more precise unit costs, it is not possible to estimate the total cost to society associated with these offences.

In year ending March 2022, there were 282 homicides committed using knives and sharp objects, and 28 homicide victims killed by shooting.[298] Applying the unit cost of homicide to these figures provides a total economic and social cost to society of £1.1 billion.

[296] Table E1: Unit costs of crime by cost category, Economic and social costs of crime (Home Office, 2018)

[297] Table E1: Unit costs of crime by cost category, Economic and social costs of crime (Home Office, 2018)

[298] Excludes homicides committed using blunt instruments, including where firearms are used as blunt instruments.

A literature review into how gangs use social media for serious violence found that the evidence base in this area is limited. Few robust studies have been carried out which look at the nature and extent of the problem of online gang activity, and the violence that is triggered as a result.

Children's exposure to pornographic content

319. A significant proportion of children access pornography online both inadvertently and intentionally. 51% of children as young as 11-13 years old have seen pornography, with this rising to 66% and 79% for 14-15 year olds and 16-17 year olds respectively. Many children - some as young as 7 years old - stumble upon pornography online. 61% of the 11-13 year olds who have seen pornography describe their viewing as mostly unintentional. Children's exposure to pornography can result in children feeling 'grossed out', 'confused', 'disturbed' or upset. Many of the children who had seen pornography at such a young age felt that it was unhealthy to have seen such content at that age.

320. Current evidence does not allow robust quantification of the baseline impacts of children's access to pornography. However, there are a range of short and long term impacts that clearly demonstrate the scale of the problem. For example, evidence suggests that pornography can create damaging insecurities in children and young people. 35% of children said they worry about what other people think of their body because they do not look like the actors they see in pornography. 19% of girls, and 17% of the boys said that they had "learnt if I look normal naked" from watching porn and 29% said that pornography makes them feel bad about their body. The content of pornography can also skew young people's view of sex, 30% of boys and girls agree that "real sex hasn't lived up to my expectations from watching porn".

321. Pornography can also influence young people's sexual behaviours and expectations towards more "rough" and "forceful" sexual encounters. Meta-analysis from 2017 shows how those who consume porn frequently are more likely to hold sexually aggressive attitudes and be engaged in sexual behaviour that is conducive to sexual aggression. Some young girls are worried that boys who watch porn will think it is normal to proceed being rough and forceful when a woman's body language indicates that they do not want sex.

322. Some children also feel that porn has affected their or their partner's view of consent, since it is often only implied in porn and not explicitly given. Children who intentionally sought out pornography had the most worrying ideas around consent (by a factor of between three and six in comparison to those who had mostly seen it by accident). 29% of these children did not think consent was needed if "you knew the person really fancies you", in comparison to only 5% of those that had mostly seen pornography by accident.

323. Research also finds that boys who consumed pornography when they were 12-14 years of age, are more likely to have engaged in aggressive, sexual behaviour. In a longitudinal study, 10-15 year olds that consumed violent pornography were six times more likely to be sexually aggressive than those who did not consume it, or than those who consumed less aggressive pornography.

324. While it is not possible to monetise the impact of children's exposure to pornography, it has a clear and significant effect on children's attitudes and behaviours. Both the core child safety duties and pornography provision will ensure platforms protect children from this content. On this basis, Option 1 is expected to result in material reductions in the short and long term impacts of children's access to pornography.

Other non-monetised benefits

325. In addition to a reduction in online harm over the appraisal period, Option 1 is also expected to result in the following non-monetised benefits:

- **benefit to law enforcement**: these benefits are expected to accrue both in terms of a general reduction in online crime and through creating a transparent regulatory system, making it easier for law enforcement to tackle crime online. Requirements to report CSEA content to law enforcement, transparency reporting, and Ofcom's information gathering powers are expected to contribute towards the accrual of these benefits, which are likely to take the form of cost savings or efficiency gains. The level of online crime reduction and the way in which both platforms and law enforcement operate within the online safety framework is unknown at this stage and this IA is therefore unable to monetise this potential benefit.
- **increase in media literacy levels**: many of the steps businesses take to comply with the OSA are likely to result in improvements in media literacy levels. For example, these benefits may accrue from steps taken by platforms to keep users safe, such as warnings and flags, giving users the ability to control the content they see, and other tools related to literacy by design. Furthermore, Ofcom's expanded media literacy duty has the potential to empower users to both keep themselves safe online but also help others to keep themselves safe. In addition, the government's related non-regulatory media literacy interventions are aimed at improving core media literacy skills and giving users the ability to keep themselves safe online. Given that platform actions are unknown at this stage and measurement of media literacy is still evolving, this potential benefit remains non-monetised.
- **safety technology**: Option 1 is expected to result in an increase in demand for safety technology and the government is supporting the sector through a series of non-regulatory interventions, such as research, investment, and challenge funds. Modelling conducted by Perspective Economics[299] estimated that a combination of the incoming online safety regulations and non-regulatory initiatives could create an additional £900 million in revenue and 3,500 FTE jobs in the lead up to the regime. This benefit remains outside of the scope of this IA for two reasons: first, benefits are expected to accrue before the appraisal period for this IA and long-term modelling has not been conducted; and second, results do not distinguish between the impact of the legislation and the impact of non-regulatory initiatives. Any benefit to the Safety Technology sector resulting from the legislation would be considered 'resources used to comply with regulation' as set out in RPC guidance.
- **evidence**: Option 1 is expected to result in an increase in the evidence base underpinning online harm through greater transparency and data availability.

Break even and scenario analysis

326. As outlined in the above sections, this IA quantifies the annual social cost - under baseline - of a subset of online harm, including both illegal and legal but harmful to children. Online harm is assumed to grow at 3% per year and the table below outlines the estimated cost in the first year of the appraisal period.

[299] Internal modelling conducted for DCMS

Table 40: Social cost of online harm (2019 prices, 2020 base year – 10-year PV)

Harm	**Cost to society**
Contact CSEA	£7,630 million
Modern slavery	£267 million
Hate crime	£102 million
Illegal sale of drugs	£3,850 million
Cyberstalking	£218,000 million
Fraud (central)	£18,900 million
Cyberbullying	£5,880 million
Total across the appraisal period (10-year PV)	**£254,000 million**

327. Given the difficulty in providing an evidenced estimate for a percentage reduction in online harm resulting from Option 1, the benefits remain purely illustrative and are not considered in the calculation of the NPSV.

328. The illustrative benefit is the value of a reduction in online harm. Given the data limitations described above, this IA has only been able to quantify estimated benefits for a reduction in the subset of online harm outlined above. It is assumed that, once enacted, a policy will start to reduce online harm in the second year of the appraisal period.[300]

329. As outlined in the previous IA, evidence on the likelihood of benefits occurring remains limited. Similar regulations abroad are either planned and not yet implemented or have not been fully assessed, as is the case for the German NetzDG. Additionally, it is difficult to highlight specific incidences of harm that have occurred in the past but would not have done so under Option 1. This is due to the complex nature of online harm, especially in relation to how they lead to realised impact. For example, hate speech aimed at an individual, impacts both the direct victim but also other users who may see it. The level of harm mitigation achieved from user safety measures will depend on the type of harm and the point at which it is addressed, this makes it difficult to determine the precise likelihood of a reduction in online harm resulting from platforms' responses to Option 1. However, the OSA is expected to lead to a reduction in online harm compared to a *do-nothing* baseline through the following mechanisms (this is not an exhaustive list):

[300] The reduction is relative to the estimates of harm under BAU and is not applied cumulatively. The year in which reductions would start will depend on the year in which regulation is enacted.

Table 41: Qualitative assessment of why the OSA is expected to result in reduced harm

Outcome	Harm reduction
Content moderation	In 2020 for example, Facebook acted on 35.9 million pieces of content relating to child nudity and sexual exploitation of children, around 99% of which was found and flagged before users reported it[301]. This highlights how systems and processes to moderate content can mitigate the impact of online harm. The OSA is expected to lead to some platforms conducting additional content moderation to address online harm. This could be through bolstering existing content moderation processes or implementing new ones for platforms that do not currently moderate content.
User reporting	In 2020, nearly 1.4 million YouTube videos were removed because of user reporting mechanisms on the platform[302][303]. This means that nearly 1.4 million potentially harmful videos (or videos that did not comply with YouTube's community guidelines) were removed from the platform which is likely to have mitigated their impact. Under the OSA, platforms will be expected to accommodate user reporting of harm and therefore, some platforms without these systems will be required to implement them and those that do have them may be required to make improvements.
Age assurance	Both the core child safety duties and pornography provider provision will result in increased age assurance processes online. This will ensure children are protected from age inappropriate material and mitigate the short and long-term impact of harms such as children's exposure to pornographic content. Under the *status quo*, the vast majority of pornography sites (and sites where pornography can be accessed) do not have age assurance systems. Where they do exist, they are light-touch and ineffective, such as a user confirming they are over 18 by ticking a box. Option 1 will ensure that platforms hosting or publishing pornography have effective age verification and age estimation processes in place minimising children's access to pornography online.
Anti-fraud measures	Fraud facilitated both by UGC and advertisements online, lead to significant victim losses. Reporting and content moderation measures alongside increased customer due diligence on advertisers is likely to result in a material reduction in online scams.
Transparency and user behaviour	Category 1, 2A and 2B services will be expected to publish transparency reports under the OSA and Ofcom will have a range of information gathering powers as well as a

[301] Community Standards Enforcement Report - Facebook (2021)

[302] User reporting was the first source of detection

[303] YouTube Community Guidelines Enforcement - YouTube (2020)

Outcome	Harm reduction
	responsibility to conduct research into online harm. Furthermore, the OSA expands on Ofcom's existing statutory duty to promote media literacy, under the 2003 Communications Act. Ofcom will now be required to take steps to increase awareness and understanding of how the public can keep themselves and others safe whilst using regulated services. In addition, alongside the OSA the government is undertaking several projects and initiatives aimed at improving media literacy. The range of initiatives aim to give users more information about the risks and prevalence of online harm on platforms. The government's initiatives related to media literacy are both expected to increase user safety online and mitigate some of the impacts associated with online harm.
Risk assessments	The OSA requires platforms to undertake risk assessments to assess risks corresponding to the type of content and activity a business is required to address. Many platforms already conduct risk assessments; however, there will be some that do not and these assessments could result in more or better targeted content moderation leading to a more efficient allocation of resources and greater harm mitigation.

330. Given the uncertainty around the reduction in online harm that could be achieved under Option 1 (as described above), this IA estimates the reduction in the subset of quantified online harm required to exactly match the costs, that is, the scale of the reduction of harm required to deliver a benefit-cost ratio of precisely 1. The results are shown in the table below.

Table 42: Break-even point[304]

	Low	Central	High
Option 1	0.9%	1.3%	1.7%

331. To further inform the analysis, this section considers how the benefit-cost ratio would change if different illustrative assumptions were made about the effectiveness of Option 1 in reducing harm:

- ○ low reduction scenario = 1% per year compared to a *do nothing* counterfactual
- ○ mid reduction scenario = 3% per year compared to a *do nothing* counterfactual
- ○ high reduction scenario = 5% per year compared to a *do nothing* counterfactual

332. Based on these scenarios, the table below compares the costs and benefits.

[304] The difference from the break-even point with previous IAs is driven by revised valuations of harms, particularly from cyberstalking, supplied by the Home Office.

Table 43: Benefit cost ratios (BCR) under illustrative scenario (central estimate only)

Low reduction scenario					
Low estimate		**Central estimate**		**High estimate**	
Implied BCR	1.1	Implied BCR	0.8	Implied BCR	0.6
Mid reduction scenario					
Low estimate		**Central estimate**		**High estimate**	
Implied BCR	3.2	Implied BCR	2.3	Implied BCR	1.8
High reduction scenario					
Low estimate		**Central estimate**		**High estimate**	
Implied BCR	5.3	Implied BCR	3.8	Implied BCR	3.0

Indirect costs and benefits

Freedom of expression

333. This legislation is designed to protect freedom of expression and has built in safeguards to avoid any potential negative impacts. This section sets out how the proposals are designed to enhance freedom of expression online rather than limit it.

334. Under the *status quo*, major technology companies already exercise significant power over what lawful speech is considered acceptable online. Many users complain about the opaque, arbitrary removal of their legitimate content and the lack of clear routes to appeal the takedown. Decisions on how to moderate content involve trade-offs with freedom of expression and absent regulation these decisions are being made by companies without democratic oversight. The requirements on Category 1 platforms to ensure they have clear and accessible terms of service and user redress mechanisms are expected to minimise freedom of expression impacts currently inherent under the *status quo.* These platforms will not be able to arbitrarily remove content. They will need to be clear what content is acceptable on their services and enforce the rules consistently and users will have access to effective mechanisms to appeal content that is removed without good reason. They will also be required to have regard for freedom of expression when fulfilling their safety duties.

335. In addition, some individuals and groups do not engage online through fear of being the targets of online abuse. For example, an international survey of female journalists found 64% had experienced online abuse – death or rape threats, sexist comments, cyberstalking, account impersonation, and obscene messages.[305] Almost half (47%) did not report the abuse they had received, and two fifths (38%) said they had self-censored in the face of this abuse. Additionally, in

[305] IFJ global survey shows massive impact of online abuse on women journalists - IFJ (2018)

the 2017 Annual Bullying Survey,[306] of those that had been the victims of cyberbullying, 26% deleted their social media profiles and 24% stopped using social media altogether.[307] The framework takes an approach which benefits and protects all users. It will empower adults, including vulnerable users, to keep themselves safe online, and to enjoy their right to freedom of expression, reducing the risk of bullying or being attacked based on their identity.

336. Given that the OSA is likely to result in increased (or more effective) content moderation, several stakeholders have raised concerns relating to potential negative impacts on freedom of expression. The table below sets out some of the main concerns:

Table 44: Main stakeholder concerns around potential impacts on freedom of expression

The OSA forces platforms to delete legal content which will have a negative impact on freedom of expression	Since the final stage IA, Category 1 services will no longer be required to risk assess for, and set and enforce, terms of service for certain categories of legal content. Instead, they will be required to carry out an assessment of the incidence of certain kinds of content on their services, and offer users user-empowerment tools for these kinds of content where relevant. They will also be required to have systems and processes to ensure they enforce their own terms of services consistently and transparently, and only remove or restrict access to content, or ban or suspend users in accordance with these terms of service. This duty will prevent those services from arbitrarily removing or restricting legal content or suspending or banning users except where this is in accordance with the service's express terms of service. The OSA also contains protections for freedom of expression that require platforms to consider the importance of free expression when fulfilling their safety duties under the OSA. Similar protections apply to content of democratic importance and journalistic content on Category 1 services.
The OSA's definitions of harmful content are too vague and could result in the over removal of content.	The OSA requires platforms to act against illegal content on their service where it is an existing UK offence that gives rise to harm to an individual. To clarify the 'illegal content duties', provisions have been added establishing how providers should determine whether content amounts to illegal content. This will provide greater clarity about how service providers should make judgements about content on their service, including whether it amounts to illegal content and must be removed. This, alongside freedom of expression provisions in the OSA, will safeguard against the over-removal of content. The OSA requires platforms that are likely to be accessed by children to protect children on their service. Companies will need to take action to protect children against content that poses a material risk of it having - or indirectly having - a significant adverse physical or psychological impact on a child of ordinary sensibilities. Action may include restricting children's access to that content (rather than removing such content entirely). The OSA also requires Category 1 services to set out their policies in relation to content or activity that is prohibited on their service, and to only remove this where it goes against their terms of service.

[306] The 2017 survey is used here as it included a deep dive on cyberbullying specifically. It is not possible to disaggregate the impacts of traditional bullying and cyberbullying in more recent editions.
[307] Annual Bullying Survey 2017 (Ditch the Label, 2017)

Large fines will cause platforms to overreact and remove content that is legal.	Platforms are required to remove ‘illegal content’ under the OSA, and they will have concomitant duties to take freedom of expression into account when carrying out their safety duties. Ofcom enforcement will apply equally to all duties in the OSA, including those regarding freedom of expression, such that the OSA ensures against platforms ‘overreacting’. Ofcom has the option of imposing substantial fines to encourage compliance (and to reflect instances of serious user harm). However, the cap is a ceiling. Ofcom will only impose fines proportionate and appropriate to the breach that has occurred. Escalating enforcement sanctions will avoid incentivising content takedown, with judicial oversight required for the most severe sanctions.
The protections for news publisher content, journalistic content and content of democratic importance provide some people with a higher level of protection, creating a two-tier system online.	The protections for journalistic content and content of democratic importance focus on the content, not the actor. Anyone who posts this content will benefit from the protections. The protections themselves are important to ensure democratic debate is protected online and users have access to quality journalism. The protections for news publishers apply to content produced by organisations that meet the definition of a ‘recognised news publisher’ (“RNP”) as set out in the OSA. Clause 50 of the OSA sets out a range of criteria that an organisation must meet to qualify as a news publisher. These include that organisations have publication of news as their principal purpose; are subject to a standards code; and that their content is created by different persons. The government is committed to protecting media freedom and the invaluable role of a free press in our society and democracy, and the criteria for news publishers were created with this in mind. However, we are clear that bad actors should not benefit from the protections and that is why we have taken steps such as ensuring that sanctioned news outlets such as RT must not benefit from these protections.
The OSA provides Ofcom with too much power and allows it to regulate free speech.	Ofcom is accountable to Parliament in how it exercises its functions. It is required to present its annual report and accounts before both houses and to appear before Select Committees to answer questions about its regulatory operations. Parliament will have a role in approving several aspects of the regulatory framework through its scrutiny of both the primary and secondary legislation. The government has ensured that, in addition to judicial review through the High Court, there is an accessible and affordable alternative means of appealing the regulator's decisions. The OSA will establish the Upper Tribunal as the alternative route to appeal Ofcom’s decisions. As a public body, Ofcom is bound by the European Convention on Human Rights, including the Article 10 right to freedom of expression, under the Human Rights Act 1998 (HRA). It has an obligation not to act in a way which is incompatible with the right to freedom of expression when carrying out its duties, for which it can be held to account. This means that Ofcom will not be able to put in place any measures that restrict users’ freedom of expression unless it is lawful, necessary, and proportionate to do so.

Removing the right of an individual to remain anonymous online will limit freedom of expression	Option 1 does not remove the right of an individual to remain anonymous online. The government agrees that placing restrictions on anonymity online could disproportionately impact users without official ID (such as refugees, migrants and those from lower socio-economic backgrounds), or those who are reliant on ID from family members, and would experience a serious restriction of their online experience, freedom of expression and rights. The OSA requires platforms to provide optional user verification and allow users to determine the content and kinds of users they interact with online. Under the user verification duty, users are still able to be completely or pseudo anonymous online, verifying their identity only if they wish to do so.

337. While Option 1 is expected to enhance certain aspects of freedom of expression online it also includes several protections - both in the design and specific safeguards - to ensure any negative impacts are mitigated. In its comparative analysis of online harm regulations in eight jurisdictions,[308] Linklaters identified that regimes can broadly be divided into those that focus on individual pieces of content and those that instead focus on the 'systems and processes' that platforms must have in place. The online safety framework is a 'system and processes' approach which means that providers will not be punished for a failure to remove individual items of content within a certain time period - rather for a failure to put in adequately performing systems and processes (e.g. content moderation processes) to safeguard their users . For example, Germany's NetzDG requires platforms to remove illegal content within 24 hours. This approach was copied in France's "Avia Law" (see international context section) but was deemed by the French Constitutional Court to be incompatible with the right to freedom of expression, given the risk that platforms would "over-block" to avoid enforcement action. By focusing on the aggregate performance of providers' systems and processes (rather than on liability for individual items of content), there will be less of an incentive for platforms to take too cautious an approach and immediately take down any content with a mere hint of illegality to avoid sanction, restricting freedom of expression online as a consequence.

338. Finally, Option 1 includes several built-in safeguards to protect freedom of expression, these include:

- ○ all in-scope companies must have regard to the importance of protecting freedom of expression when implementing safety policies and procedures. This mitigates the risk that companies adopt highly restrictive measures to fulfil their statutory duties.
- ○ codes of practice will set out steps relating to companies' processes for considering the balance between user safety and freedom of expression when introducing content moderation or other online safety measures. Companies will be assessed as having fulfilled their duty to have regard to the importance of protecting freedom of expression if they follow these steps.
- ○ companies must have systems and processes in place to enable users to complain and seek redress if their content has been unfairly removed or restricted, or if they have been suspended or banned from a service.
- ○ effective transparency reporting will help ensure content removal is well-founded, as the decisions platforms make on content removal and user appeals on content removal will have greater visibility.

[308] Online harms a comparative analysis (Linklaters)

- escalating enforcement sanctions will avoid incentivising content takedown, with judicial oversight to safeguard the most severe sanctions like access restriction.
- super-complaints will allow organisations to lodge concerns on behalf of users, which can include concerns about limits on freedom of expression.
- companies must make clear in their terms of service that users have a right to bring a claim in court for breach of contract where their content is removed in breach of that company's terms of service.
- Category 1 services will have new duties to implement systems and processes to ensure they only remove or restrict access to content, or ban or suspend users, except where in accordance with their terms of service, or where they otherwise have a legal obligation to do so. This will prevent those services from arbitrarily removing or restricting legal content, however controversial, or suspending or banning users where this is not in accordance with the service's express terms of service. Platforms must also ensure these terms of service are clear, easy to understand and consistently enforced.
- Category 1 services will need to assess the impact on freedom of expression and privacy both when deciding on safety policies and after they have adopted those policies. They will also need to demonstrate they have taken positive steps to mitigate this impact.
- Category 1 services are required to put in place clear policies about how they will protect users' access to content of democratic importance[309] when making content moderation decisions. Providers must consider the importance of users' free expression in relation to content of this kind.
- content of democratic importance will apply to content, not people. Therefore, content that supports or opposes government policy will be captured whether the creator of that content is a government minister or an individual political campaigner. This definition of democratic content does not, therefore, privilege politicians and/or specific political parties. For example, a service cannot provide a higher level of protection for left-wing views compared to right-wing ones.
- users will be able to appeal to the platforms if they consider that the platform is not complying with its duties to protect content of democratic importance.
- Category 1 services will be required to put in place clear policies to protect journalistic content[310] and recognised news publishers' content when making content moderation decisions. Protections must include an expedited complaints procedure for users who are the creators of such content (including recognised news publishers) to appeal against decisions companies have taken regarding journalistic content they have generated, shared or created**.**
- Ofcom must fulfil its new functions in a way that protects users' rights to freedom of expression. There will be a robust appeals process against regulator decisions for anyone materially affected by a decision by the regulator.

339. The online safety framework limits platforms' ability to arbitrarily remove lawful content, and is designed to protect freedom of expression online. Based on the above qualitative assessment of freedom of expression implications, Option 1 is expected to enhance freedom of expression online rather than limit it.

[309] 'Content of democratic importance' is defined as content, including news publisher content, which is, or appears to be, intended to contribute to democratic political debate in the UK at a national or local level. This includes content promoting or opposing government policy and content promoting or opposing a political party.

[310] 'Journalistic content' will apply to content, including news publisher content, which is generated for the purpose of journalism and which is UK-linked

Privacy impacts

340. There are several areas within the OSA that have the potential to result in privacy implications. For this reason, it includes strong privacy protections and Ofcom and the ICO will work together to ensure consideration of how personal data is processed as part of the duties.

341. The regulatory framework will apply to public communication channels and services where users expect a greater degree of privacy - for example online instant messaging services and closed social media groups. The regulator will set out how businesses can fulfil their duties in codes of practice, including what measures are likely to be appropriate in the context of private communications. This could include steps to make services safer by design, such as limiting the ability for anonymous adults to contact children.

342. End-to-end encrypted services are in scope of the OSA and Ofcom will take steps to ensure that these services are meeting their obligations under the duties. The government is supportive of strong encryption to protect user privacy, however, there are concerns that a move to end-to-end encrypted systems, when public safety issues are not considered, is eroding several existing online safety methodologies. This could have significant consequences for tech companies' ability to tackle grooming, sharing of CSEA material, and other harmful or illegal behaviours on their platforms. Companies will need to regularly assess the risk of harm on their services, including the risks around end-to-end encryption. They would also need to assess the risks ahead of any significant design changes such as a move to end-to-end encryption. Service providers will then need to take reasonably practicable steps to mitigate the risks they identify.

343. In addition, given the severity of the threat, the legislation will also enable Ofcom to require businesses to use technology that is highly accurate to identify, take down, and prevent and remove tightly defined categories of illegal material relating to CSEA and terrorism on public and, where proportionate, CSEA content on private communications. The regulator will also have the power to require companies to use their best endeavours to develop or source new technology to tackle CSEA, which will ensure they will have the flexibility to find the best fit method of tackling CSEA on their service.

Age assurance requirements in the OSA have the potential to require the use of users' personal data - depending on the specific solution used. Under Option 1, platforms that host pornographic content will likely be required to verify the age of their users to prevent children from accessing this content. Concerns related to user privacy were raised under Part 3 of the Digital Economy Act 2017; however, the OSA, combined with existing data protection law, will provide strong legal safeguards for user privacy. The Data Protection Act 2018 already provides a high standard of data protection legislation in the UK, which age verification providers will need to comply with and which has strong sanctions for malpractice. The Information Commissioner's Office recently published an opinion about the use of age assurance technologies and compliance with data protection law, which makes clear that providers using age verification must comply with data protection principles of transparency, fairness, lawfulness, accuracy, data minimisation and purpose limitation. The ICO also suggests companies use appropriately certified solutions. The OSA will also place an explicit duty on providers to carry out privacy impact assessments. In addition, the ICO recently approved a new certification scheme for age assurance, through which companies can demonstrate their commitment to following the DPA 2018 when using age assurance technologies. Furthermore, there is a growing range of solutions available on the market that minimise the amount of personal data users are required to share and which can provide platforms with an anonymised 'yes/no' answer to whether the user is over 18. M

344. More broadly, all in-scope companies must have regard to the importance of protecting users from unwarranted infringements of privacy when implementing safety policies and procedures. Codes of practice will set out steps relating to companies' processes for considering the

balance between user safety and privacy when introducing content moderation or other online safety measures. Companies will be assessed as having fulfilled their duty to have regard to the importance of protecting users from unwarranted infringements of privacy if they follow these steps.

345. The then government consulted a range of stakeholders on end-to-end encryption and privacy implications more generally. This included businesses, Parliament, charities, and privacy-focussed organisations. Proposals have included banning end-to-end encryption or greater consequences for companies when illegal material such as CSEA is found on their systems. However, there are also several privacy-focussed organisations who are concerned about how the regulatory framework will impact on user privacy.

346. Recognising the potential risk of an impact to users' privacy, the preferred option includes several protections for privacy and mitigations against potential privacy implications.

Table 45: Overview of mitigations against privacy impacts

Mitigation	Description
Platforms must take steps to protect against unwarranted infringements of privacy when carrying out their safety duties.	The OSA includes specific provisions that require service providers to protect against unwarranted infringements of privacy in the fulfilment of their safety duties and reporting and redress duties. This is to ensure service providers do not, for example: ○ actively monitor more content than is necessary for safety features to function ○ track the activity of children more than it is needed to ensure they are only served appropriate content
Platforms must take steps to enable users and other affected persons to report concerns about a platform's non-compliance with their duties.	This includes a platform protecting against unwarranted infringements of privacy. If a complaint is upheld, platforms are expected to seek to rectify the issue by making changes to their policies and procedures to bring themselves into compliance.
Ofcom must put together codes of practice that explain how platforms can comply with their duties. Ofcom must consult on these codes. Platforms must comply with these codes or take alternative steps that achieve the same ends.	Companies will be expected to be clear about how they can protect against unwarranted infringements of privacy when fulfilling their duties. Throughout the codes, Ofcom would set out how platforms can fulfil each of their safety and redress duties in such a way that protects users from unwarranted infringements of privacy. For example, Ofcom may refer to service providers' existing duties under data protection law and include specific steps that service providers can take to guard against privacy infringements when implementing safety systems and processes.
Stringent safeguards relating to Ofcom requiring the use of technology	This will only be used as a last resort where alternative measures are not working and will be subject to stringent safeguards to protect users' rights. The regulator will advise the government on the accuracy of tools and make operational decisions regarding whether or not a specific business should be required to use them. Before the regulator can use these powers, it will need to seek approval from ministers on the basis that sufficiently accurate tools exist.

Mitigation	**Description**
Ofcom can enforce the privacy duties on platforms.	Ofcom will be able to enforce the privacy duties to hold platforms to account.
Collaboration between Ofcom and the ICO	Ofcom will work closely with the ICO when developing codes so that platforms are clear about what they must do to comply with both regimes and inefficiencies are reduced. Each regulator will provide guidance for platforms and users about how the regimes interact. Operationally, both regulators will work closely together to resolve issues as they arise, for example, flagging complaints that are relevant to the other regulator and passing on complaints that are for the other regulator to investigate. Ofcom will be required to consult the ICO before producing or updating guidance on how they will exercise their enforcement powers in relation to the OSA.
In addition to the measures within the OSA, the government is supporting the development of technological solutions to mitigate against the public safety challenges arising from the use of end-to-end encryption.	As part of the government's related non-regulatory interventions, the Safety Tech Challenge Fund awarded five organisations funding to prototype and evaluate innovative ways in which sexually explicit images or videos of children can be detected and addressed within end-to-end encrypted environments, while ensuring user privacy is respected.
The Secretary of State must review how effective the regulatory framework is at protecting users from unwarranted infringements of privacy.	Given the novelty and complexity of the regime, monitoring work and the post-implementation review will consider freedom of expression and privacy implications

347. There are inevitably trade-offs between user safety and technologies such as end-to-end encryption which seek to increase user privacy. Option 1 recognises this and includes strong protections for user privacy online. Alongside Ofcom, the government will continue to consult with stakeholders through implementation of the regime and beyond to ensure any potential privacy implications are minimised.

Impact on user experiences

348. While the OSA is expected to improve user experiences overall, the measures implemented by in-scope services in response to the requirements set out in the OSA have the potential to introduce additional frictions into the user experience, which could have adverse impacts on both users and platforms.

349.

Table 46: Stakeholder concerns around potential impacts on user experience

Additional frictions introduced by age assurance could deter users from using their service and result in a subsequent reduction in revenue.	Ofcom research suggests that age verification is generally accepted by adult users for certain online activities such as preventing under-18s from accessing online pornography.[311] Some solutions will be more privacy preserving than others and therefore more palatable for the user and less likely to deter them, for example, some solutions will offer anonymised tokenised exchange between the third party age assurance provider and the platform being accessed, so no personally identifiable information is shared with the platform. The level of user friction introduced by these measures is also likely to depend on the amount of interoperability in the market. As previously mentioned, there are significant movements towards interoperability with solutions that can work across several platforms. While this is not yet an established approach, it is something that the government is supporting through its work on standards, including the Digital Identity and Attributes Trust Framework, which will support interoperable solutions to function.
The requirement for Category 1 services to proactively acquire confirmation from registered users on whether they wish to use the user empowerment features has the potential to introduce greater friction into the user journey. It may require a company to acquire confirmation every time a user accesses the service.	This requirement is driven by a desire to ensure these tools are easy for users to opt in or out of. The government has sought to mitigate this impact by requiring providers to proactively ask only their registered users whether they wish to use the features. As such, providers can choose to store the preference given by a registered user and avoid repeatedly asking them to confirm their preference.
User verification, which allows users to filter out content from non-verified users and prevent non-verified users from interacting with their content, could have negative effects through reducing users' exposure to different ideas or viewpoints.	The non-verified users' duties only apply to Category 1 services and are entirely optional for users. If a user found that these tools impacted their exposure to different content, they could turn them off. In addition, Ofcom will produce codes of practice for the non-verified users' duties, at which point they will consult with services about how the duty might interact with various functionalities. It is important that users can be better protected from online abuse, particularly from anonymous accounts, which these duties will help to do.

350. Organisations report greater synergies between safety measures and engaging user experiences if considerations relating to user safety have been factored in throughout the design process through a 'safety by design' approach.[312] Therefore, a 'safety by design' approach is integral to compliance with the OSA. The then government published 'Principles of safer online platform design' guidance in Summer 2021, which sets out best practice platform design for user

[311] Adult Users' Attitudes to Age Verification on Adult Sites (Ofcom, 2022)

[312] Trust, Safety and the Digital Economy - The Commercial Value of Healthy Online Communities (Ipsos, 2022)

safety. This will support companies to build safer online services from the outset and platforms while minimising the effects on user experiences.

Calculations

351. Under requirements set out in the Better Regulation Framework, this IA calculates an illustrative overarching EANDCB covering the whole policy, including best estimates for requirements resulting from future codes of practice. The illustrative EANDCB includes all monetised direct costs to business. Under Option 1, the NPSV is estimated to be -£2,970 million (central) with an EANDCB of £263 million (central). This NPSV and EANDCB are illustrative only and are based on a best estimate of likely business requirements stemming from future codes of practice. It will be for Ofcom to determine specific requirements and is required in the legislation to conduct consultations and produce IAs.

352. Estimated costs have increased since the consultation stage IA as follows –

	NPSV	EANDCB
Consultation stage	-£2,120 million	£206 million
Final stage	-£2,510 million	£251 million
Enactment stage	-£2,970 million	£263 million

The main factors include:

- **content moderation costs**: while there has been no change to the methodology or analytical assumptions here, increases in the number of businesses in scope (to reflect an implementation date of 2024) and inclusion of the latest revenue data has led to increases in this cost.
- **additional costs**: the inclusion of the pornography provision, the fraudulent advertising duty, and duties related to user verification and empowerment have added to the NPSV by approximately £190 million.
- **consultation and new evidence**: familiarisation costs and transition costs have also been increased to reflect input from stakeholders in areas such as the potential need for legal advice and representing SMB staff time with Chief Executive wage estimates instead of estimates for regulatory professionals. These costs have also increased to reflect published guidance.

353. Given that specific business requirements are unknown at this stage, the EANDCB calculated here remains largely illustrative and aims to indicate the potential scale or nature of impacts of the whole policy (scenario 2 in the RPC's primary legislation guidance).[313]

Key assumption sensitivity analysis

354. This IA presents low, central, and high estimates throughout to reflect the range of potential impact scenarios on business. Additionally, this section brings attention to the key assumptions used in the production of the estimates and varies them in isolation to outline how sensitive the central estimate is to each.

[313] RPC Case Histories: assessment and scoring of primary legislation measures (2019)

Table 47: Sensitivity analysis

<table>
<tr><th>Assumption</th><th>Lower bound sensitivity</th><th>Central</th><th>Upper bound sensitivity</th></tr>
<tr><td>Number of businesses in scope of the regulation</td><td>19,400</td><td>25,100</td><td>139,000</td></tr>
<tr><td>Illustrative EANDCB</td><td>£210m</td><td>£263m</td><td>£1,380m</td></tr>
<tr><td>Illustrative NPSV</td><td>-£2,350m</td><td>-£2,970m</td><td>-£12,400m</td></tr>
<tr><td colspan="4">Evidence: The central estimate is based on a random stratified sample of 500 businesses from the IDBR. This is varied to reflect the number of organisations identified by RR prior to supplementing with additional known types of organisations likely to be in scope and the upper bound of RR estimates.</td></tr>
<tr><th>Assumption</th><th>Lower bound sensitivity</th><th>Central</th><th>Upper bound sensitivity</th></tr>
<tr><td>Cost to Category 1 organisations of additional content moderation</td><td>1% of turnover</td><td>7.5% of turnover</td><td>15% of turnover</td></tr>
<tr><td>Illustrative EANDCB</td><td>£194m</td><td>£263m</td><td>£341m</td></tr>
<tr><td>Illustrative NPSV</td><td>-£2,210m</td><td>-£2,970m</td><td>-£3,480m</td></tr>
<tr><td colspan="4">Evidence: The central estimate is based on the midpoint of estimates provided by in-scope businesses during the interview phase of RR research project. This is varied to reflect the range of responses.</td></tr>
<tr><th>Assumption</th><th>Lower bound sensitivity</th><th>Central</th><th>Upper bound sensitivity</th></tr>
<tr><td>Cost to Category 2 organisations of additional content moderation</td><td>0.3% of turnover</td><td>1.8% of turnover</td><td>3.8% of turnover</td></tr>
<tr><td>Illustrative EANDCB</td><td>£143m</td><td>£263m</td><td>£429m</td></tr>
<tr><td>Illustrative NPSV</td><td>-£1,770m</td><td>-£2,970m</td><td>-£4,230m</td></tr>
<tr><td colspan="4">Evidence: The central estimate of 1.9% used above is 25% of the midpoint of estimates provided by businesses in interviews (7.5% of turnover). This reflects the proxied volume of illegal vs harmful content actioned by social media businesses (25% illegal content). This is varied to reflect the range of responses.</td></tr>
<tr><th>Assumption</th><th>Low</th><th>Central</th><th>High</th></tr>
<tr><td>Growth rate of online harms under the baseline</td><td>1.3%</td><td>3.0%</td><td>5.1%</td></tr>
<tr><td>Break-even point</td><td>1.4%</td><td>1.3%</td><td>1.1%</td></tr>
<tr><td colspan="4">Evidence: The central estimate is in line with growth in the number of hours spent online. This is varied to reflect a realistic range in terms of potential growth rates. This has a small effect on outcomes</td></tr>
</table>

<table>
<tr><td colspan="4">because later years with much higher levels of harm are heavily discounted.</td></tr>
<tr><th>Assumption</th><th>Low</th><th>Central</th><th>High</th></tr>
<tr><td>Percentage of fraud within scope of the OSA</td><td>30%</td><td>45%</td><td>60%</td></tr>
<tr><td>Break-even point</td><td>1.3%</td><td>1.3%</td><td>1.3%</td></tr>
<tr><td colspan="4">Evidence: This reflects an illustrative range of between 30%-60% (central: 45%) of the relevant fraud offences likely to be in scope of the OSA. However, fraud is a relatively small (7%) driver of overall harms that are expected to be reduced.</td></tr>
</table>

Impact on small, micro, and medium-sized businesses

Small and micro business assessment

Justification for non-exemption

355. As explained in guidance from the RPC, the default position is to exempt SMBs fully from the requirements of new regulatory measures.[314] However, the evidence suggests that the objectives of the regulations would be compromised by exempting SMBs.

356. First, there is evidence of harm occurring on smaller platforms. Law enforcement and NGOs regularly see CSEA offenders active on small chat forums, live streaming apps and file sharing/hosting services. The IWF notes that online harm exists 'in vast quantities' on smaller platforms.[315] 87% of the content the IWF removes from the internet is from small and medium size sites including file sharing sites, image hosting boards and cyberlockers.

357. In addition, terrorist actors have sought to 'exploit an overlapping ecosystem of services', taking advantage of the fact that smaller businesses 'don't have the scale or resources to handle the challenge on their own'. The Tech against Terrorism project indicated that Daesh supporters use larger, well-known platforms (e.g. X) to share links to smaller, less well-resourced platforms, where it is easier to exchange terrorist content.[316] Second, there is a limited relationship between the size of an organisation in terms of turnover and employees and the reach and impact of a given organisation. Third, given the fluidity of the online space, it would be possible for individuals to migrate from large to small platforms in a short time frame.

Impacts on SMBs

358. This IA estimates that there are around 21,500 SMBs within scope of the OSA. The in-scope SMBs are estimated to fall within the following risk categories:

[314] Small and Micro Business Assessments: guidance for departments, with case history examples - RPC (2019).

[315] IWF Online Harms White Paper Response (2021)

[316] UK launch of tech against terrorism at Chatham House - Tech Against Terrorism (2017).

Table 48: Estimated number of SMBs and medium businesses in each risk tier (rounded to the nearest ten)

	Low risk	Mid risk	High risk	Category 1
Micro	10,090	10,090	60	0
Small	580	580	60	0
Medium	1,210	1,210	490	0

359. Tables 49 to 52 outline the indicative estimated average costs that SMBs are expected to incur because of the regulations (with medium and large businesses included for comparison). Final outturn could be different due to business size, structure and platform functionality:

Table 49: Indicative estimated average SMB transition costs excluding additional user reporting costs (2019 prices)[317]

	Low risk	Mid risk	High risk
Micro	£3,720	£3,740	£3,900
Small	£7,200	£7,250	£7,460
Medium	£7,000	£7,040	£7,300
Large	£18,300	£18,400	£18,700

This table represents the per business transition costs. It does not reflect costs to the 10% of businesses in the central estimate that are expected to incur higher costs because of not currently enabling user reporting.

Table 50: Indicative estimated average SMB transition costs including additional user reporting costs (2019 prices)

	Low risk	Mid risk	High risk
Micro	£4,620	£4,650	£4,810
Small	£8,100	£8,160	£8,370
Medium	£7,890	£7,940	£8,200
Large	£19,300	£19,300	£19,600

This table represents the per business transition costs. It reflects the cost to 10% of businesses in the central estimate that are expected to incur higher costs because of not currently enabling user reporting.

360. As the tables above illustrate, the largest per-business transition costs are expected to fall on large businesses who are better placed to absorb them. High transition costs are also driven by a conservative assumption, consistent with earlier versions of the IA, that chief executives in small businesses will be responsible for familiarisation and the illegal content judgment, which were lower in previous IAs. In reality, businesses with 10-49 employees are likely to have a range of options to defray this cost. While costs are expected to be higher for medium and large businesses in absolute terms, small and micro businesses that do not currently allow users to report harm are expected to incur comparable costs when considered in relative terms. Allowing users to report harm is

[317] During their scrutiny of this IA, RPC flagged the significant increase in indicative transition costs for small and micro businesses, compared to the OSB final-stage IA. This increase derives from a more accurate estimate of the cost of all platforms familiarising themselves with Ofcom's guidance, as well as the new illegal content judgement. Both these duties apply to all platforms.

fundamental to the success of the OSA and to keeping users safe online and therefore, the government considers these costs to be proportionate.

361. Both the government and Ofcom will work with small and micro businesses through implementation to ensure transition costs are minimised through for example, clear and accessible codes and guidance and proportionate expectations based on the size of business and risk of harm.

Table 51: Indicative estimated average SMB compliance costs excluding additional content moderation and risk assessment costs (2019 prices)

	Low risk	Mid risk	High risk	Category 1
Micro	£87	£87	£87	
Small	£87	£87	£87	
Medium	£300	£300	£300	
Large	£599	£599	£599	£390,000

This table represents the per business compliance costs. It does not reflect the cost to the 2.5% of businesses in the central estimate that are expected to incur higher costs because of not currently assessing risks. It also does not reflect the 10% of larger mid-risk firms and the 25% of high-risk firms expected to conduct additional content moderation.

Table 52: Indicative estimated average SMB compliance costs including additional content moderation and risk assessment costs (2019 prices)

	Low risk	Mid risk	High risk	Category 1
Micro	£6,040	£6,040	£8,500	
Small	£6,040	£6,040	£50,800	
Medium	£6,250	£321,000	£321,000	
Large	£6,550	£4,240,000	£4,240,000	£18,400,000

This table represents the per business compliance costs. It reflects both the cost to the 2.5% of businesses in the central estimate that are expected to incur higher costs because of not currently assessing risks and the 10% of larger mid-risk firms and the 25% of high-risk firms expected to conduct additional content moderation.

362. While per business costs are expected to be higher for medium and large businesses, it is important to consider the possibility that some in-scope SMBs will have limited resources for compliance. To minimise burdens on SMBs, it will be vital for Ofcom to work with businesses and to ensure both requirements and enforcement are proportionate to the risk of harm and resources available to businesses. Proportionality in the context of effective safety measures must be balanced against the risk of harmful content being displaced to smaller and less well-equipped platforms. The government and Ofcom will work with SMBs to ensure that steps taken are effective in both reducing harms and minimising compliance costs. The government's Safety by Design framework and guidance is targeted at SMBs to help them design in user-safety to their online services and products from the start thereby minimising compliance costs.

363. The pornography provision is estimated to bring into scope an additional 12 SMBs (11 micro businesses and one small business),[318] made up of high risk UK-based pornography providers. These businesses will only incur costs associated with preventing children from accessing

[318] Based on business demographics within creative industries - DCMS Sectors Economic Estimates 2022: Business Demographics (DCMS)

pornography. This impact assessment was only able to estimate illustrative site costs and a total economic cost of the pornography provision, and it is not possible to determine the per business cost on these 12 SMBs. Ofcom - through its assessment of codes and regulator guidance - will further consider potential impacts on these businesses.

Unregulated SMBs

364. While the fraud advertising duty only applies to Category 1 and 2A platforms (costs reflected in Table 21 above) it is also expected to result in costs to a significant number of out-of-scope SMBs that advertise on in-scope platforms. Costs here are expected to occur because of providing information to support anti-fraud checks. This impact assessment estimates that approximately 3.1 million micro businesses and 0.2 million small businesses will incur some costs in the first year. The two tables below outline the per business cost to SMBs expected to undergo standard and enhanced CDD which is made up of staff time to provide necessary information:

Table 53: Fraudulent advertising duty per business costs (standard CDD)

	Low risk	**Mid risk**	**High risk**
Micro	£7.60	£15.20	£22.80
Small	£3.20	£6.30	£9.50
Medium	£3.20	£6.30	£9.50

Table 54: Fraudulent advertising duty per business costs (enhanced CDD)

	Low risk	**Mid risk**	**High risk**
Micro	£22.80	£34.30	£45.70
Small	£9.50	£14.30	£19.00
Medium	£9.50	£14.30	£19.00

365. It should be noted that while a significant number of SMBs are expected to undergo CDD, these costs are one-off and once a business is verified to advertise on Category 1 and 2A platforms, they are not expected to incur any additional costs in the appraisal period. 95% of these businesses are expected to undergo standard CDD with 5% incurring costs associated with enhanced CDD.

366. Given the proportionate and risk-based design of the regulations, the vast majority of costs fall on medium and large businesses. Based on the cost distribution across size bands in the table below (and the per business cost in the table above), costs are not expected to fall disproportionately on SMBs.

Table 55: Total costs for each size band

	Total costs (10-year PV)	Number of businesses (to nearest ten)	Percentage of total costs	Percentage of in-scope businesses
Micro	£111.0 million	20,300	5.3%	80.8%
Small	£15.3 million	1,220	0.7%	4.9%
Medium	£585.0 million	2,910	28.0%	11.6%
Large	£1,380.0 million	680	65.9%	2.7%

Please note: These costs do not include the industry fee, as well as certain Category 1, 2A and 2B-specific costs as it is not clear which businesses are likely to contribute or be Categorised; however, given the revenue threshold aspect of the fee, the majority are expected to fall on medium and large businesses.

Findings from SMB engagement

367. The then government engaged extensively with industry including with SMBs since the OHWP and more recently during pre-legislative scrutiny. SMBs (either themselves or through trade and industry associations) noted the following key concerns relating to ensuring that the OSA does not disproportionately affect smaller platforms:

Table 56: SMB concerns and mitigations

Potential impacts on competition: the need to ensure that innovative and smaller companies are not disproportionately negatively impacted. Large in-scope companies are more likely to design products already in line with regulatory requirements.	Ofcom has a proven track record of balancing robust consumer protection with the need to ensure the regulatory environment is conducive to growth and innovation. Under Option 1, Ofcom will have a legal duty to assess the impact on SMBs and have regard to innovation in production of its codes.
SMB awareness: the need to reach out to SMBs to ensure they understand their obligations and to reduce the cost of familiarisation.	Ofcom and the government will work together to engage SMBs throughout implementation and ensure obligations are clear and aimed at SMBs.
Technology requirements: the need to ensure a balance between mandating technology and ensuring SMBs are not required to employ technology which they cannot afford.	The government will only mandate specific technologies in very limited circumstances such as to identify and remove illegal terrorist content or CSEA content and only where this is the only effective, proportionate, and necessary action available, and the regulator is confident that the tools available are highly accurate
Clear codes and guidance: the need for clear and easy to understand codes and guidance. Most SMBs do not have teams of regulatory compliance staff and prefer things such as checklists.	Guidance and codes produced by Ofcom will be clear, accessible, and easy to understand. It will also ensure guidance is aimed specifically at SMBs.
Transparency reporting: the need to ensure thresholds are set at such a point to avoid the	Thresholds for designation as Category 1, 2A, and 2B will be set out in secondary legislation. It

unnecessary inclusion of SMBs within this requirement.	is likely that the information requested will vary between different. In every case, however, Ofcom must take account of the capacity of the provider of the service.
Non-prescriptive guidelines for risk assessments and transparency reports: prescriptive requirements related to the way in which platforms assess risk and report on harm is likely to disproportionately impact SMBs.	While certain information will be required in both a platform’s assessment of risk and reporting of harm, overall, the information requested, and the systems and processes used by platforms will vary greatly. Ofcom will consult SMBs to ensure guidelines are not overly prescriptive.
Alignment with existing global regulations: the need to avoid creating unnecessary burdens on SMBs and ensure requirements align with other countries’ regulations.	The government and Ofcom are continuing to assess potential areas of alignment in terms of compliance activities and are working closely with many international partners to address this shared challenge in order to build consensus around shared approaches to internet safety and to learn from other nations’ experiences of tackling online harm.
Proportionality: the need to ensure the principle of proportionality through implementation of the legislation.	Proportionality is at the heart of Option 1 and Ofcom will work closely with affected SMBs to ensure requirements are feasible and proportionate.
Continued engagement with SMBs: the need for the government and Ofcom to continue to engage with SMBs.	Engagement with SMBs is ongoing and will continue throughout implementation of the regime.

368. SMB concerns raised during engagement have been instrumental in the design of Option 1 and the government’s commitment to proportionality. Ofcom will continue to engage SMBs on future codes in an attempt to ensure impacts are proportionate to both the risk of harm and a platform’s resources.

SMB mitigations

369. This section sets out how the potential mitigations for SMBs identified by the RPC have been considered.[319]

[319] Checklist tool for a high-quality SaMBA - RPC (August 2019)

370.

Table 57: SMB mitigations

Potential mitigations (as suggested by the RPC)	How they have been considered in the OSA
Differentiated regulatory approach and requirements, which will likely apply to most small businesses	Most in-scope businesses will only be required to respond to illegal content and put in place measures to protect children (including from online content/activity which may be legal for adults, e.g. pornographic). A narrower range of service providers (Category 1) will be additionally required to consistently enforce their terms of service relating to the restriction or removal of user-generated content. This will form a broader duty regarding the safety of *all* users. Additionally, only Category 1, 2A, and 2B businesses will be required to publish transparency reports. We expect a small number of only large businesses to be designated.
Partial exemptions - use of derogations and de minimis measures (e.g. use of warnings to businesses rather than applying sanctions where non-compliance is identified)	Exemptions will apply to online product and service reviews as well as ‘below the line’ comments. This will reduce the regulatory burden on many low risk businesses who have a low degree of user interactions and UGC. Many of these will be SMBs. Enforcement measures will begin with confirmation decisions ahead of any sanctions being issued. The regulator will have the discretion to set the level of fines which will consider the size of the business (revenue, users, staff) alongside the actual or potential harm caused.
More discretion for smaller businesses to meet regulatory requirements* (e.g. extended transition period or temporary exemption)	This was not considered separately as the preferred approach already builds in significant discretion for businesses to decide how to meet regulatory requirements. businesses will not face prescriptive requirements but will be expected to assess their level of risk and put in place proportionate measures to address this. Laying of codes will undergo consultation and IAs and will be staggered allowing time for SMBs to comply with individual codes, as opposed to a specific date in which the whole regime comes into force at once.

Potential mitigations (as suggested by the RPC)	How they have been considered in the OSA
Simpler and clearer guidance on how to comply. More compliance support for small businesses from the government and regulators	As well as the requirement to be consistent with the principle of risk-based and proportionate action, Ofcom will also be required to have regard to the need to: ○ ensure all businesses can understand and fulfil their responsibilities and ○ cater for all businesses whatever their risk level and capacity (for example by providing support to start-ups and SMBs, drawing on best practice in other sectors). Businesses will not be obliged to comply directly with all the contents of the codes of practice; they may implement alternative approaches provided they can demonstrate that these are as effective or are more effective. The government is also developing a Safety by Design framework targeted at SMBs that will support businesses in adopting a “Safety by Design” approach, helping them design in user-safety to their online services and products. This work will produce practical online guidance tailored to SMBs. The framework will support SMBs to prepare for the introduction of their duties. In addition, the government is undertaking several measures to stimulate and grow the UK commercial market in products and services supporting online safety, so that businesses in scope of the OSA have a greater choice of tools they need to monitor online behaviour or protect users, at appropriate price levels.
Stronger culture of transparency and learning*	Ofcom is a centralised body with a clear remit and responsibility to lead efforts to share learning and encourage collaboration between businesses and between sectors and to promote innovation and best practice. It will have a dedicated digital, data and innovation function to lead these efforts. Ofcom has a culture of proactive monitoring, evaluation and improvement, working with a range of stakeholders including industry, civil society and users to be continuously improving, refining and innovating. For example, a rigorous approach to understanding business impact based on on-the-ground research would help it to understand what’s working well and where businesses might need more support. It will also focus on collaborative methods for policy and implementation and focus on inclusion of a broad range of stakeholders. In addition, Ofcom will be required to conduct IAs on all new (or revised) codes of practice with further requirements to specifically

Potential mitigations (as suggested by the RPC)	How they have been considered in the OSA
	assess the impacts on SMBs and innovation - this goes beyond normal regulatory requirements as set out under the SBEE Act 2015.
Different requirements for different sizes of businesses	As mentioned above, not all businesses will be expected to respond to all categories of harm: many, and most SMBs, will only be required to respond to illegal harm and to protect children online. Furthermore, the regulator's codes of practices will set out proportionate requirements. For example, the legislative requirement to have effective and accessible mechanisms for user redress will vary between businesses; the smallest and lowest-risk businesses might only be expected to have an email address for contact (which is already a legal requirement under the Electronic Commerce Regulations 2002). SMBs will unlikely be required to pay the annual fee or notify the regulator as they will fall under the notification threshold set by the regulator.
Financial aid (e.g. reimbursement of compliance costs)	Whilst there may not be reimbursement of payments from businesses to the regulator, there are mechanisms in place to ensure that any non-enforcement related payments from businesses are not disproportionate. The fee will be tiered and informed by the regulator's regulatory timesheet data. The annual fees charged to industry will therefore be informed by the total quantum of costs incurred by the regulator in running the online safety regime, therefore the fee is proportionate. The regulator should not be in a position to reimburse businesses or not be able to cover any regulatory costs.
Opt-in and voluntary solutions	Voluntary approaches have been tested in the sector but have not been successful (see rationale for intervention).

Medium-sized business assessment

371. To comply with the latest government guidance, this assessment examines the feasibility of applying exemptions to medium-sized businesses.[320] For the purpose of this assessment, the RPC defines a medium-business as a business that has between 50–499 employees, while the evidence from engagement with platforms to inform this IA uses the medium-businesses definition as 50-249 employees, as defined within the Companies Act 2006.[321] While this IA has not considered any evidence specific to businesses with 250 to 499 employees as they have been categorised as large-

[320] Medium-sized business regulatory exemption assessment: supplementary guidance

[321] Companies Act 2006

businesses. It is still expected that many of the points raised would be relevant for these businesses.

Justification for non-exemption

372. As with small and micro businesses, RPC guidance states that the default position is to exempt medium-businesses fully from the requirements of new regulatory measures.[322] However, similarly to the Small and Micro Business Assessment (SaMBA), the same evidence suggests that the objectives of the regulations would be compromised by exempting medium-businesses.

373. The scope of the policy has been intentionally designed to encompass a wide range of services that are linked to causing harm. The primary focus is on mitigating harm, which has a limited relationship between attributable harm and the categorisation of size. This approach is consistent with the principles outlined in SaMBA. Furthermore, introducing exemptions could lead to users shifting from larger platforms to medium-sized ones. This assumes platforms with similar resources and functionalities can be seen as viable alternatives to one another. Considering this, transitioning from a larger platform to a medium-sized one would likely be an easier transition compared to moving to a small or micro platform. This potential migration would undermine the objectives of policy and highlights the implications any exemptions would have on the objective of harm reduction.

Impacts on medium-sized businesses

374. This IA estimates that there are around 2,910 shown by Table 55, medium businesses within scope of the OSA. The in-scope medium businesses are estimated to fall within the following risk categories.

375. As illustrated by Tables 43 and 44, the largest costs are expected to fall on medium and large businesses, who are better resourced to absorb them.

376. To minimise transition costs for medium businesses, a proportional approach is crucial due to their varied resources. Medium businesses operate in a transitional phase, positioned between the more established large enterprises and the relatively smaller, emerging businesses. This transitional nature can result in a diverse mix of resources. Considering the wide spectrum of resources within this category, it is essential to adopt a proportional approach when it comes to complying with codes and expectations.

377.
Medium-sized businesses are expected to receive clear guidance, have proportionate expectations, and access codes tailored to their specific size and risk levels. Nevertheless, these provisions may vary to some extent when compared to the treatment of small and micro businesses. This ensures effective harm reduction while minimising compliance costs. The government and Ofcom will work with medium businesses, recognising their diverse resources, to establish proportional requirements and enforcement, lessening burdens on these businesses.

378. The OSA's regulated services are classified into different categories, which are not necessarily determined by the size of the business. Business sizes are based on employee numbers, while categories such as category 1 platforms are expected to be platforms with higher reach and greatest influence over public discourse. Ofcom will access this based on relevant characteristics such as how quickly, easily, and widely user-generated content is disseminated, along with the number of users and functionalities of the service. It is unlikely that medium business would be included in this or any category. However, if included, these services will have to comply

[322] Small and Micro Business Assessments: guidance for departments, with case history examples - RPC (2019)

with additional measures and consequently face larger costs. Medium businesses may also have more limited resources when compared to larger firms to adhere to these measures. Ofcom will have a legal duty to assess the impact of codes on businesses. As the regulator, Ofcom will take proportionate and targeted approaches on identified platforms where the risk of harm is the highest. The codes will provide flexibility for different sizes of businesses and resources.

Unregulated medium-sized businesses

379. Like the SaMBA section, as fraud advertising duty only applies to Category 1 and 2A platforms it is also expected to result in costs to a significant number of out-of-scope medium-businesses that advertise on in-scope platforms. With costs representing providing information to support anti-fraud checks. Table 55 shows these costs to medium businesses. This impact assessment estimates approximately 70,000 medium businesses will incur some costs in the first year. The two tables (Table 53 and 54) outline the per-business cost to medium businesses expected to undergo standard and enhanced CDD which is made up of staff time to provide necessary information.

380. Shown by Table 55 medium businesses face higher costs with more varied resources when compared to larger businesses. However, when using the proportionate and risk-based design Ofcom will consider a platform's resources in relation to compliance to decide what is proportionate.

Findings from medium-sized businesses engagement and mitigations

381. During the government’s engagement with platforms, medium businesses also noted key concerns to ensure the OSA does not disproportionately affect them. All of the concerns and mitigations mentioned in the SaMBAs (Table 51) apply to medium businesses with some small differences included below.

382. Some medium-businesses have stated that they do not have the resource or technical expertise to meet the user-verification duties, and/or that doing so would disproportionately impact the platform functionalities. The government acknowledges this concern, however the user-verification duties are only applicable to Category 1 platforms, so as not to place excessive burden on smaller platforms. If a platform is designated as Category 1, it is likely that the duties would not be disproportionate.

383. In terms of clear and user-friendly guidelines, medium businesses may have more resources to allocate to regulatory compliance staff. However, considerations will still need to be considered to account for the differing resource levels among medium businesses. The government will strive to provide guidelines that are easily understandable and accessible, considering the varying levels of resources.

384. Regarding support on risk assessments and transparency reports, medium businesses are unlikely to receive the same level of assistance as small and micro businesses. The thresholds for these requirements may be larger for medium businesses, but the intention is not to make these requirements overly prescriptive. Ofcom will work with medium businesses to ensure resourcing and proportionality are taken into account.

385. The RPC mitigations as mentioned in the SaMBA (Table 55), while specific to small and micro businesses, still provide useful concerns and mitigations to consider for medium businesses.

Wider impacts

Trade impacts

Does this measure have potential impacts on [the value of] imports or exports of a specific good or service, or groups of goods or services?

386. The OSA will apply to any in-scope service provided to UK users regardless of where the service is based. The scope of the framework's core duties is functionality based, i.e. it is both goods/service and sector agnostic. It is difficult in the context of online platforms and online harm to apply the import/export framework to assess potential impacts. For example, the UGC/P2P interaction functionality offered by an online platform could be the service itself – in which case a normal trade in services framework would apply – or it could be a minor part of the online presence of a business which attains revenue from an unrelated good or service.

Where UGC/P2P interaction is the main offering

387. The UK is an important market for many of the most affected types of organisations, for example:

- **Social media and search engines**: social media businesses' main offering to users and advertisers is the UGC and P2P interaction[323]. As a result, Facebook accounts for over 50% of the display advertising market and with regard to search engines, Google controls 90% of the search advertising market.[324] The UK is the 13th largest market in terms of user base for Facebook[325], 8th for Instagram[326], 5th for X [327], 4th for Snapchat[328], and 3rd for Pinterest. In terms of traffic, the UK is responsible for 4.1% of traffic to Google (the third largest share behind the US and Brazil).[329]
- **Online marketplaces**: the UK is the third largest market for Amazon and second largest market for eBay representing 8% and 18% of total global traffic to the sites respectively. The UK market is of equal importance to smaller online marketplaces placing second for its share of global traffic for platforms such as Etsy and Wayfair[330].

388. Given the value of the UK market to these businesses, it is unlikely that the OSA would lead to a reduction in services offered to UK users (or UK advertisers). Platforms offering UGC and P2P services to UK users will not be at a significant disadvantage from those that operate elsewhere as the regulatory landscape for online platforms is evolving internationally. Similar regulations to the OSA are being developed or have already been implemented internationally - Germany's NetzDG was implemented in 2018 and the EU is developing their Digital Services Act. Other countries are also expected to follow suit.

389. Compliance costs associated with Option 1's fraudulent advertising duties, will make the process of advertising to UK consumers more expensive (if compliance costs are passed on to

[323] It could be argued that the main offering to advertisers is the user base (rather than UGC and P2P interaction specifically); however, the ability of users to react, like, discuss and share is what sets social media advertising apart from traditional forms.
[324] Online platforms and digital advertising - Market study final report (CMA, 2020)
[325] Leading Countries Based on Facebook Audience Size as of January 2021 - Statista
[326] Instagram Demographic Statistics: How many people use Instagram in 2021? - Brian Dean
[327] Leading Countries Based on Number of Twitter Users as of January 2021 - Statista
[328] Leading Countries Based on Snapchat Audience Size as of January 2021 - Statista
[329] Regional distribution of desktop traffic to Google.com as of June 2021, by country (Statista, 2021)
[330] https://www.webretailer.com/b/online-marketplaces-uk/

advertisers). However, many platforms are already deciding to implement anti-fraud measures and the cost of both conducting (for in scope platforms) and undergoing (for advertisers) a customer due diligence check is expected to be relatively modest for each individual business. This impact assessment does not expect anti-fraud measures to negatively impact the provision of advertising space or the decision of non-UK-based companies advertising to UK consumers.

390. Unlike a business providing an online service, if the cost of regulatory compliance becomes excessive in one country for a business manufacturing goods, given the business's finite productive capacity, it would be worthwhile instead selling the goods elsewhere where regulatory burdens are lower. This is not the same for businesses in the digital markets whose main offering is UGC and P2P interaction. Due to the nature of digital markets, there are limited constraints on the provision of an online service, e.g. on the number of users/consumers. In a digital market the decision to provide a service is solely based on whether the benefits from providing the service in that country, for example, ad revenue or similar, exceed the cost of compliance. This IA estimates a relatively modest per business cost of compliance which is proportional to business risk, the likelihood of online platforms withdrawing their services from the UK in favour of providing their services elsewhere because of the proposed regulation is minimal.

391. For services currently offered to UK users only, who may in the future, look to enter other markets, this IA does not expect compliance costs to put them at a competitive disadvantage. The cost of complying with the regulation will increase business costs; however, businesses will be in a more favourable position to compete on user safety. Over half of respondents to an Ofcom survey have spontaneous (not prompted by the interview question) concerns about interaction with other people/content online.[331] Moreover in Ofcom's Online Nation 2021 report, 61% of respondents agreed with the statement: "Internet users must be protected from seeing inappropriate or offensive content".[332] Over the past year there has been increasing public pressure on platforms to take further steps in addressing online harm, particularly for categories of harm such as disinformation and online abuse. Given the general public's concerns about internet safety, compliance with the OSA could be a competitive advantage for UK providers[333] on the international stage.

Where UGC/P2P interactions are secondary

392. Some businesses - that may not be considered traditional digital businesses - will be within scope of the regulations solely due to offering UGC or P2P interaction functionality on their website. For example, a business which sells a traditional good or service (retailers, legal services etc) but that offers a forum function on its website could be in scope. As noted earlier, compliance with the OSA will increase the cost of doing business for these organisations. However, given the risk-based design of the framework, any compliance costs are expected to be proportionate. Further, the introduction of the 'low risk' functionality exemption has removed a large proportion of these types of businesses from scope, for example, small hospitality, beauty and health businesses, where there is simply a comment function for reviews on their products. At the margins, some of these businesses - still in scope after all the exemptions - may remove some functionalities from their websites instead of incurring compliance costs. This could result in a reduction in the quality of the customer experience when engaging with such businesses. 73% of customers find live chat the most satisfactory form of communication with a company[334].

[331] Internet Users' Experience of Potential Online Harms: Summary of Survey Research - Ofcom (2020)

[332] Online Nation 2021 report - Ofcom (2021)

[333] 'UK providers' here refers to platforms providing services to UK users only.

[334] https://99firms.com/blog/live-chat-statistics/#gref

Does this measure include different requirements for domestic and foreign businesses?

393. The framework will apply to any in-scope business worldwide that provides services to UK users. There are no differing requirements for domestic and foreign businesses. Applying this policy to all businesses providing services in the UK will help to ensure a level playing field between businesses that have a legal presence in the UK, and those who operate entirely from overseas. The UK is paving the way in this regulatory landscape with countries worldwide following suit. There may consequently not be a marked difference in operating costs between similar jurisdictions as other countries look to align.

Does this measure have potential impacts on [the flow or value of] investment into and out of the UK?

394. There is a risk that the regulation could dissuade foreign investment and/or encourage UK based organisations to disinvest in the UK if the compliance costs are too high. The arguments presented above on trade apply equally for investment in so far as businesses are not expected to stop providing services to UK users and compliance costs are not expected to stop platforms who provide services to UK users to be able to provide services to non-UK users.

395. There is evidence to suggest that, in the short- to medium-term, there will not be a large net outflow of investment, especially from digital sectors. The largest businesses have large and sticky investments in the UK market. They also have large investments in value-add employment (that is, not just selling to UK customers but services that can be exported): the UK hosts the largest Facebook engineering base outside of the US, and Apple has a large R&D centre in Cambridge. Large businesses are already taking measures to combat online harm, the government would therefore expect there to be a minimal impact upon their investment and business activity within the UK.

WTO Notification

396. The WTO requires members to "promptly or at least annually issue notifications of new or amended legislation that will 'significantly affect' international trade in services under the GATS". On advice from the Department for International Trade the government will not be required to notify the WTO about this legislation.

Competition assessment

Competition in digital markets

397. In July 2020, the Competition Markets Authority (CMA) published the final report of its market study into online platforms and digital advertising.[335] The findings highlight several characteristics of digital markets that inhibit entry and expansion by rivals, undermining effective competition. These include network effects and economies of scale, the power of default placement (for example being assigned the default search engine on an internet browser),[336] unequal access

[335] Online Platforms and Digital Advertising Market Study - final report - CMA (2020)

[336] In 2019, Google paid around £1.2 billion in return for default position in the UK, a majority of which was to Apple for being the default on the Safari browser. Such payments are one of the most significant factors inhibiting competition in the search engine market.

to user data,[337] lack of transparency (in terms of decisions made by platforms), ecosystems of complementary products and services, and vertical integration as large digital platforms are present at multiple stages of supply chains. On this basis, there are significant barriers to competition present under the baseline.

398. In terms of online search engines, Google has persistently had a high and stable share of the general search market, with a share of supply between 89% and 93% over the last three years.[338] In June 2021 Bing and Yahoo Search had the next two highest shares at 5.6% and 2.7% respectively.[339] Similarly to Google, Bing holds extensive default positions through Microsoft's agreements with Windows PC manufacturers. These extensive default positions limit the expansion of rival search engines through limiting their accessibility to consumers, preventing new entrants from developing into strong competitors. Existing smaller platforms in the market are often syndication partners of Google or Bing, relying on the larger search engines for their search results and adverts[340]. These businesses seek to attract customers through other means, for example DuckDuckGo's unique selling point is its focus on privacy. In search advertising, Google is by far the largest player with the CMA stating that potential rivals can no longer compete on equal terms.[341]

399. In the social media market, the extent of competition between platforms is dependent on the degree to which users consider them as substitutes. Social media platforms offer similar types of functionality although they are differentiated based on particular consumer needs based on the type of communication and content consumption provided. Despite this, evidence indicates that Facebook has a significant and enduring market power in social media. Between July 2015 and February 2020, Facebook had a share of 54% of user time spent in social media[342].

400. When looking at the market for VSPs, based on analysis of the number of users watching videos on platforms and the number of video views on such platforms, the Herfindahl–Hirschman Index (HHI)[343] is greater than 2,500 indicating a highly concentrated sector, this has been the case since 2017. Only a limited number of platforms have entered the sector and achieved scale in recent years. Consumers use a limited number of platforms to view videos online, 38% of people said that the main reason they use a platform to watch videos is because it was the first platform they used, suggesting a degree of consumer inertia.[344]

401. While it is important to acknowledge potential competition impacts of Option 1, many of the main online markets are highly uncompetitive currently. Many of the requirements under Option 1 such as transparency reporting, user redress, and privacy protections may go some way to mitigating some of the current problems.

[337] Analysis of a trial run by Google in 2019, comparing the revenue publishers received from personalised advertising with revenue from non-personalised ads, suggests that UK publishers earned 70% less revenue when they were unable to sell personalised ads. The inability of smaller platforms and publishers to access user data therefore creates a significant barrier to entry.

[338] Online Platforms and Digital Advertising Market Study - final report - CMA (2020)

[339] Worldwide market share of search engines - Statista (2010-2021)

[340] Online Platforms and Digital Advertising Market Study - final report - CMA (2020)

[341] Online Platforms and Digital Advertising Market Study - final report - CMA (2020)

[342] Online Platforms and Digital Advertising Market Study - final report - CMA (2020)

[343] The Herfindahl–Hirschman Index (HHI) is used to assess the level of concentration in a sector.

[344] Understanding how platforms with video sharing capabilities protect users from harmful content online - EY, DCMS (2021)

Potential impacts on competition of the OSA

402. While the rapid evidence assessment of Germany's NetzDG did not find any evidence that the policy had any impact on market competition, the proposals under Option 1 are not limited to large social media companies. Option 1 could potentially impact competition in the market if:

- ○ compliance costs create – or are viewed by potential new entrants as - a barrier to entry; or
- ○ costs fall disproportionality on SMBs, i.e. they are not able to absorb the costs (in unit terms) as easily as larger businesses; or
- ○ compliance costs dissuade foreign investment and/or encourage UK based businesses to disinvest in the UK; or
- ○ compliance with the legislation creates friction for users' consumption of online platforms.

Will the measure indirectly or directly limit the range or number of suppliers?

403. The proposals could indirectly limit the number of suppliers if for example, compliance costs are seen by potential entrants to the market as barriers to entry or realised costs of compliance force some providers out. The growth of the UK digital economy outpaces that of most other sectors.[345] The fast-paced nature of this evolving market can result in platforms scaling rapidly; however, the financial benefits of the achieved scale can be delayed. Therefore, it is possible that a firm be deemed high-risk and not yet have the financial resources available to comply with the legislation. This could potentially result in realised costs of compliance forcing platforms out of the market. The proportionate enforcement expected of the regulator will be essential in minimising this impact.

404. For a low risk in scope micro-businesses, beyond familiarising themselves with the regulations, they may only be required to produce a risk assessment, ensure it has an email address for potential user reporting and conduct no or minimal additional content moderation (one small low risk organisation interviewed for example, noted that moderating was already a part of business as usual and 'negligible'). The impact on such businesses is expected to be limited. Given the differentiated requirements on businesses (of size and risk) and the proportionate enforcement expected of the regulator, these impacts are expected to be minimal.

405. Option 1 is expected to result in impacts on some out-of-scope SMBs through requirements under the fraudulent advertising duty. These SMBs - that participate in paid-for advertising on Category 1 and 2A platforms - will incur costs associated with providing necessary information to ensure they are legitimate businesses. Platforms like Google and Facebook do offer low friction advertising opportunities especially to small businesses with an estimated 63% of SMEs advertising this way.[346] These costs on out-of-scope SMBs are expected to be minimal, involving between 10-30 minutes of staff time only. It is still the case that Ofcom must engage with SMB advertisers as it develops codes of practice and ensure any compliance burdens are minimal.

Will the measure limit the ability of suppliers to compete?

[345] DCMS Sectors Economic Estimates 2019 (provisional) Gross Value Added - DCMS (2020)

[346] Powering Up: Helping UK SMEs unlock the value of digital advertising (IAB, 2020)

406. For platforms where UGC and P2P interaction is secondary to the good or service being sold, this measure is not expected to limit their ability to compete given the main areas of competition (price and quality) are largely unrelated to that aspect of their website. These businesses may find that the cost of compliance is not worth the benefits of having this functionality on their site and they may remove it.

407. However, for platforms where UGC and P2P interaction is the service, this proposal may reduce smaller businesses' ability to compete. For example, size is not a perfect proxy for risk of online harm (although there is a link) and therefore, a business like Facebook may be in the same risk tier as a much smaller (in terms of employees and revenue generation) social media business. Businesses in the same risk category are bound by the same duties and given that Facebook (in our example) will find it much easier to absorb compliance costs than the smaller social media platforms there may be distortionary effects. To limit this, there will be differentiated requirements within duties - for example, while all Category 1 businesses will have to report on transparency, the information they are required to collect and publish may vary proportionately depending on the requirements set out in future codes. Additionally, based on the intention of the policy, small or micro businesses are not expected to be designated as Category 1. It should be noted that the pornography provision ensures that all businesses regardless of size will be required to prevent children from accessing pornography. While this has the potential to result in burdens on SMB platforms that host pornography, the government considers the protection of children a core objective of the OSA. Further, given the nature of costs for age verification solutions (largely based on the number of checks), sites with a larger user base will pay more.[347]

Will the measure limit the suppliers' incentives to compete vigorously?

408. Regulation of online platforms will have a minimal impact upon the suppliers' incentive to compete. There is a risk that the regulation could inadvertently encourage collusion (e.g. sharing data, forming research groups and sharing technology), however, this risk is expected to be negligible. By introducing a minimal level of online harm action this proposal could potentially limit businesses' ability to compete on that aspect of their services, i.e. user safety. However, a thriving digital economy is at the heart of the government's vision for long-term economic growth. As such, the growth of digital markets will be supported by initiatives including the pro-competition regime for digital markets which will encourage competition in this sector.[348]

Will the measure limit the choices or information available to consumers?

409. The policy will increase information available to consumers through bridging the information gap between businesses and consumers through increased transparency, as detailed in the Rationale for Intervention. This will allow consumers to make informed decisions about their use of online platforms and purchase of online goods and services, driving greater competition between businesses to implement measures meeting regulatory and consumer demands for increased safety on online platforms.

[347] The number of users is not a perfect proxy for platform size, but they are related.

[348] A new pro-competition regime for digital markets - DCMS (2021)

Innovation test

Innovation in digital markets

410. Investment in primary technologies, including artificial intelligence and machine learning, provide an indication of the level of innovation within digital markets. In 2020, the UK had the second highest proportion of venture capital investment into these foundational technologies, accounting for 54% of total venture capital investment.[349] UK investment in the technology sector has significantly increased over recent years. Impact tech investment[350] in the UK has more than doubled since 2018, a 106% increase, in the same period the US saw only a 15% increase.[351] The UK is the third in the world for impact tech investment. These large-scale investments into the technology sector indicate high levels of innovation, providing the resources for innovation in digital markets.

411. The success of online marketplaces illustrates the value of eCommerce innovation. In the UK the largest marketplaces such as Amazon and eBay accommodate millions of customers with 407 million visits and 298 visits in April 2021.[352] Marketplaces can provide a streamlined process of servicing and selling with access to an extensive global consumer base. Innovation in this market over the years has enhanced consumers' experiences. This includes the use of smart eCommerce which enables the supply of a customised list of recommendations based on consumer behaviour and history to provide a tailored online experience. AI has enabled marketplaces to provide 24/7 customer service using chatbots, providing instant answers to simple questions.

412. There has also been considerable innovation in the gaming industry. In the past year Fortnite have hosted events including in-game concerts and movie trailer premiers.[353] It also anticipated that these innovations could develop into the creation of a digital metaverse, a virtual experience going beyond gaming to provide an array of media experiences.[354] The development of virtual reality has also augmented the gaming experience through providing an immersive gaming environment.

413. There is currently large-scale investment in research and development among the largest online platforms, indicating significant levels of innovation. In 2020 Amazon's R&D expenditure amounted to $42.7billion[355] (£33.3billion), similar levels of investment in the same period were seen among other platforms including Google $27.5billion (£21.4billion) and Facebook $18.4billion (£14.3billion).[356] In the past GAFAM companies have delivered breakthrough and disruptive innovations improving consumers' lives and creating jobs. Digital firms have also disrupted existing markets including the taxi and hotel industries.

414. However, there are concerns that the dominant firms that have emerged from the growth of digital markets are constraining further innovation.[357] As explored in the 'Competition Assessment',

[349] The Future UK Tech Built, Tech Nation Report 2021 - Tech Nation

[350] Impact tech investments are investments in technology made to generate positive social and environmental impacts alongside a financial return#.

[351] The Future UK Tech Built, Tech Nation Report 2021 - Tech Nation

[352] Leading online marketplaces in the United Kingdom as of April 2021, based on number of monthly visits - Statista

[353] The 10 most innovative social media companies of 2021 - Fast Company

[354] The 10 most innovative social media companies of 2021 - Fast Company

[355] Amazon Research and Development Expenses 2006-2021 - Macrotrends

[356] The average exchange rate (1 USD = 0.7798 GBP) in 2020 was used to present figures in GBP.

[357] Competition and Innovation in Digital Markets - BEIS (2020)

certain characteristics of digital markets inhibit the entry and expansion by rivals. In digital markets innovation requires access to data, users and fair returns.[358] The biggest digital platforms control some, if not all, of these elements. Established platforms have access and control over data,[359] a loyal consumer base,[360] and can exploit their market power to extract an unfair share of returns from successful innovation.[361] Therefore, the aim of Option 1 is to minimise any indirect impacts of regulatory compliance on wider innovation.

Potential impacts on innovation of the OSA

415. While sector agnostic in its design, Option 1 is risk-based and therefore, most requirements will fall on businesses with websites offering high levels of UGC and P2P interaction functionalities, for example, social media and other digital technology businesses. These types of businesses are high-growth and highly profitable businesses, as such these companies invest considerably into research and development. The compliance requirements of this framework will therefore disproportionately fall on highly innovative sectors. However, these platforms are already investing substantially into user safety, and it is therefore assumed that they do not necessarily see a trade-off between user safety and innovation.

416. The impact on smaller businesses and start-ups will depend on the degree to which proportionality is built into the system, and the ways in which the independent regulator is able to reduce the burden on SMBs. The SaMBA above outlined a number of potential mitigations for SMBs - these include: partial exemptions; proportionate enforcement; duties with significant discretion for businesses to decide how to meet the requirements; clear and tailored guidance for SMBs, including in advance of legislation, a voluntary Safety by Design framework targeted at SMBs; a practical compliance support function for SMBs built into the regulator; and a proportionate fee structure which considers business size.

417. Protecting and encouraging innovation is a key consideration for the framework. The policy has been designed from the start with innovation at the forefront:

- by implementing through primary legislation and codes of practice, it gives the regulator flexibility to lay and revise codes of practice as new technologies emerge
- Ofcom will have a legal duty to pay due regard to innovation in the exercise of all of its functions
- there is a specific requirement on the regulator to produce IAs for all new and revised codes of practice and to ensure within these, that the impact on innovation is considered.
- the framework is principles-based and businesses are given the freedom to meet high-level requirements in the most efficient way allowing them to undertake alternative measures that prove to be sufficiently effective.
- options analysis considered the adaptability to future technological changes as a key criteria and impact on innovation.

[358] Big Tech: how can we promote competition in digital platform markets? - Amelia Fletcher, Economics Observatory (2021)

[359] Big Tech: how can we promote competition in digital platform markets? - Amelia Fletcher, Economics Observatory (2021). Amazon, Apple, Facebook and Google are estimated to hold around 1.2 billion gigabytes of data between them.

[360] Understanding how platforms with video sharing capabilities protect users from harmful content online - EY, DCMS (2021) There is evidence to suggest that a degree of inertia exists among consumers in the VSP industry.

[361] Big Tech: how can we promote competition in digital platform markets? - Amelia Fletcher, Economics Observatory (2021)

- implementation of the policy will be risk-based so the regulator can focus resources on the most serious categories of online harm (even if that changes).
- the approach taken will be technology neutral and therefore encompass future changes to how the architecture of the internet functions.
- development of the online safety implementation measures which will focus on researching emerging harm and the working safety technology sector to encourage innovative solutions to the problems.
- proportionate system (e.g. smaller and less risky businesses have to do less), this will minimise the disincentive effects of the regulation and minimise the impact on new entrants.
- partial exemptions will be implemented to reduce the regulatory burden on many low risk businesses who have a low degree of user interactions and UGC. Many of these will be SMBs.

418. Consideration of innovation has been at the forefront of policy design and will continue to be during its implementation. For the reasons noted above, indirect impacts on innovation are expected to be negligible. Finally, the monitoring and evaluation (M&E) section outlines a detailed plan which will consider the policy's impact on innovation and any unintended effects in this area.

Equalities impact

Statutory Equalities Duties	Completed
Proposals set out in the OSA to make the internet a safe place for all users are expected to have an overall positive impact on individuals with protected characteristics. The government is not aware of any possible direct discrimination, in relation to the OSA, and when considering indirect discrimination various elements of framework are expected to positively impact users with protected characteristics. These elements include a higher level of protections for children, requirements to assess risks to users, requirements for major platforms to clearly state what content is considered acceptable in their terms of service and to enforce these consistently and transparently, further promotion of media literacy, the establishment of a super-complaint function, and the requirement for all services to have easily accessible user redress mechanisms. **Overall, the proposed framework will help advance the protections of the Equality Act 2010 online and make the internet a safer place for all, including those with protected characteristics.** **The Senior Responsible Officer has agreed with these findings.**	**Yes**

419. The government has a legal obligation to consider the effects of policies on those with protected characteristics[362] under the Public Sector Equality Duty 2011 and the Equality Act 2010.

420. Overall, these proposals are expected to have a positive impact on users with protected characteristics. This is incorporated in the overarching aim of the policy; to make the internet a safe place for all users. Reducing online harm is particularly important for those with protected characteristics, many of whom are disproportionately more likely to be victims of online abuse and discrimination, for example:

[362] Age, disability, sex, gender reassignment, pregnancy and maternity, race, religion or belief and sexual orientation

- a 2019 report by the Alan Turing Institute found that Black people and those of 'Other' ethnicities are far more likely to be targeted by, and exposed to, online abuse than White and Asian people, with 39% of Black people having observed hateful/cruel content online compared to 27% of White People.
- between January and June 2021, Community Security Trust recorded 1,308 anti-Jewish hate incidents nationwide in the first half of this year. This is a 49% increase from the 875 incidents recorded in the first six months of 2020, and is the highest total CST has ever recorded in the first half of any year.[363]
- users with disabilities have been forced to leave social media because of the abuse they had experienced online.[364]
- women tend to be disproportionately affected by online offences like harassment, stalking, revenge pornography.

421. Vulnerable groups, particularly those with mental health problems, are at a much higher risk of falling victim to online scams. A 2021 report found that people who have experienced mental health problems are nearly three times more likely to have been a victim of an online scam than the rest of the population (23% of those with mental health problems were victims of online scams vs 8% of the wider population)[365].

422. It should also be acknowledged that there are potential distributional impacts because of the possibility of introducing age assurance processes. For example, those from disadvantaged backgrounds, including those with disabilities, are often less likely to hold form identification[366]. How much of an impact will be heavily dependent upon the level of verification and identification required.

423. The assessment of prospective equality impacts that Option 1 may have on those with protected characteristics is considered regarding both direct and indirect discrimination:

424. At present, the government is not aware of any possible direct discrimination, in relation to each of the protected characteristics, which will result from this policy.

425. Additionally, when considering indirect discrimination, various elements of the regulatory framework indicate ways in which the policy will positively impact users with protected characteristics. These include:

- **requirement to have clear terms of service and to enforce them effectively and transparently**: platforms will be required to have clear guidance in their terms of service about what is acceptable behaviour on their platform. These may contain explicit guidance about unacceptable behaviours relating to people with protected characteristics.
- **improving media literacy**: some individuals from protected characteristic groups, for example children, the elderly or in some cases disabled people, have been identified as more vulnerable to online harm. The media literacy efforts incorporated in this policy may therefore be particularly important to enable these users to be able to keep themselves safe online.

[363] Antisemitic incidents January - June 2021 (CST, 2021)
[364] House of Commons Petitions Committee report (2018)
[365] Caught in the Web - Online Scams and Mental Health (Money and Mental Health Policy Institute, 2020)
[366] Public Opinion Tracker 2021, Electoral Commission

- ○ **super-complaints**: this function would be open to organisations, who meet a set eligibility criterion, wishing to report systemic failures to comply with the duties across two or more services (or in exceptional circumstances one or more services).
- ○ **requesting that redress mechanisms are easily accessible**: this would ensure that report functions are clear and accessible to all users, including those with protected characteristics who may be otherwise less likely to navigate and pursue them.

426. The government does not expect this policy to impact negatively on people with protected characteristics. This will be monitored post-commencement. However, the focus of the framework on systems and processes, as opposed to content, is intended to avoid this.

427. Overall, the proposed framework will help advance the protections of the Equality Act 2010 online and make the internet a safer place for all, including those with protected characteristics.

Devolution test

428. Internet law and regulation is a reserved policy area under all three devolution settlements. The online safety regime will apply across the whole of the UK.

429. The online safety legislation is reserved, however, there are a number of areas within the regime where there is possible interaction with devolved competencies, and so government is working closely with the Territorial Offices (TOs) and Devolved Administrations (DAs) to ensure that such issues are taken into account. This includes issues such as categories of harm in scope and media literacy.

430. While some of the categories of harm relate to offences in Scottish or Northern Irish Law, and therefore involve devolved competences, the legislation is not seeking to change the law in relation to these offences. Instead, Option 1 clarifies the responsibility of businesses to tackle this activity on their services.

431. The government has engaged regularly with the DAs, TOs, and Ofcom's offices in the devolved nations as proposals have been developed, and it will continue to engage throughout implementation.

Monitoring and evaluation

432. This section lays out the current proposed plans for monitoring and evaluation (M&E); however, these are subject to change as M&E work commences and the programme of research underpinning it progresses. The approach will be iterative and will draw on expertise from across government and external experts.

433. The first stage of the monitoring and evaluation plan discussed in the final stage IA, involving the development of an interim evaluation framework for the OSA has been completed, setting out a theory of change, evaluation questions, and metrics and evidence to be collected as part of the evaluation process.

434. Any review will take a holistic approach and will evaluate the entirety of the online safety framework, including the OSA, Ofcom as the regulator, future codes of practice and secondary legislation and the impact on the digital sector more broadly. There are three main areas of evaluation:

- a review of the wider online safety framework;
- evidence from the implementation of individual codes of practice; and
- an assessment of the government's overall online safety strategy, including the online safety implementation measures, such as media literacy initiatives, child and adult online safety initiatives, investments in the safety tech sector, and safety by design interventions.

Review clause

435. The OSA contains a statutory review clause and a post-implementation review (PIR) will be conducted within 5 years of implementation. At this stage, it would not be wise to provide a more explicit timeline for the review given the fast-moving nature of the policy area and the iterative process of producing codes of practice. It will be for the Secretary of State to determine the specific point at which a review is necessary, this is expected to be between 2-5 years of implementation (and within 5 years) unless there is a clear and obvious reason for delaying or expediting the review.

Review governance

436. The review will be led by the DSIT Secretary of State and they will be responsible for delivering the PIR. However, given that the OSA is a joint policy, both DSIT and the Home Office will share responsibility and work closely with Ofcom to ensure appropriate monitoring and develop the underlying evidence base for online harm. In addition, the review will require input from:

- other government departments;
- regulated online platforms;
- civil society groups; and
- wider society.

437. DSIT, the Home Office and Ofcom are expected to set up an analytical evaluation working group to coordinate on baselining activities, the development of online harm metrics, and research pipelines. This work will be overseen by a senior evaluation steering group, again with representation from the three main stakeholders. Advice and potential involvement will also be sought from established Whitehall expert groups such as the cross-government evaluation group, the Cabinet Office's evaluation task force, and the independent RPC.

Review plans

438. At a high level, the review will consider:

- whether the online safety framework has achieved its stated objectives
- whether the impacts of the policy were in line with those estimated in previous IAs (both primary and codes of practice)
- whether the policy has resulted in any unintended consequences
- how well the regime is functioning in practice and whether there are any areas which could be improved through changes to legislation (or recommendations to the regulator)

439. Most of the initial M&E work is focussed on baselining, developing key metrics, and ensuring that there is a coordinated programme of research to fill evidence gaps. A key strand in the evaluation work will be an assessment of the policy's stated objectives:

- ○ **Objective 1 - to increase user safety online**: Work to understand and baseline the current prevalence of several key types of online harm is underway and this work is expected to result in clear and measurable indicators for illegal and child priority harms.
- ○ **Objective 2 - to preserve and enhance freedom of speech online:** This will be monitored through the collection and reporting of transparency data, such as the amount of content removed/restored; and user satisfaction, such as measuring the effectiveness of redress mechanisms**.** Ofcom already conducts regular high-quality user attitude surveys which will be key indicators for this objective and further research will be undertaken to address any existing evidence gaps.
- ○ **Objective 3 - to improve law enforcement's ability to tackle illegal content online**: This is expected to materialise as efficiency gains or cost savings for law enforcement. This can be measured using crime data and the level of understanding of the drivers of crime, including the specific role of activities in scope in facilitating crime. Addressing online crime will help drive economic growth and enable a stronger online business environment. Assessing the policy against this objective will require consultation with law enforcement and relevant enforcement authorities.
- ○ **Objective 4 - to improve users' ability to keep themselves safe online:** This will draw on Ofcom's comprehensive programme of media literacy and internet use-related research and evaluation of media literacy initiatives.
- ○ **Objective 5 - to improve society's understanding of the harm landscape**: This links closely to Objective 1 and the need to have a clear understanding of how harm manifests and how it can be measured. While important, success against this objective is more subjective than the others. However, the government will draw on Ofcom's programme of user experience research to assess wider understanding of online harm and the joint programme of harm research planned.

Key measures and sources of data

440. The table below outlines some of the potential key measures for the OSA evaluation. As noted, these will largely depend on both government and Ofcom research programmes between now and implementation.

Table 58: Potential metrics for evaluation

Link to objective	Metric/measures	Sources (non-exhaustive)
1 & 2 & 3	Reductions in prevalence of priority and non-priority online harms on in-scope platforms	● Ofcom's adult and child media literacy trackers ● Annual bullying survey ● Police recorded crime data (online flag) ● Counter disinformation monitoring (HMG) ● NFIB fraud reports ● Home Office and Ministry of Justice data ● Ofcom's online experiences tracker survey
1 & 3	Reductions in the spread and flow of illegal content within and across platforms	● Platform transparency reports ● Counter disinformation monitoring (HMG) ● Ofcom's online experiences

Link to objective	Metric/measures	Sources (non-exhaustive)
		tracker survey ● Ofcom's information gathering powers
1 & 3 & 4	Reductions in children's exposure to illegal content and age inappropriate content such as pornography	● Ofcom's child media literacy tracker ● Ofcom's online experiences tracker survey ● Independent research on children's pornography use
3	Improvements in platform performance in areas such as responding to user reports, content moderation, and minimising the algorithmic spread of harmful content	● Ofcom's information gathering powers ● Platform transparency reporting ● Ofcom's compliance reporting
4 & 5	Increases in media literacy indicators, such as awareness of safety features, and interacting with other users safely online	● Ofcom's child media literacy tracker ● Independent media literacy research
2	Improvements in platforms' handling of content takedown challenges	● Ofcom's information gathering powers ● Platform transparency reporting
1 & 2 & 4	Improvements in users' experience of the online environment, with particular focus on children's experiences	● Ofcom's programme of user experience research ● Ofcom's child media literacy tracker

441. It is expected that planned M&E work will be structured around three core phases: Scoping, Mid-term and Final evaluation. Further evaluation work has begun in full following the Act receiving Royal Assent. An overview of the key objectives for each of these phases are set out in the table below.

442. This phased approach allows for flexibility in the evaluation to evolve in tandem with the technologies and behaviours under consideration - both in scope and approach - and to the rapidly changing technology, harm and market context. The approach allows for a future evaluation to be structured around regular review points, drawing on developmental evaluation techniques, designed to ensure lines of inquiry and new hypotheses can be incorporated into the evaluation design as and when they arise. This approach will be particularly amenable to an evaluation of the OSA given the rapidly changing technological and societal context, and potential amendments to the legislation.

Table 59: Proposed evaluation phases

Evaluation Phase	Description
Scoping (Year 1)	● Revise the Theory of Change[367] developed for the OSA, evaluation questions and finalise the evaluation framework based on the scope of the final OSA, after Royal Assent. ● Appraise impact evaluation approaches (incl. a rapid review of feasibility of a quasi-experimental design in case a viable option has emerged since this framework was drafted). ● Develop a data collection plan including research with different impacted groups. ● Capture initial process learning including insights related to: ○ The immediate implementation of the online safety regime. ○ The establishment of Ofcom as the relevant regulatory body. ○ The roll-out of new offences and establishment in case law. ○ The initial responses of regulated services (and associated costs).
Mid-term evaluation (Years 2-4)	● Capture further process learning related to the implementation of the online safety regime. ● Build an understanding of the impact of the online safety regime on different affected groups, exploring emerging evidence of outcomes. ● Identify additional outcomes generated during data collection and revise the evaluation framework accordingly.
Final evaluation (Year 5)	● Provide an overall evidence base for the delivery and impact of the online safety regime including what has worked well and less well. ● Assess the extent to which the OSA has met the original policy objectives set out in the Impact Assessment, and any new outcomes identified during the evaluation. ● Provide evidence to demonstrate whether the online safety regime represents value for money in achieving its objectives.

Proposed Evaluation Questions

443. The following sub-sections present the high-level evaluation questions for the process, impact and economic evaluations. Under each of these high-level evaluation questions will sit a set of more specific and in-depth sub-questions.

Process evaluation questions

444. The process evaluation will aim to establish how the online safety regime was implemented and what can be learned from this, including what worked well or less well, for whom and for what reasons.

- ○ to what extent was the online safety regime implemented as set out in the legislation?
- ○ does the OSA give Ofcom the necessary powers and legislative tools to act effectively as the regulatory body for the online safety regime?
- ○ how effective was the implementation of new and updated offences?

[367] Theories of change are tools to enable causal link monitoring and allow the government to identify assumptions made and fill evidence gaps

Impact evaluation questions

445. The impact evaluation has been structured to assess whether the anticipated immediate and longer-term outcomes as set out in the Theory of Change have been achieved, and the extent to which the online safety regime has led to any unintended consequences. This will build understanding of the difference or impact made by the OSA.

- to what extent, and how, are regulated services making changes because of the online safety regime?
- to what extent, and how, are users changing their behaviours because of the online safety regime?
- in what ways has the online safety regime achieved its intended outcomes and impacts (e.g., increase user safety online, protect privacy, maintain freedom of expression online, reduce the economic costs of online harm)?
- in what ways has the online safety regime produced or contributed to any unintended consequences?

Economic evaluation questions

446. Economic evaluation activity will need to enhance and support an understanding of the costs and benefits involved with the implementation of the OSA, and provide an OSA Monitoring & Evaluation Framework assessment regarding the effectiveness, efficiency and equity of the implementation. Any economic evaluation activity should be compliant with HMT Green Book / Magenta Book guidance.

447. These questions will vary depending on the final scope of evaluation activity; though this is expected to include:

- what costs are incurred by regulated services (and wider organisations)?
- what benefits (and to whom and how) have been generated by the implementation of the online safety regime? How does this compare with the costs of implementation?
- what is the nature and scale of direct costs incurred by regulated services to comply with the OSA?
- what is the nature and scale of indirect costs of the OSA to businesses, markets and/or the wider economy?

The online harm landscape is a fast-moving policy area and the OSA is ground-breaking and novel in its approach. The government recognises the need for a comprehensive and adaptable M&E framework to ensure the policy achieves its objectives and minimises the potential for unintended consequences.