

Title: Measures to Address Fake Online Reviews IA No: DBT032(C)-23-CCP RPC Reference No: N/A Lead department or agency: Department for Business and Trade Other departments or agencies: Competition and Markets Authority	Impact Assessment (IA)			
	Date: 01/09/2023			
	Stage: Development			
	Source of intervention: Domestic			
	Type of measure: Secondary legislation			
Contact for enquiries:				
RPC Opinion: N/A				

Summary: Intervention and Options

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
£1,838m	-£1,846m	£48m	£238m

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

Fake reviews are reviews which have been designed intentionally to mislead consumers and therefore lead to the market failure of asymmetric information. Despite moderation efforts by platforms who host reviews, recent DBT research estimated that approximately 11% to 15% of product reviews on e-commerce platforms are still fake. The research also found that well-written fake reviews make consumers more likely to purchase the product. These findings are aligned with other studies conducted outside of government. Given a growing amount of shopping activity is taking place online, it is important to uphold the authenticity of online reviews. Despite the existence of platform terms and conditions and guidance on reviews, fake reviews are still prevalent and misleading consumers. Consequently, there is a strong rationale for government intervention to set clear rules concerning the trading of fake reviews and the obligations on those who host reviews to take adequate steps to ensure they are not misleading consumers.

The proposals have four key objectives:

1. Reduce the prevalence of fake online reviews and the misleading presentation of reviews.
2. Reduce the consumer harm arising from fake online reviews and the misleading presentation of reviews.
3. More efficient enforcement against those that facilitate the trading of fake online reviews.
4. Increase moderation efforts from review hosting businesses to ensure reviews are genuine and not presented in a way that can mislead consumers.

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

Do-nothing – This option takes no steps to prevent the trading of fake reviews or to provide obligations for businesses who host reviews to ensure they are genuine.


Option 1 - This option uses legislation to outline reasonable and proportionate steps businesses who host reviews must take to ensure reviews are genuine.

Option 2 - This option uses guidance alone to outline reasonable and proportionate steps businesses who host reviews should take to ensure reviews are genuine.

Option 3 (preferred option) - This option uses legislation supplemented by dedicated guidance to outline reasonable and proportionate steps businesses who host reviews should take to ensure reviews are genuine.

Will the policy be reviewed? It N/A be reviewed. If applicable, set review date: N/A				
Is this measure likely to impact on international trade and investment?		No		
Are any of these organisations in scope?	Micro Yes	Small Yes	Medium Yes	Large Yes
What is the CO ₂ equivalent change in greenhouse gas emissions? (Million tonnes CO ₂ equivalent)		Traded: N/A		Non-traded: N/A

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 the responsible Minister:  Date: 20/09/2023

Summary: Analysis & Evidence

Policy Option 1

Description: This option bans practices related to the trading of fake reviews. This option also uses legislation to outline reasonable and proportionate steps businesses who host reviews must take to ensure reviews are genuine.

FULL ECONOMIC ASSESSMENT

Price Base Year 2019	PV Base Year 2020	Time Period Years 10 y	Net Benefit (Present Value (PV)) (£m)		
			Low: - 3, 360	High: 8,809	Best Estimate: 2,265

COSTS (£m)	Total Transition (Constant Price) Years	Average Annual (excl. Transition) (Constant Price)	Total Cost (Present Value)
Low	1.8	102.1	866.8
High	3.0	552.1	4,681.3
Best Estimate	2.4	285.6	2,422.5

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

- Small one-off familiarisation costs (~£2m) for businesses who use or host online reviews
- Moderate automated review moderation costs (~£3m 10y PV) to businesses following legislative obligations to ensure reviews are genuine and not misleading
- Human review moderation costs (~£387m 10y PV) to business
- Foregone revenue costs (~£2,030m 10y PV) to businesses arising following a reduction in the prevalence of fake reviews

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

- Minor additional enforcement costs to following more cases related to fake reviews

BENEFITS (£m)	Total Transition (Constant Price) Years	Average Annual (excl. Transition) (Constant Price)	Total Benefit (Present Value)
Low	0	156.3	1,321.4
High	0	1,142.4	9,675.5
Best Estimate	0	553.8	4,687.0

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

- Consumer benefit (~£2,030m 10y PV) through forgoing overspending on products and services with fake reviews
- Consumer benefit (~£2,657m 10y PV) through forgoing the costs of issue resolution following expectations not being met

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

- Large amount of spending redirected towards fair practicing traders
- Improved competition in the digital landscape following improvements in review authenticity
- Improved consumer trust and satisfaction
- Business savings from reduced issue resolution costs

Key assumptions/sensitivities/risks

Discount rate (%)

3.5

- Consumer spending in relevant sectors
- The yearly number of reviews posted on the largest review hosting platforms
- The prevalence of fake reviews and their impact on consumer decision making
- The proportion of detriment arising from fake reviews
- Algorithmic and human moderation costs to businesses
- The reduction in the prevalence of fake reviews following policy intervention

BUSINESS ASSESSMENT (Option 1)

Direct impact on business (Equivalent Annual) £m:			Score for Business Impact Target (qualifying provisions only) £m:
Costs: 69.1	Benefits: 0	Net: - 69.1	
			-345.7

Summary: Analysis & Evidence

Policy Option 2

Description: This option bans practices related to the trading of fake reviews. This option also uses guidance alone to outline reasonable and proportionate steps businesses who host reviews must take to ensure reviews are genuine.

FULL ECONOMIC ASSESSMENT

Price Base Year 2019	PV Base Year 2020	Time Period Years 10	Net Benefit (Present Value (PV)) (£m)		
			Low: -1,382	High: 4,207	Best Estimate: 1,199

COSTS (£m)	Total Transition (Constant Price) Years	Average Annual (excl. Transition) (Constant Price)	Total Cost (Present Value)
Low	1.8	34.6	294.3
High	3.0	235.4	1,996.8
Best Estimate	2.4	115.7	982.0

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

- Small one-off familiarisation costs (~£2m) for businesses who use or host online reviews
- Small automated review moderation costs (~£0.2m 10y PV) to businesses following guidance to ensure reviews are genuine and not misleading
- Moderate human review moderation costs (~£35m 10y PV) to business
- Foregone revenue costs (~£945m 10y PV) to businesses arising following a reduction in the prevalence of fake reviews

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

- Minor additional enforcement costs to following more cases related to fake reviews

BENEFITS (£m)	Total Transition (Constant Price) Years	Average Annual (excl. Transition) (Constant Price)	Total Benefit (Present Value)
Low	0	72.7	614.8
High	0	531.5	4,501.6
Best Estimate	0	257.7	2,180.7

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

- Consumer benefit (~£945m 10y PV) through forgoing overspending on products and services with fake reviews
- Consumer benefit (~£1,236m 10y PV) through forgoing the costs of issue resolution following expectations not being met

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

- Large amount of spending redirected towards fair practicing traders
- Improved competition in the digital landscape following improvements in review authenticity
- Improved consumer trust and satisfaction
- Business savings from reduced issue resolution costs

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

- Consumer spending in relevant sectors
- The yearly number of reviews posted on the largest review hosting platforms
- The prevalence of fake reviews and their impact on consumer decision making
- The proportion of detriment arising from fake reviews
- Algorithmic and human moderation costs to businesses
- The reduction in the prevalence of fake reviews following policy intervention

BUSINESS ASSESSMENT (Option 2)

Direct impact on business (Equivalent Annual) £m:			Score for Business Impact Target (qualifying provisions only) £m:
Costs: 15.3	Benefits: 0	Net: -15.3	
			-76.5

Summary: Analysis & Evidence

Policy Option 3

Description: This option bans practices related to the trading of fake reviews. This option also uses legislation and guidance alone to outline reasonable and proportionate steps businesses who host reviews must take to ensure reviews are genuine.

FULL ECONOMIC ASSESSMENT

Price Base Year 2019	PV Base Year 2020	Time Period Years 10	Net Benefit (Present Value (PV)) (£m)		
			Low: -2,568	High: 6,965.9	Best Estimate: 1,837.7
COSTS (£m)	Total Transition (Constant Price) Years		Average Annual (excl. Transition) (Constant Price)	Total Cost (Present Value)	
Low	1.8		75.1	637.7	
High	3.0		425.3	3,606.5	
Best Estimate	2.4		217.6	1,845.7	
Description and scale of key monetised costs by 'main affected groups'					
<ul style="list-style-type: none"> - Small one-off familiarisation costs (~£2m) for businesses who use or host online reviews - Small automated review moderation costs (~£2m 10y PV) to businesses following legislative obligations to ensure reviews are genuine and not misleading - Human review moderation costs (~£246m 10y PV) to business - Foregone revenue costs (~£1,596m 10y PV) to businesses arising following a reduction in the prevalence of fake reviews 					
Other key non-monetised costs by 'main affected groups'					
<ul style="list-style-type: none"> - Minor additional enforcement costs to following more cases related to fake reviews 					
BENEFITS (£m)	Total Transition (Constant Price) Years		Average Annual (excl. Transition) (Constant Price)	Total Benefit (Present Value)	
Low	0		122.8	1,038.4	
High	0		897.8	7,603.6	
Best Estimate	0		435.2	3,683.4	
Description and scale of key monetised benefits by 'main affected groups'					
<ul style="list-style-type: none"> - Consumer benefit (~£1,596m 10y PV) through forgoing overspending on products and services with fake reviews - Consumer benefit (~£2,088m 10y PV) through forgoing the costs of issue resolution following expectations not being met 					
Other key non-monetised benefits by 'main affected groups'					
<ul style="list-style-type: none"> - Large amount of spending redirected towards fair practicing traders - Improved competition in the digital landscape following improvements in review authenticity - Improved consumer trust and satisfaction - Business savings from reduced issue resolution costs 					
Key assumptions/sensitivities/risks				Discount rate (%)	3.5
<ul style="list-style-type: none"> - Consumer spending in relevant sectors - The yearly number of reviews posted on the largest review hosting platforms - The prevalence of fake reviews and their impact on consumer decision making - The proportion of detriment arising from fake reviews - Algorithmic and human moderation costs to businesses - The reduction in the prevalence of fake reviews following policy intervention 					

BUSINESS ASSESSMENT (Option 3)

Direct impact on business (Equivalent Annual) £m:			Score for Business Impact Target (qualifying provisions only) £m:
Costs: 47.6	Benefits: 0.0	Net: 47.6	-237.9

Table of Contents

Evidence Base	6
Background	6
UK Consumer Protection Law	7
Problem Under Consideration	8
Rationale for intervention.....	11
Rationale and evidence to justify the level of analysis used in the IA (proportionality approach)	12
Description of options considered	13
Do-nothing	14
Option 1	14
Option 2	16
Option 3	17
Policy objective.....	19
Summary and preferred option with description of implementation plan	20
Cost and benefit analysis framework.....	21
Familiarisation Cost Method	26
Monetised and non-monetised costs and benefits of each option (including administrative burden)	27
Option 1	28
Option 2	30
Option 3	32
Impacts Summary	33
Direct costs and benefits to business calculations	34
Sensitivity analysis	34
Risks and assumptions.....	36
General Assumptions.....	36
Quantitative Assumptions	37
Risk of Unintended Consequences.....	40
Impact on small and micro businesses.....	42
Wider impacts.....	43
A summary of the potential trade implications of measure	44
Monitoring and Evaluation	45
Public Sector and Equality Duty	45
Appendix 1 - Technical Annex - Inferring Additionality of Policy Intervention	48
Categorising of fake reviews	48
Summary	53
Appendix 2 – List of Questions	54

Evidence Base

Background

1. Online shopping has seen remarkable growth in recent years, markedly since Covid-19, where consumers turned to purchasing goods and services online following the introduction of social distancing measures. In 2022, UK consumers spent £224 billion in online retail markets.¹ This is almost double the reported level of £115 billion in 2015. Amazon alone reported £24.4bn of net sales in the UK in 2022.² This increase in online shopping activity has led to an increasing number of reviews left online. Many of these reviews are left on large online platforms such as Amazon and Google, and the above figures demonstrate the magnitude of consumer spending which is potentially influenced by reviews.
2. Consumer reviews influence whether consumers decide to purchase a good or service. Studies suggest consumers increasingly use reviews to make purchasing decisions with 97 per cent of adults in a survey undertaken by Which? in 2018 saying they use online customer reviews when researching a product³. This is supported by research⁴ conducted on behalf of the Department for Business and Trade (DBT) which found that consumers rank star ratings and text reviews as the second and fourth most important pieces of information, respectively, when making purchases online⁵.
3. The role of reviews in consumer decision making can be attributed to the pre-conceived assumption that reviews provide a fair and unbiased assessment of a good or service as they are expected to have been provided by a party other than the seller. For this reason, reviews differ from standard advertising or promotional material. Reviews therefore should offer a balanced form of 'social proof' on the quality of a product or service, and when consumers see that others have had a positive experience, they feel reassured they will to.
4. Whilst reviews are useful for consumers, they also impact businesses. Positive reviews can boost sales not only through 'social proof', but through better visibility online which increases their exposure to prospective customers. Many platforms use algorithms which consider a trader's reviews when determining ranking, assessing traders with a higher volume of positive reviews as more trustworthy and relevant, leading to better search visibility on the platform. Furthermore, whilst positive reviews can build a business's trust and credibility, negative reviews still provide valuable market feedback on how a business can enhance its offering.
5. Reviews are usually descriptive in nature but frequently also feature a system of quantification, such as a 'star rating' system where reviewers can rate the quality of a product on a scale (often from 1 to 5). This allows for products to be ranked in terms of average rating, or for reviews to be searched and sorted by product rating, for instance if a customer wanted to specifically view what had been written by reviewers who had posted 1-star reviews. Further to this, average ratings may influence the prominence with which a product appears on search listings due to algorithms that are employed to promote the 'top-rated' products.

¹ Capgemini. (January 10, 2023). Online retail spending in the United Kingdom (UK) from 2011 to 2022 (in billion GBP).

² Amazon. (February 3, 2023). Amazon annual net sales in the United Kingdom (UK) from 2010 to 2022 (in million U.S. dollars).

³ <https://www.which.co.uk/news/2018/10/the-facts-about-fake-reviews/>

⁴ [Investigating the Impact and Prevalence of Fake Online Reviews \(2023\), Department for Business and Trade](#)

⁵ The reported top 5 factors driving purchasing decisions (from first to fifth) were price, star rating, product description, brand and content of reviews

6. Review systems also include features such as purchase verification badges or upvote buttons for helpful reviews. By incorporating these systems, websites empower users to access the most relevant information and make more informed decisions when making purchases online.
7. Reviews are not limited to product or services pages sold directly by platforms as there are various types of review websites catering to different industries and niches. For example, TripAdvisor, is a specialist review site focused on the travel and hospitality industry. It provides reviews and ratings for hotels, restaurants, attractions, and other travel-related services without directly selling these services. This contrasts with Amazon, an e-commerce platform hosting reviews for products it sold directly to users by third-party sellers.
8. Reviews can be hosted on several different types of review sites, such as:
 - a. Retailers (e.g., Amazon), service providers or manufacturers, who sell their own products online and whose websites are mainly platforms for marketing and sales, but which allow customers to post reviews.
 - b. Booking agents, whose websites are designed to allow consumers to book or order services online from a third party but allow customers to post reviews.
 - c. Business review websites allow users to review and rate businesses, ranging from local establishments like restaurants and shops to larger companies and corporations. They offer valuable insights into a company's reputation and customer satisfaction.
 - d. Trusted trader schemes, which enable consumers to find tradespeople and host reviews to help consumers decide which tradesperson to hire.
 - e. Specialist review sites (e.g., TripAdvisor), whose core business model is to provide a platform through which consumers can view reviews of products and services from various providers, and in some cases are used directly by retailers and service providers to host reviews of their products.
 - f. Price comparison websites, whose core business model is based around comparing prices or other characteristics of products or services from third parties, but which host reviews to facilitate comparison.
9. Whilst different types of review hosting websites have different business models and serve different purposes, ultimately, they aim to provide consumers with information to make purchase decisions with. Therefore, it is crucial that the integrity of online reviews is protected to ensure that reviews reflect an accurate depiction of authentic experiences.

UK Consumer Protection Law

10. The Digital Markets, Competition and Consumers (DMCC) Bill prohibits unfair commercial practices (as defined in the regulations). These provisions originate from the Consumer Protection from Unfair Trading Regulations (CPRs) retained from EU law and restated into UK law.
11. Specifically, it requires traders to exercise professional diligence towards consumers as well as prohibiting misleading actions or omissions and aggressive practices where these are likely to have an impact on the economic decision making of the notional 'average consumer'. These concepts are defined in the DMCC Bill. Schedule 18 of the DMCC Bill sets out 31 'banned practices', which will be unfair in all circumstances, without the need to consider their effect on consumer decision making e.g., where a trader falsely represents themselves as a consumer. The list includes bait advertising,

bait and switch, limited offers, false free offers, pressure selling and aggressive doorstop selling.

12. Even where a practice is not already specified on the list of 'banned practices' in Schedule 18, it may nevertheless be unfair under the other substantive prohibitions of the DMCC Bill. For example, where a trader commissions or incentivises another person to write or submit a misleading fake review and this is likely to affect the economic decision making of the average consumer, this could (depending on the specific facts and circumstances) amount to a misleading omission or action, and/or a contravention of the standards of professional diligence.

Problem Under Consideration

13. Sellers will often have more information about the true quality of a product or service than buyers, particularly in an online setting where buyers may only be able to assess quality through the information provided online as opposed to assessing it in person. This mismatch between the information buyers and sellers have access to is called asymmetric information. This market failure leads to socially inefficient economic outcomes as informed purchasing decisions can only be made by consumers when they have access to high quality information on whether a product or service meets their preferences. In other words, without high quality information, a consumer may purchase something which ultimately disappoints them or forgo purchasing something they would have valued more highly.
14. Genuine online reviews which offer additional information to consumers on the true quality of a good or service therefore levels the playing field between buyers and sellers in terms of information on the quality of products and services for sale. Genuine reviews therefore promote market efficiency by acting as a market-based solution to the problem of asymmetric information through providing social proof.
15. As such, genuine online reviews help to increase competition on aspects of quality and service that consumers value by reducing information asymmetries, enabling consumers to access better information on these issues, incentivising businesses to improve aspects of their service that may not otherwise have been visible to consumers at the point of purchase.⁶ Small businesses in particular may benefit from consumer reviews because they do not have large advertising budgets to promote their business. Conversely, disparaging reviews may also have a disproportionate impact on a small business for similar reasons. Therefore, it is crucial that this market operates fairly.
16. Genuine consumer reviews are made by consumers who have used a good or a service, without pressure or incentive to provide a particular perspective. This relies on reviews being an authentic reflection of a consumer's experience, however this is not always the case and reviews can be intentionally misleading. These types of review are referred to as a "fake review". A fake review is one that does not reflect an actual consumer's genuine experience of a good or service and has been left in an attempt to manipulate consumer perception or target a particular business. They are typically seen where an individual or business has posted or purchased fake reviews to promote their own product or to express negative views on products offered by competitors.
17. Even where reviews are genuine, they can still be presented in a way which misleads consumers. For example, if a trader only presents positive reviews having removed any negative reviews.
18. The digitisation of consumer reviews and the ease of posting fake reviews have created a growing 'industry' that thrives on creating and selling fake reviews,

⁶ Online Reviews and Endorsements, Report on the CMA's call for information, CMA (2015) p 15.

misleading consumers. Recent advancements in natural language processing (NLP) and the availability of chatbots has also made producing well-written fake reviews in large quantities relatively simple for those who wish to do so. This is a concern both for businesses and consumers; bona fide UK businesses are likely to suffer lost sales and reputational damage from dishonest competitors duping customers. Such activities distort the market by undermining competition and by giving an unfair advantage to traders commissioning fake reviews.

19. Research commissioned by DBT has recently provided evidence on the prevalence of fake reviews on e-commerce platforms and the impact they have on consumer decision making.⁷ The research found that:
 - **11% to 15% of reviews on popular e-commerce platforms are fake.**
 - **The prevalence of fake reviews has been relatively flat over time** suggesting that fake review activity has increased in line with growth in online shopping.
 - Consumers are **5.3% less likely** to purchase a product with **poorly written fake reviews (consumers were able to spot that these were fake)** and **3.1% more likely** to purchase a product with **well written fake reviews.**
 - The impact of fake reviews varies based on product price. **When the price exceeds £80, consumers are 9.2% more likely to purchase a product** with well-written/subtle fake reviews.
20. Additionally, research⁸ conducted in the US into review fraud on Google, Facebook, Yelp and Trip Advisor has assessed the impact of reviews on these platforms. This study yielded similar findings to the DBT research, **finding that approximately 11% of reviews on Google, 7.1% on Yelp, and 5.2% on TripAdvisor were "suspect"**. The report also suggests that 4% of all reviews are fake and that the direct influence of fake reviews on online spending is estimated to be \$152bn globally.
21. These studies demonstrate that fake reviews are prevalent across both products and experience goods on major review platforms. Furthermore, it also demonstrates that consumers, despite typically engaging with reviews, cannot spot well written fake reviews and are actively misled by them. This leads to detriment through misinformed spending decisions where consumers purchase products or services that do not meet their expectations or needs. This can result in lost money and dissatisfaction for consumers as well as the costs associated with issue resolution (for example the time spent obtaining a refund or the cost of sending a product back). **The research by DBT estimated that fake review text on products alone causes up to £312 million of consumer detriment a year**⁹. This is a conservative estimate as the true detriment caused is likely to be higher once reviews on services and the effect of inflated star ratings have been accounted for.
22. Given the capacity for reviews to be manipulated, larger platforms take steps to moderate reviews to ensure they are genuine and within platform terms and conditions. Moderation is the process of screening user-generated content, such as reviews, to ensure that it aligns with the platform's guidelines, policies, and standards. This ensures that platforms can maintain consumer trust and ensures users do not encounter harmful content. Amazon has pursued legal action against more than 90 fake review brokers and has sued over 10,000 Facebook group administrators who facilitated the sale of fake reviews¹⁰. In 2022, Amazon received 1.5 billion reviews and

⁷ Research commissioned by DBT to investigate the prevalence and impact of fake reviews
<https://www.gov.uk/government/publications/investigating-the-prevalence-and-impact-of-fake-reviews>.

⁸ The State of Online Review Fraud, An Analysis of 4 Million Reviews on Google, Facebook, Yelp and TripAdvisor, Uberall (2022)

⁹ <https://www.gov.uk/government/publications/investigating-the-prevalence-and-impact-of-fake-reviews>

¹⁰ Amazon Review Blueprint 2023 - <https://www.aboutamazon.eu/news/policy/a-blueprint-for-private-and-public-sector-partnership-to-stop-fake-reviews>

ratings from customers, of these 200 million suspected fake reviews were proactively blocked. Trustpilot claims to have removed 2.7 million reviews in 2021, making up 5.8 per cent of all reviews received¹¹.

23. Moderation by large platforms typically involves a team of moderators and/or an automated system that reviews each submitted piece of content before it is publicly displayed. Moderators assess whether the content is genuine, relevant, and complies with the platform's rules. They look for potential instances of fake or misleading reviews, offensive language, spam, or any other content that violates the platform's terms of service. Effective review moderation helps create a reliable and valuable resource for consumers and contributes to the platform's reputation for transparency and authenticity. Content moderation practices can vary widely between websites based on their access to technical capabilities and varying levels of stringency.
24. For large platforms, evidence suggests that a balanced approach that combines both automated and manual processes is advantageous for review management. Automation offers efficiency, scalability, and real-time moderation, especially for high volumes of content. Automated systems can ensure accuracy and consistency in adhering to predefined rules and platform policies, making them cost-effective and capable of handling straightforward content. However, manual moderation is crucial for nuanced and context-dependent decisions, addressing complex review cases and mitigating potential biases.
25. Moderation can be performed manually by human moderators or through automated systems that use algorithms and machine learning techniques on user and review data which can flag potentially misleading content. Existing evidence suggests that websites may use a combination of both approaches, with human moderators reviewing flagged content and making the final decision on 'edge' cases.
26. A study investigating fake reviews purchased from review brokers on Facebook in real time found that whilst Amazon ultimately delete one in two reviews, there was an average lag of 100 days between posting and removal¹². This lag allowed the fake reviews to provide a boost in sales rank to the product so that the firms purchasing the fake reviews gained an unfair advantage for a significant period.
27. Given that studies suggest 20 to 30 per cent of pre-screened reviews are fake¹³, the above figures are concordant with DBT research which estimated that 12 per cent of remaining reviews on Amazon are fake.
28. The varying estimates of fake review prevalence estimated across platforms suggest that there are differing levels of fake review tolerance across review hosting businesses. Prevalence was estimated to range from 25 to 35 per cent across eight large e-commerce platforms in the UK by DBT research. Government acknowledges the current moderation efforts of certain platforms and that past a certain point, diminishing returns to review moderation exist due to the complex nature of the problem. However, studies have found that review screening activities on platforms may ramp up and down sporadically. Review 'purge' events have been observed on Amazon where the flow of review deletions increases above the normal rate¹⁴. This inconsistency seen in fake review detection suggests that even for a platform which undertakes significant screening activity, there is further capacity to make detection efforts more consistent without having to introduce new systems.

¹¹ Trustpilot Transparency Report 2022

¹² Sherry He, Brett Hollenbeck, Davide Proserpio (2022) The Market for Fake Reviews. *Marketing Science* 41(5):896-921. <https://doi.org/10.1287/mksc.2022.1353>

¹³ Martínez Otero, J.M. Fake reviews on online platforms: perspectives from the US, UK and EU legislations. *SN Soc Sci* 1, 181 (2021). <https://doi.org/10.1007/s43545-021-00193-8>

¹⁴ Sherry He, Brett Hollenbeck, Davide Proserpio (2022) The Market for Fake Reviews - Appendix. *Marketing Science* 41(5):896-921. <https://doi.org/10.1287/mksc.2022.1353>

29. Smaller businesses who host online reviews have less resource and in-house capability to implement advanced review screening systems. However, whilst these businesses undertake less moderation activity, potentially misleading reviews hosted by these businesses also present far less risk to consumers given their much smaller visibility and user base leads to far less exposure. Furthermore, there may be stronger incentives to manipulate reviews on large platforms where lots of sellers of similar products or services are competing for business within the platform (with reviews significantly impacting sales). As a result, and for the analysis in this Impact Assessment (IA), there is more focus on major review sites or platforms where consumers are most active. This is reflected in the policy proposals, where smaller review hosting businesses will only be expected to take very simple steps to ensure reviews are genuine.

Rationale for intervention

30. Despite existing guidelines which make it clear that reviews should only reflect genuine experiences of goods and services, as well as the current moderation efforts of platforms and websites, fake online reviews and manipulation of the presentation of reviews is still prevalent. Given the persistent prevalence of fake online reviews over time, and manner in which they distort consumer decisions, there is a strong rationale for government intervention to make it clear that such practices are outlawed.
31. Genuine reviews help to alleviate mismatches in the level of information held by traders and consumers, as a result, reviews lead to more efficient markets through the information they reveal to consumers who can then make more informed decisions. Genuine reviews therefore relieve the market failure of asymmetric information.
32. Conversely, fake reviews worsen the discrepancy between the information available to consumers and traders and subsequently leads to less efficient markets as consumers make misinformed decisions. As a result, fake reviews worsen the market failure of asymmetric information. Under this market failure, consumer purchase decisions are based on misleading information and consequently do not reflect their preferences. Subsequently, consumer detriment arises where consumers purchase goods or services which do not meet their expectations.
33. It is reasonable to assume that the perverse incentives arising from the benefits to businesses who use misleading review content currently outweigh the potential costs of doing so. Furthermore, projected increases in online shopping activity and the development of chatbots which drastically reduce the time and resource needed to produce fake reviews are likely to further fuel the perverse business incentives to employ such practices.
34. Given that large platforms already take steps to moderate online reviews, the free market goes some way in creating incentives which partially internalise the market failure. These incentives likely concern creating the trust and long-term consumer engagement needed to operate a profitable online business. However, the inconsistency in prevalence seen between large review hosting platforms and the irregular nature in which screening activity is conducted demonstrates the need for clear rules to bring moderation standards up to a more socially efficient level. The persistent prevalence of fake reviews over time suggests that without government intervention, the free market will likely not go further to address fake reviews despite the review screening technologies and procedures already existing.
35. Further to this, DBT research has indicated that consumers cannot spot well-written fake reviews, and that the best predictors of review fakery are not textual or metadata features, but the characteristics of the reviewers themselves. Therefore, given platform operators hold administrative access to reviewer data they are better placed than consumers to identify fake reviews.

36. Considering the above, there is a strong rationale for government to introduce a legal framework to address fake reviews which:
 - a. Increases the cost of facilitating the buying and selling of fake reviews.
 - b. Provide businesses who host reviews with further incentives to moderate reviews in a manner proportionate to their available resource and the level of risk of review manipulation on the website presents.
37. By consulting with stakeholders on adding certain practices to Schedule 18 - and other specific harmful practices - to the list of automatically unfair 'banned practices', government will explore potential solutions to help promote review authenticity, facilitate more efficient enforcement by regulators and send a clear signal to traders.

Rationale and evidence to justify the level of analysis used in the IA (proportionality approach)

38. This Impact Assessment (IA) assesses the impact of options being publicly consulted on to address fake online reviews and misleading review practices in the 'Improving Consumer Transparency' consultation.
39. The analysis assesses the impacts a reduction in the prevalence of fake online reviews will have on consumers, as well as the impacts of further requirements to moderate reviews placed on businesses.
40. Following the Reforming Competition and Consumer Policy (2020) consultation, government has gathered further evidence to understand the problem and the impacts of policy options. This includes procuring primary research on the prevalence of fake online reviews and the impact they have on consumer decision making¹⁵ and engagement with industry stakeholders.
41. The estimated impact on consumers is underpinned by an online experiment with 4,800 participants conducted during the research. This experiment used a fully functioning, interactive online shopping platform that closely resembled real-world shopping experiences. It was conducted using a large sample of UK adults that were representative of ethnicity and gender. Therefore, the approach in this IA towards estimating consumer impacts is based on the best available experimental evidence.
42. The costs of automated review moderation to businesses are underpinned by available data on the market price of text moderation charged by large content moderation providers. The costs of manual moderation have been informed by available data on the human resource large platforms direct towards review moderation in combination with their automated review screening systems. There is inherent uncertainty present in these estimates given limited evidence on these practices and variations between large review hosting platforms in the resource they dedicate towards review screening. That said, government has engaged with industry stakeholders to test the assumptions and the underlying logic of the analysis. The consultation presents a further opportunity to test and strengthen assumptions with stakeholders.
43. The scope of the analysis focuses on the largest review hosting platforms only. This IA provides evidence on why this approach is justified given the concentrated market share seen in markets using online reviews. This also ensures the assumptions made to facilitate the assessment of impacts are based on the highest quality evidence available. Further to this, any provisions set out in this IA requiring businesses to take steps to moderate reviews are designed in a manner which accounts for businesses of different sizes and capabilities. Therefore, any obligations on smaller review hosting businesses will be relatively straightforward and non-technical. As a result, a small and

¹⁵ <https://www.gov.uk/government/publications/investigating-the-prevalence-and-impact-of-fake-reviews>

micro sized business exemption is not planned at this stage as a disproportionate amount of burden is not expected on this group. This is reflected in the analysis, which only quantifies the impact of the activities of large review hosting platforms, given this is where most activity is expected to arise if the proposals are implemented, and hence the largest impacts.

44. Although the extent to which fake reviews will be removed or prevented following policy action is inherently uncertain, the best available evidence has been used to infer additionality through using surveys on business compliance as well as a dataset of 'known truth' intercepted fake reviews. Therefore, assumptions concerning the additionality of each option (in terms of the extent to which they will reduce the prevalence of fake reviews) is based on the features of true fake reviews.
45. In light of the above, the evidence presented and the data gathered is appropriate for a consultation stage IA. Furthermore, wide sensitivity ranges have been used where inherent uncertainty is present to avoid conveying false precision. The areas where evidence is weaker are clearly flagged in the Assumptions log presented in the 'Risks and assumptions' section of this document where RAG ratings are presented for the impact and quality of each key assumption. Further to this, key assumptions are being tested with stakeholders during the consultation.

Description of options considered

46. Government has publicly committed to taking steps to tackle fake online reviews. Therefore, this impact assessment considers four options where the three main options consulted on explore the specific details and balance of legislation and guidance which may be used:
 - a. **Do-nothing** – This acts as the counterfactual.
 - b. **Option 1** – Outline reasonable and proportionate steps in Schedule 18 of the DMCC Bill list of automatically unfair practices.
 - c. **Option 2** – Outline reasonable and proportionate steps in guidance.
 - d. **Option 3** – Outline that traders will be required to take reasonable and proportionate steps to remove and prevent consumers encountering fake reviews; and prevent any other information presented on the platform that is determined or influenced or determined by reviews from being false or in any way being capable of misleading consumers. Use supplementary guidance to detail what the reasonable and proportionate steps should be. (**Preferred option**)
47. Options 1, 2 and 3 all add the following practices to Schedule 18 of the DMCC Bill:
 - a. Submitting a fake review, or commissioning or incentivising any person to submit a fake review.
 - b. Offering or advertising to submit, commission, or facilitate a fake review.
48. Options to amend Schedule 18 of the DMCC Bill offer viable provisions to tackle fake reviews as it would make it clear that these practices are explicitly ban under UK consumer law. Schedule 18 could also be used to outline that businesses who host reviews are not allowed to do so unless reasonable and proportionate steps have been taken to ensure they are genuine.
49. Through setting out the rules that determine what type of review content is allowed, enforcement bodies can act against parties that are responsible for producing or proliferating misleading review content. Effective enforcement will deter the production of fake reviews by making the cost to the party responsible (for instance due to being subject to enforcement sanctions) greater than the benefit of producing the fake

reviews, including where sites that host reviews deliberately distort the selection or presentation of reviews.

Do-nothing

50. This option acts as the appraisal's counterfactual scenario. Given that this option takes no steps to legislate or use guidance to address fake online reviews and misleading review practices it is assumed there will be no impact on the prevalence of fake online reviews.
51. As a result, consumers will continue to be exposed to fake reviews and the distortive impact they have on decision making. DBT research found that fake review text on products alone causes up to £300 million pounds of consumer detriment a year¹⁶.
52. Furthermore, advancements in natural language processing technologies and the emergence of highly sophisticated chat bots have greatly reduced the time and costs of producing large amounts of well written fake reviews. The increasing availability of these technologies may drive increases in the prevalence of fake online reviews and the harm they cause to UK consumers as review brokers use them to scale up activity.
53. Content moderation on platforms is a multi-faceted and often contentious issue due to the ethical, legal, cultural and technical considerations involved. As a result, platforms are not always fully transparent on the extent to which, and how, content on their websites is moderated. That said, most platforms appear to undertake significant activity through automated and manual means to moderate content, including reviews on their platforms.
54. Despite this, whilst the free market creates some incentives for platforms to moderate reviews in the interest of upholding reputation and credibility, fake and misleading review practices are still highly prevalent. Unless there is government intervention, review hosting platforms will not have further incentives to strengthen review moderation activity. These incentives may be dampened further if fake reviews become more difficult, and therefore costly, to detect. Consequently, consumers operating in the market will continue to make purchase decisions based on misleading review information under the do-nothing scenario.

Option 1

55. Schedule 18 of the DMCC Bill outlines 31 commercial practices prohibited in all circumstances and unlawful (i.e., without any need to prove likely or actual impact on the average consumer).
56. Option 1 proposes to legislate against fake and misleading reviews by adding the following practices to Schedule 18:
 - a. Submitting a fake review, or commissioning or incentivising any person to write and/or submit a fake review of products or traders.
 - b. Offering or advertising to submit, commission or facilitate a fake review.
 - c. Misrepresenting reviews, or publishing or providing access to reviews of products and/or traders without:
 - i. taking reasonable and proportionate steps to remove and prevent consumers from encountering fake reviews; and

¹⁶ Investigating the prevalence and impact of fake online reviews. Alma Economics on behalf of Department for Business and Trade, 2023

- ii. taking reasonable and proportionate steps to prevent any other information presented on the platform that is determined or influenced by reviews from being false or in any way capable of misleading consumers.
57. In the interest of brevity, this IA often refers to provisions outlined in para. 56a and 56b collectively, as the ban on trading fake reviews. For the purposes of quantification, the impact of these measures in reducing the prevalence of fake reviews is also assessed together given a lack of available evidence on the balance of fake reviews which are commissioned and those exchanged for advertising.
58. The significance of adding these commercial practices to Schedule 18 is that carrying out these practices will be automatically unfair in all circumstances. There would be no need for the practice to be likely to cause the average consumer to take a different transactional decision for it to be unlawful.
59. As the Government is proposing to use a delegated legislative power to add to Schedule 18, these new practices will be subject to civil liability. As a result, violations of consumer law relating to fake reviews will be easier to enforce against, making it more costly for businesses who utilise these unfair practices.
60. This option takes a multi-pronged approach in addressing fake reviews through legislating against the market for fake reviews, mitigating the harm at the source, whilst placing requirements on review hosting platforms to ensure reviews are genuine and not misleading. This prevents fake and misleading reviews from culminating when some traders continue to use the practice¹⁷.
61. As discussed earlier in this IA, a large source for the prevalence of fake reviews is the emerging market of review brokers. It is expected that the banning of buying or selling fake reviews and offering or advertising to submit, commission or facilitate a fake review will lead to a reduction in demand for reviews through more efficient enforcement against those partaking in review broking as well as a deterrent effect. The deterrent effect is expected to arise from the clear signal to UK businesses that buying fake reviews is illegal. Therefore, this provision will reduce the prevalence of fake reviews through reducing those arising from the source of the problem, the review broking market.
62. Despite the provisions banning the exchange of fake online reviews, given evidence indicating that large amounts of this activity originates in other jurisdictions and is clearly non-compliant with existing platform terms and conditions, it is assumed that not all of this activity will cease following the ban. As a result, in the interest of protecting consumers, measures placing accountability on platforms to detect and remove fake reviews after they have been posted are being proposed in conjunction with the ban on buying and selling reviews.
63. As discussed earlier in this IA, large platforms are already undertaking various steps to detect and remove fake reviews to varying extents, however research indicates that fake reviews are still prevalent and screening efforts are inconsistent. Therefore, the provision set out in para. 56c seeks to set out clear expectations review hosting businesses are required to take to ensure all businesses are taking proportionate steps to ensure reviews are genuine and not presented in a misleading way. What constitutes reasonable and proportionate steps is fact specific. Policy intent is that the risk, size and capabilities of the review hosting businesses will be relevant factors. Therefore, a smaller review hosting business would be required to take simpler steps to ensure reviews are genuine relative to a large one. Consequently, most activity

¹⁷ Full compliance is not assumed given the evidence demonstrating the scale at which this practice is used.

following the reforms is expected to arise from large review hosting platforms as opposed to smaller businesses.

64. This new rule is expected to lead to a preferable and proportionate level of review moderation following the incentives the potential penalties associated with non-compliance provide, whilst recognising it is unreasonable to expect every business to screen reviews before publication. Whilst moderation may employ significant resource for larger platforms, the technologies and procedures to screen reviews already exists as platforms already undertake some moderation activity. Therefore, whilst additional costs on review hosting platforms are expected to arise if they are not already at the required standard, these costs are expected to be marginal costs from increasing activity as opposed to costs developing new screening systems altogether.
65. Further to this, small and medium sized businesses will be within the scope of this provision. However, it is important that regulation is proportionate. So given a risk-based approach is being taken to outlining reasonable and proportionate steps, these businesses will not be expected to implement advanced and expensive review screening procedures. Rather, the intention is that steps for these businesses will be to take simple steps to ensure that reviews they host are genuine and not presented in a way that could be misleading for example.
66. Review moderation by business who host reviews creates a more transparent online review system for all users. Consequently, although review moderation is not considered a public good in traditional academic terms¹⁸, the benefits to society of effective moderation against misleading content likely exceed the private costs to platforms as consumer trust builds and spending is redirected to fair practicing traders.
67. This option defines 'reasonable and proportionate steps' solely in the legislation. As a result, this option is expected to lead to the largest reduction in the prevalence of fake reviews given the unambiguous nature of the law relative to guidance and the sanctions associated with non-compliance. As a result, there will be greater incentives for platforms to undertake additional moderation activity due to the penalties associated with non-compliance.

Option 2

68. Option 2 differs to Option 1 only in its approach towards defining reasonable and proportionate steps for businesses who host reviews to take to ensure they are genuine. Option 2 will still add the following practices to Schedule 18:
 - a. Submitting, commissioning or incentivising any person to write and/or submit a fake consumer review of goods or services.
 - b. Offering or advertising to submit, commission or facilitate fake consumer reviews.
69. The expected channels and mechanisms of impact of these provisions are unchanged from the descriptions above.
70. However, this approach would use guidance alone to define the reasonable and proportionate steps businesses who host reviews are expected to take to ensure they are genuine and not presented in a misleading manner (as opposed to a definition set out in legislation).
71. This option therefore takes a more flexible approach towards outlining these steps as the guidance will be subject to businesses' self-assessment on the steps they need to take to be in accordance with the principles set out. This approach would also allow for flexibility given guidance can be updated more regularly than legislation. This offers

¹⁸ Public goods are defined as goods which are non-rivalrous and non-excludable. Platforms who host reviews may make their content excludable through, for example, age or membership restrictions.

considerable benefits given that the technological environment of reviews is rapidly evolving, driven by advancements in AI and NLP. These trends are revolutionising how reviews can be produced and moderated which means the reasonable and proportionate steps businesses can take to tackle reviews will also likely evolve.

72. Given guidance is not legally binding the cost of not following it are smaller (in comparison to legislative provisions) as there are no associated sanctions. Consequently, businesses may not be incentivised to undertake further moderation activity or may undertake less than they would have done under a legislative approach.
73. Guidance would be issued by the Competition and Markets Authority (CMA) given their role as competition and consumer law enforcer. Issuing guidance will allow the CMA to update what reasonable and proportionate steps may be as fake review practices develop over time, and tailor guidance based on the different types of businesses in scope.

Option 3

74. This option is the preferred option.

75. Again, Option 3 differs to Options 1 and 2 only in its treatment of defining reasonable and proportionate steps for businesses to take to ensure reviews are genuine. This option will add the following practices to Schedule 18:
 - a. Submitting, commissioning or incentivising any person to write and/or submit a fake consumer review of goods or services.
 - b. Offering or advertising to submit, commission or facilitate fake consumer reviews.
76. The description of the above additions to Schedule 18 and the expected policy mechanism through which they will have an effect is the same as described above.
77. This approach would use a combination of both guidance and legislation to define the reasonable and proportionate steps for businesses who host reviews to take to ensure they are genuine and not presented in a misleading manner.
78. This approach is intended to combine the benefits of the approaches taken in Options 1 and 2 whilst minimising the associated risks of each option. These risks being that a solely legislative approach may lack flexibility and introduce undue costs on businesses whilst an approach using only non-legally binding guidance risks failing to achieve the stated objectives.
79. This option would add the following to Schedule 18:
 - a. Misrepresenting reviews, or publishing or providing access to reviews of products and/or traders without:
 - i. taking reasonable and proportionate steps to remove and prevent consumers from encountering fake reviews; and
 - ii. taking reasonable and proportionate steps to prevent any other information presented on the platform that is determined or influenced by reviews from being false or in any way capable of misleading consumers.
80. Meanwhile, this option would complement the above legislative provision with supplementary guidance which would add detail and advice on the steps businesses are expected to take to ensure reviews are genuine. This will ensure that businesses can take a proportionate and risk-based approach to ensuring reviews are genuine whilst still having adequate incentives to comply with the law due to the associated penalty of non-compliance.

81. The key difference in comparison to Option 1, is that this guidance will be bespoke and dedicated to the provisions on fake and misleading reviews. This is opposed to updating existing guidance on the ban list which would cover the schedule in entirety and therefore contain less detail on fake reviews.
82. Further to this, given dedicated guidance on the provisions will be produced by the CMA, any final legislative provisions will be less prescriptive relative to Option 1 given the dedicated guidance will allow for more detail and thus offer greater flexibility to businesses.
83. Like Option 1, small and micro businesses are not exempt from the above provisions. Reasonable and proportionate steps for these businesses will be more straight forward and not require technical solutions given the lower level of risk reviews hosted by businesses of this size present. These businesses will not be expected to implement complex review screening procedures, rather there will be an expectation that steps are taken to ensure any reviews they do host are genuine and not presented in a misleading manner.
84. This option is therefore deemed to offer a satisfactory balance between the benefits to consumers of improved review authenticity whilst minimising the costs placed on business.

Other options considered

Developments from Reforming Competition and Consumer Policy Consultation

85. Following its consultation in 2021, the Government committed to consult on adding the following practices to the list of banned practices in Schedule 1 of the Consumer Protection from Unfair Trading Regulations (CPRs):
 - a. Commissioning or incentivising any person to write and/or submit a fake consumer review of goods and services.
 - b. Hosting consumer reviews without taking reasonable and proportionate steps to check they are genuine.
 - c. Offering or advertising to submit, commission or facilitate fake reviews.
86. Whilst the options proposed in this IA are similar to what was consulted on in 2021, following responses to the consultation and further engagement with industry stakeholders and the CMA, further detail has been added to the measure relating to reasonable and proportionate steps for businesses to take to ensure reviews are not misleading.
87. These developments relate to ensuring that, as well as being genuine, reviews are not presented in a manner that could mislead consumers. These practices, as being consulted on, could include:
 - a. Deleting or suppressing negative reviews;
 - b. only publishing positive reviews;
 - c. applying different weightings to reviews based on the source consumer;
 - d. publishing or providing access to incentivised reviews that are not clearly labelled as such;
 - e. disabling the consumer from changing default sorting options; and

- f. presenting reviews of a different product as relating to the product a consumer is considering (sometimes known as review hijacking, review merging, or catalogue abuse).

Alternatives to Adding Fake Review Practices to Schedule 18

88. During early policy development, it was considered whether the policy objectives could be achieved through other means than to adding practices to Schedule 18 of the DMCC Bill.
89. A bespoke chapter in the DMCC Bill on fake reviews was considered which would use its own definitions and rules relating to fake and misleading reviews. However, this option was discounted given traders are already familiar with Schedule 18 of the DMCC Bill as this has been directly retained from the CPRs which have been in place since 2008. Consequently, traders may already be familiar with the list of automatically unfair practices and would therefore find it easier to understand the implications of adding practices related to fake reviews to this schedule. Furthermore, through adding to Schedule 18, the existing definitions in the legislation can be relied upon as opposed to adding new definitions and rules that traders would have to familiarise with.
90. Considering the above, this option was discounted from the longlist of options being consulted on.

Non-regulatory interventions

91. During DBT research into the impact of fake reviews on consumer decision making, a non-regulatory intervention and the impact of fake reviews on consumer trust were examined.
92. The non-regulatory intervention examined experimentally was informed silence. Within this intervention condition, participants were able to see a text box displayed prominently on all product-specific pages. The text stated that steps had been taken to moderate misleading content (such as misleading customer reviews) on the platform. Please note that no specific reviews were flagged in this condition. This intervention was intended to explicitly inform consumers that some reviews could be fake or misleading to see whether it helped consumers discern the fake reviews and adjust their decisions accordingly.
93. The research found that:
 - a. Informing consumers that steps have been taken to moderate misleading content on the platform **does not** impact consumer purchasing behaviour.
 - b. Exposure to fake reviews generally **does not** impact consumer trust and future behaviour. Despite being exposed to fake reviews on the online platform, consumers were not observed to adapt their future behaviour, leaving them potentially susceptible to being affected by further misinformation in the future.
94. Given that the research found that consumers cannot spot well-written fake reviews, even when directly alluded to in a non-regulatory intervention, further consumer facing non-regulatory interventions were not considered at the longlist stage. This decision was further supported by the finding that consumers do not adjust their future behaviour after being exposed to fake reviews. Therefore, given the effect of fake reviews is not mitigated through trying to assist consumers to spot them with the help of non-regulatory interventions, legislative approaches are considered at longlist stage.

Policy objective

95. The proposals have four key objectives:

- a. Reduce the prevalence of fake online reviews and the misleading presentation of reviews.
 - b. Reduce the consumer harm arising from fake online reviews and the misleading presentation of reviews.
 - c. More efficient enforcement against those that facilitate the trading of fake online reviews.
 - d. Increase moderation efforts from review hosting businesses to ensure reviews are genuine and not presented in a way that can mislead consumers.
96. The measurability of the above objectives relies on having detailed and accurate data on the prevalence of fake reviews and the moderation activities of review hosting platforms. Given the difficulty in regularly monitoring these, robust SMART objectives are not possible at this stage.
97. These data gaps hinder an accurate measurement of what is achievable and in the interest of realistic goal setting, specific targets have not been attached to the objectives at consultation stage.
98. Following further evidence gathering activities and engagement with stakeholders, the final stage IA will outline the arrangements planned to monitor and evaluate the impact of any final proposals. This will likely include the identification of viable benefits indicators that can be used to underpin smart objectives. Benefits indicators could include metrics such as the reported number of fake reviews removed by platforms or the number of enforcement cases raised against those suspected of using fake review practices.

Summary and preferred option with description of implementation plan

99. The reforms will be implemented through secondary legislation which amends Schedule 18 of the DMCC Bill. The statutory instrument amending Schedule 18 is expected to progress to royal assent concurrently alongside the DMCC Bill.
100. Through adding the practices to Schedule 18, businesses operating in the UK will have to follow the provisions to comply or risk facing the civil sanctions associated with breaches related to the trading of fake reviews and taking sufficient steps to ensure reviews are genuine and not misleading.
101. The CMA is an independent authority responsible for promoting fair competition and protecting consumers' interests in the UK. Therefore, the CMA would be expected to be the primary enforcer of these rules. Other government agencies and organisations that also play a role in enforcing specific aspects of consumer law, such as trading bodies, may also be active in enforcing the new arrangements.
102. The CMA will prepare supplementary guidance on the reasonable and proportionate steps businesses who host reviews are expected to take to ensure reviews are genuine and not presented in a misleading manner.
103. There will be a one-year transition period for businesses to prepare any arrangements needed to comply with the provisions added to Schedule 18.
104. The CMA will also be responsible for regularly reviewing and updating the supplementary guidance to maintain its relevance and effectiveness in light of ongoing trends, ensuring that any guidance on procedures and practices is refined to align with the evolving technological landscape.
105. Considering the above, the provisions are expected to take effect in the year 2025.

Cost and benefit analysis framework

106. The cost-benefit analysis (CBA) models the impact of the proposals for products and services separately. This allows bespoke assumptions to be used for each category. These assumptions include the amount of UK consumer spending in each sector, the number of online reviews posted in each sector and the estimated prevalence of online fake reviews.

107. The scope of each category is as follows:

- a) Products includes all retail expenditure by UK consumers on the largest third-party e-commerce platforms. These platforms include Amazon, eBay, Etsy and Wayfair¹⁹. It is assumed that first-party e-commerce platforms (e.g., Argos and John Lewis) are not manipulating reviews. Larger businesses have a more significant public profile which increases the reputational risk of non-compliance with consumer laws. Furthermore, large first-party platforms are under more scrutiny from regulatory bodies which would make this type of activity challenging to sustain without being detected. Furthermore, if a platform was found to be manipulating reviews, they would face a much higher reputational cost which could have a long-term impact on brand perception and profitability. This assumption is supported by existing studies on the prevalence of fake reviews which have never found evidence of review manipulation by platform operators.
- b) Services includes all expenditure for household services, recreation and culture, and restaurants and hotels. This scope was chosen as these sectors regularly service consumers and form most providers on review hosting platforms for services such as TripAdvisor and Trustpilot²⁰. Government acknowledges there are other service sectors where reviews may play a prominent role in consumer decision making, however given a lack of granular consumer expenditure data to estimate impacts with, they have not been included in the interest of not overstating the expected impacts through inflated scope.

108. For both goods and services, the cost-benefit analysis quantifies the expected impacts in the following steps:

- a. **Estimates the total spending by UK consumers each year in sectors where third party hosted reviews are common.** This includes third party retail for products and household services, recreation and culture and restaurant and hotels for experience goods. This is gathered from outturn data sources on consumer expenditure²¹.
- b. **Estimates the total amount of consumer expenditure exposed to fake reviews in each sector**, i.e., the amount of spending where consumers were exposed to review fakery during their decision to purchase a product/service and were therefore potentially misinformed. This is estimated by multiplying the total expenditure data gathered above by the estimated prevalence of fake reviews for each sector produced by previous research papers. This assumes fake reviews are uniformly spread across the e-commerce and experience good sectors.
- c. **The next step is to estimate the amount of spending which can actively be attributed to fake reviews in each sector, i.e., the causal impact of fake**

¹⁹ UK online marketplace monthly site visits as recorded by Webretailer: 407m Amazon, 298m eBay, Etsy 36m, Wayfair 24m. Using monthly visits as a proxy for market it share this demonstrates how the market is predominantly dominated by Amazon and eBay

²⁰ <https://uk.trustpilot.com/categories>

²¹ ONS Household expenditure based on COICOP classification, Table 4.1

reviews on consumer purchasing decisions. Not all consumer spending exposed to fake reviews will be directly attributable to the fake reviews themselves as various other factors feed into consumer decision making. The amount of misinformed consumer spending arising from fake reviews in the baseline do-nothing (counterfactual) scenario is estimated by multiplying the estimated amount of expenditure exposed to fake reviews (estimated in previous step) by assumptions on the direct impact fake reviews have on consumer decision making. This impact on decision making has been estimated during experiments conducted by DBT²² and Which?²³. The estimates from these studies have been used to form assumption ranges. DBT assessed the impact of fake review text alone, so forms a lower bound estimate of the impact of fake reviews on decision making. This estimated that well-written fake reviews make consumers 3.1% more likely to purchase a product. The Which study assessed the impact of inflated star rating and fake review text on consumer decision making, and therefore forms an upper bound of the assumption. This study estimated that fake reviews made consumer 12.6% more likely to buy a product.

- d. **Consumer detriment is estimated from the spend arising directly from the distortive impact of fake reviews on decision making.** The analysis assumes that not all misinformed spending arising from the distortive impact of fake reviews is true detriment. This is because the product purchased may hold some value and fulfil some of the consumer's needs. This assumption is aligned with the methodology used to estimate net detriment in the Consumer protection study 2022²⁴ (CPS 2022). It is assumed that there are **two channels of detriment to consumers**:
- i. Fake positive reviews likely imply the quality of a product is higher than it actually is to consumers. Assuming consumers are willing to pay more for products they believe to be of higher quality, where consumer decisions are swayed by fake reviews, they likely end up paying more for a product than they would have done had the reviews and associated information on quality been genuine. Research by Which?, through inferring willingness to pay, **estimated consumers to overspend £0.12 on every pound spent on products with fake reviews.**
 - ii. Ultimately, some consumers will receive a product or service they are unhappy with as they realise the true quality does not match the quality the fake reviews implied. In these instances, consumers may incur costs²⁵ seeking compensation or a resolution to their issue. Using data from CPS 2022, the average cost to consumers of the time and cost spent seeking resolution was 16% of the cost of the product on average. This assumption is based on an average of all instances of detriment recorded in the CPS 2022 and therefore accounts for instances where consumers never tried to seek compensation. **Therefore, it is assumed that £0.16 for every pound spent due to fake reviews leads to these costs.**
- e. By combining the two sources of detriment above with the spending directly attributable to fake reviews, an overall consumer detriment figure arising from fake

²² Investigating the Prevalence and Impact of fake Reviews, Department for Business and Trade, 2023

²³ The Impact of Fake Reviews on Demand and Welfare, Akesson, J. et al 2022

²⁴ Consumer Protection Study 2022, Department for Business, Energy and Industrial Strategy, 2022

²⁵ These costs include the costs of paying return shipping and time spent resolving the issue

reviews is obtained. This figure represents the current detriment faced by consumers based on the current prevalence of fake reviews (i.e., the baseline).

- f. **The next steps concern estimating the expected impact of each option, i.e., the additionality of the options.** Each policy option is assumed to lead to a reduction in the prevalence of fake reviews based on its policy approach i.e., legislation vs. guidance. Due to uncertainty relating to how and to what extent platforms moderate reviews already, it is inherently uncertain how each option will translate into increased moderation activity by platforms and subsequent reductions in the prevalence of fake reviews. The same uncertainty applies to the impact of each option on review trading activity given its tacit nature. However, based on the network features of known fakes²⁶, DBT has identified the proportion of fake reviews which would be more easily detected by review screening algorithms based on how clustered²⁷ they are (See technical appendix for detail on how each assumption was formulated). The CBA assumes the following reductions in fake review prevalence:
1. **Option 1** (legislative approach) – 77% reduction in fake reviews
 2. **Option 2** (guidance) – 36% reduction in fake reviews
 3. **Option 3** (legislative and guidance) – 60% reduction in fake reviews
- g. These estimates form scenarios in which the most fake reviews are removed in the most prescriptive option (legislation alone), and the least for the least prescriptive option (guidance alone). Option 3, a combination of options 1 and 2 represents a middle ground position between Option 1 and 2 in terms of expected impacts given it uses a legislative approach to defining reasonable and proportionate steps whilst also utilising complementary guidance which introduces more flexibility. The underlying rationale here is that guidance will allow for flexibility when businesses of different types determine how they take proportionate steps to ensure reviews are genuine. This results in less fakes being removed as, assuming full compliance, legislation alone is more likely to incentivise businesses to enhance compliance efforts compared to a combination of legislation and guidance that allows for flexibility. By relying solely on legislation, businesses face clearer and more stringent requirements, creating a stronger motivation to meet those standards. Option 3 aims to balance the benefits of enhanced review moderation efforts whilst ensuring undue costs are not placed on businesses. This balance is anticipated to result in less fake reviews being removed by businesses as it introduces less emphasis on legal compliance through the use of supplementary guidance which introduces flexibility.
- h. **The estimated reduction in fake reviews results in two impacts on consumers:**
- i. **It reduces the amount of spending exposed to fake reviews** and therefore the amount of misinformed consumer expenditure (in equal proportion to the assumed decrease) and associated detriment. The portion of consumer detriment foregone arising from overspending also represents a transfer from businesses as revenue is foregone.
 - ii. **The reduction in prevalence which is assumed to be the result of additional moderation activities introduces costs to businesses.**

²⁶ The Market for Fake Reviews, He et al, 2020

²⁷ Clustering is the degree to which a product is connected to other products by common reviews. Highly clustered products suggest suspicious and co-ordinated activity as this does not occur randomly.

These costs arise from operating automated screening processes as well as employing human moderators. The additional automated moderation costs are estimated by multiplying the expected number of fakes which are removed by a market implied moderation cost assumption (see assumptions log in technical appendix). Similarly, it is assumed that a portion of reviews flagged as suspicious by automated means are passed on to human moderators for validation. Here it is assumed that 1/5 of fakes which are to be removed undergo human moderation²⁸. The cost of human moderation is then estimated using an FTE equivalent wage assumption for a moderator in combination with an assumption on the number of reviews a single moderator can review in a year to estimate the additional FTE resource needed.

- iii. As covered earlier in this IA, evidence suggests that platforms already undertake steps to moderate reviews. Therefore, the analysis assumes that any additional algorithmic moderation costs are marginal costs from increasing activity if appropriate, as opposed to the sunk costs of developing or implementing new moderation systems. These cost assumptions are derived from the price of available market provided text moderation solutions and therefore have been deemed an appropriate proxy for the marginal costs of operating an algorithmic review screening system. The consultation tests these assumptions with stakeholders.
- i. **The CBA separately estimates a familiarisation cost to business** using a reading time assumption²⁹ multiplied by the total population of UK businesses reporting to make online sales or within one of the service sectors listed above³⁰.
- j. **The above quantified impacts are aggregated and discounted to arrive at Societal Net Present Value (SNPV).**
- k. **The Equivalised Annual Net Direct Cost to Business is estimated from the direct costs to business.** The direct costs to business include the familiarisation costs, additional moderation costs and the impact on business profits³¹ following foregone spending on products and services with fake reviews.
- l. The 'Risks and assumptions' section contains further detail on the necessary assumptions made and the underlying evidence which they are formed on.
- m. Figure 1 below illustrates the CBA model logic.

Questions:

- 1) **Do you agree that the costs and benefits quantified are the correct ones for the policy impact assessment? Are there any key impacts which have been overlooked?**

²⁸ This is an indicative assumption to reflect that it is logistically impractical for human moderators to screen all reviews.

²⁹ This is based on [guidance](#) produced by the Office of Fair Trading on the Schedule 1 of the CPRs

³⁰ Roughly 170,000 businesses of various sizes have been assumed to be in scope. This is likely an overestimate of businesses who will formally familiarise themselves with the reforms, however this approach has been taken to ensure costs are not underestimated.

³¹ It is assumed that 10% of the foregone revenue following intervention results in a reduction in business profits – Damodaran, A. – New York University – Stern School of Business (2022): US Profit margins by Sector 2022 (https://pages.stern.nyu.edu/~adamodar/New_Home_Page/datafile/margin.html). This data set compiles data from many sources, including Morningstar, Capital IQ and Compustat.

Figure 1 - Cost-Benefit Analysis Methodology Model

Inputs

Total annual household expenditure in relevant sectors (£)

Assumed sectoral prevalence of fake reviews (%)

Annual consumer spending exposed to fake reviews (£)

Distortive impact of fake reviews on consumer decision making (%)

Baseline/Counterfactual

Annual misinformed consumer spending (£)

Options/Additionality

Yearly no. of reviews posted

Options analysis: expected reduction in fake review prevalence (%)

Reduction in annual misinformed spending (£)

Algorithmic cost per review assumption (£/review)

Yearly no. of fake reviews removed

Human moderation FTE cost (£/review)

Detriment from overpaying (12% of spend)

Detriment from issue resolution (16% of spend)

Automated moderation cost (£)

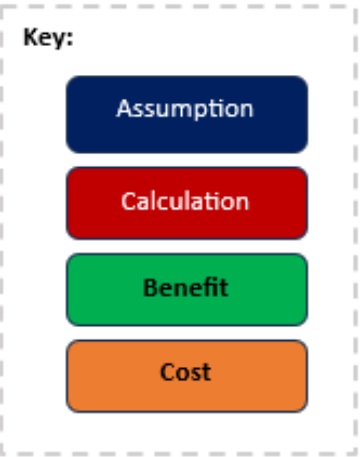
Human moderation cost (£)

Reduction in business revenue (£)

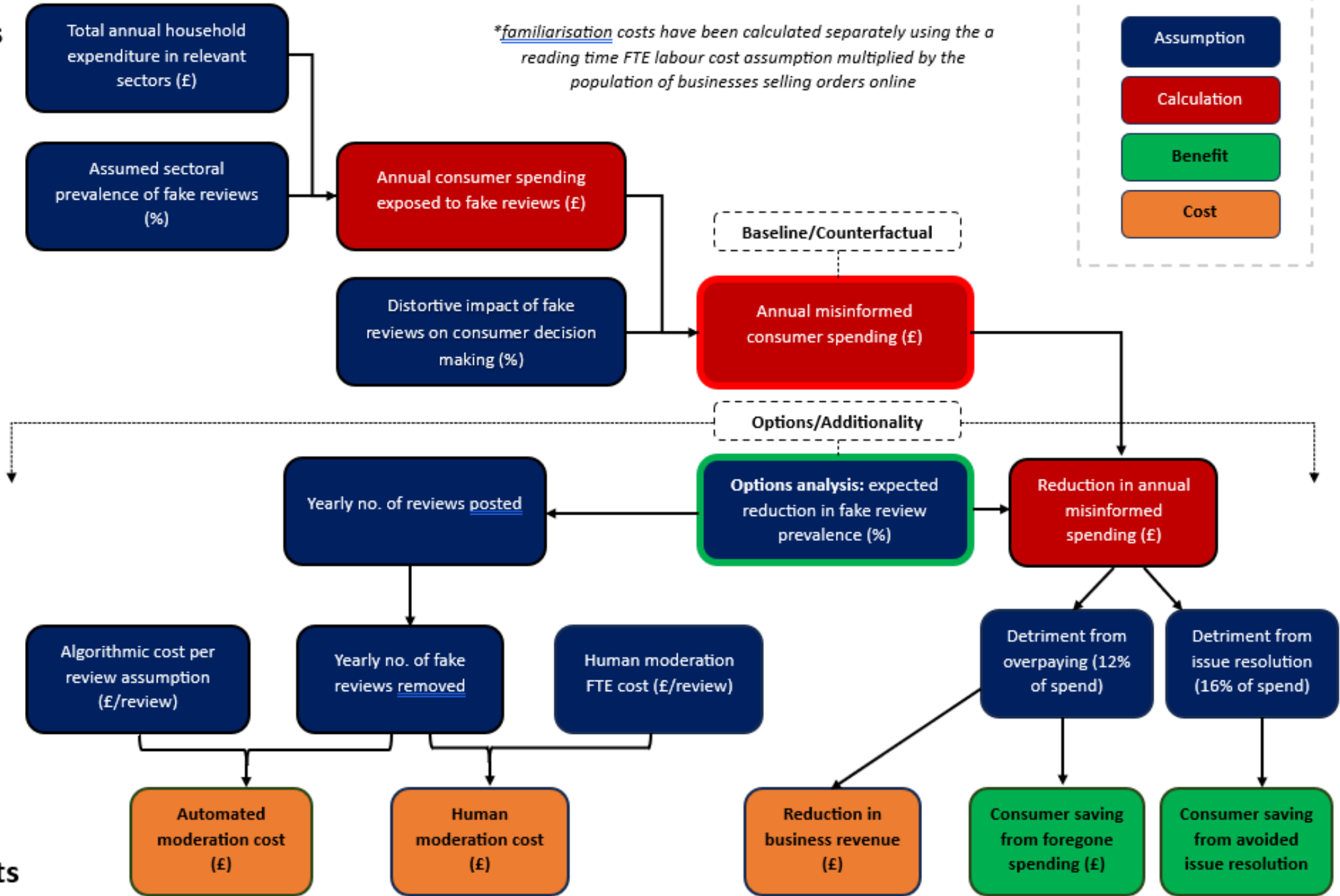
Consumer saving from foregone spending (£)

Consumer saving from avoided issue resolution

**familiarisation costs have been calculated separately using the a reading time FTE labour cost assumption multiplied by the population of businesses selling orders online*



Outputs



Familiarisation Cost Method

109. Familiarisation costs are estimated separately from other impacts using a bespoke methodology; therefore, a breakdown of the method is provided below.

$$\text{Familiarisation cost} = \text{Reading time} \times \text{Labour cost of relevant professional} \times \text{Business population}$$

Reading time

110. It is assumed that reading time is the main activity arising from businesses need to familiarise with the reforms. Given the scope for interpretability, it is assumed that familiarisation costs will predominantly arise from the measures requiring businesses hosting reviews to take reasonable and proportionate steps to ensure reviews are genuine and not misleading.
111. It is assumed that the bans on trading fake reviews require less familiarisation resource given these policies send a clear unambiguous signal to businesses that these activities are prohibited in all circumstances.
112. The guidance on reasonable and proportionate steps for review hosting businesses has not been produced at this stage of policy development. Therefore, this CBA uses previously produced OFT Guidance¹ on Schedule 1 of the CPRs (the EU law that has been restated into Schedule 18 as UK law) as a proxy for the length of any impending guidance.
113. Assuming guidance will be roughly around 2,500 words long, assuming a reading speed of 100 words a minute, a total reading time per person of 0.42 hours is derived.

Labour Costs

114. Using labour costs from the Annual Survey of Hours and Earnings (ASHE)², it is assumed that 'Corporate, managers and directors' and 'Database administrators and web content technicians' will be involved with familiarisation processes.
115. Varying levels of personnel involvement are assumed based on the size of the business to reflect the different availability of resource and capability between small and large businesses.
116. Furthermore, the hourly wage reported from ASHE, is uplifted by a non-wage cost factor to account for non-wage costs (e.g., pension contributions)³ a business pays.
117. Table 1 summarises the assumptions used.

¹ Consumer Protection From Unfair Trading Guidance, Office of Fair Trading (2008)

² Annual Survey of Hours and Earnings (2023), Office for National Statistics - Table 14.6a – Hourly pay excluding overtime,

³ Non-wage uplift factor of 1.32 – Derived from Eurostat data on wages and non-wage labour costs

Table 1 - Staff assumptions underpinning familiarisation costs

Business size	No. of staff that familiarise		Hourly wage (£)		Uplifted wage cost rates (£)	
	Corporate managers and directors	Database administrators & web content technicians	Corporate managers and directors	Database admins & web content technicians	Corporate managers and directors	Database admins & web content technicians
Micro	1	0	24.47	16.36	32.30	21.60
Small	2	0	24.47	16.36	32.30	21.60
Medium	2	10	24.47	16.36	32.30	21.60
Large	4	20	24.47	16.36	32.30	21.60

Business Population

118. Business populations have been estimated from ONS data on UK Business Counts⁴.
119. In the interest of taking a conservative approach to cost estimation, the business population in scope includes businesses of all sizes in consumer facing sectors where reviews are common⁵.
120. Overall, roughly 170,000 businesses are assumed to familiarise themselves with the policies.
121. Combining the business population estimates with the reading time and labour cost assumptions used above, **a central familiarisation cost of £2.4m is obtained.**

Monetised and non-monetised costs and benefits of each option (including administrative burden)

122. **All impacts presented below are in 2019 price years discounted to a 2020 base year** in line with Better Regulation Guidance and HMT Green Book appraisal methodology.
123. The tables present low to high impact estimates based on the sensitivities arising from assumption ranges.
124. Please see table 2 in the 'Risks and Assumptions' section of this Impact Assessment for detail on the assumptions which underpin the figures, and their uncertainty ranges, presented below.
125. The quantified impacts to businesses below include only the largest review hosting platforms, as these platforms are expected to undertake the most activity following the introduction of reasonable and proportionate steps to ensure reviews are genuine and not misleading given this is where most online reviews are concentrated. As a result, significant costs to small and micro sized businesses are not expected.

⁴ [UK Business Counts - enterprises by industry and employment size band, Office for National Statistics \(2023\)](#)

⁵ Business sectors in scope includes: Travel agency activities, Retail sale of audio and video equipment in specialised stores, Other holiday and other short-stay accommodation (not including holiday centres and villages or youth hostels), Holiday centres and villages, Retail sale of games and toys in specialised stores, Retail sale of cosmetic and toilet articles in specialised stores, Retail sale of electrical household appliances in specialised stores, Retail sale of sporting equipment in specialised stores, Other retail sale not in stores, stalls or markets, Other retail sale in non-specialised stores, Hotels and similar accommodation, Other retail sale of new goods in specialised stores (other than by opticians or commercial art galleries), Unlicensed restaurants and cafes, Licensed restaurants, Public houses and bars. Retail sale via mail order houses or via Internet

Options Comparison

126. Table 2 contains a comparison of the impact of each option below.

Table 2 - Options comparison

	Option 1	Option 2	Option 3 (Preferred option)
Reduction in Prevalence	77%	36%	60%
SNPV (£m)	2,265	1,199	1,838
Benefits to consumers (£m)	4,687	2,181	3,683
Costs to business (£m)	2,423	982	1,846
Benefit cost ratio	1.9	2.2	2.0

127. The slight difference seen across the benefit cost ratio (BCR) is driven by the balance of fake reviews which are prevented because of the provisions banning the trading of fake reviews and those prevented following the requirement for businesses to take reasonable and proportionate steps to ensure reviews are genuine and not misleading. All options ban the trading of fake reviews. This provision does not impose additional moderation costs on businesses as these reviews are prevented at the point of sale. Consequently, Option 1, which would place the least obligations on businesses to undertake additional moderation activity also imposes proportionately less moderation costs on businesses. Hence Option 2 also has the highest BCR.

128. Despite having a slightly lower BCR than Option 2, Option 3 is the preferred option. This is because of the wider unquantified benefits improved review authenticity is expected to have following greater moderation efforts by platforms. These benefits include spending being redirected to fair trading businesses and greater competition and innovation in online commerce. Option 3 is also expected to offer greater flexibility to businesses through using dedicated guidance alongside the legislation to ensure the legislation does not place undue costs on businesses through being overly prescriptive in the obligations and detail set out within law.

Option 1

129. The expected impacts of Option 1 are presented in Table 3 below. This option defines reasonable and proportionate steps for review hosting businesses to take in Schedule 18. This option also adds the practices related to trading of fake reviews to Schedule 18, Option 1 therefore takes a purely legislative approach.

Table 3 - Option 1 Impacts

		Total NPV Impact – 10 years (£m)		
		Low	Central	High
Costs	Category			
Familiarisation Cost	One off business cost	1.8	2.4	3.0
Algorithmic Moderation	Ongoing direct business cost	1.9	2.8	3.8
Human Moderation	Ongoing direct business cost	290.7	386.9	483.0
Foregone Spending (profit)	Ongoing direct business cost (transfer)	57.2	203.0	419.2
Foregone Spending (revenue minus profit)	Ongoing indirect business cost	515.2	1,827.4	3,772.4
Benefits	Category			
Consumer savings from reduced spend	Ongoing direct benefit	572.4	2,030.5	4,191.5
Consumer savings from reduced issue resolution costs	Ongoing direct benefit	748.9	2,656.5	5,483.9
Net Business Impact		- 866.8	- 2,422.5	- 4,681.3
Net Societal Impact		454.6	2,264.5	4,994.1

130. **Option 1 is expected to offer a net benefit to society of £2,264.5m in the central scenario** over the 10-year appraisal period. This is driven by the cost savings associated with issue resolution that consumers will save following a reduction in the prevalence of fake reviews, as well as a reduction in the amount of overspending on products with fake reviews.
131. For every pound of cost placed on businesses, consumers are expected to get nearly two pounds back.
132. **Option 1 introduces a cost to business of £2,422.5m over the appraisal period in the central scenario.** This is predominantly formed of foregone consumer spending on products with fake reviews which are subsequently removed following increased moderation activity. This is estimated at £2,030.5m and represents a transfer back to consumers. Whilst this estimated loss of revenue is significant, it is only experienced by businesses who are using unfair review practices. Due to a lack of clear laws specifically concerning fake reviews, this cost is not described as arising from non-compliance and is counted fully in the interest of taking a prudent approach to understanding the potential costs to business.

133. The increased moderation activity is expected to result in costs arising from paying additional human moderators as well as running algorithmic detection systems. The additional human moderation activity is expected to result in a £386.9m cost to business over the appraisal period. Algorithmic moderation costs make up a much smaller portion of the cost impact, at £2.9m over the appraisal period. This is expected given algorithmic moderation costs have lower cost assumptions associated with them as it allows for faster and more scalable review moderation without the need of human labour. Furthermore, the relative difference in cost seems reasonable given many platforms already have these systems in place.
134. The costs to business are formed of direct and indirect costs. It is assumed that not all the foregone consumer spending is a direct cost to business as it is a revenue impact, and therefore only a portion of this represents a profit impact. This is in line with RPC guidance specifies that impacts on businesses should be expressed in terms of lost profits rather than lost revenue.⁶ Consequently, although foregone sales lead to a reduction in revenue, affected businesses still hold the stock and/or are not incurring the resource cost of providing that product or service. Therefore, whilst this cost is still accounted for to ensure a prudent approach to cost assessment is taken, only the loss of profit which would have arisen from the sale is accounted for as a direct cost. Furthermore, consumer spending will likely be redirected to products with genuine reviews when fake reviews are removed, as a result this cost likely overestimates the true cost, however evidence to form assumptions on any redirected spending is weak. Therefore, this benefit is assessed qualitatively in the 'Wider impacts' section.
135. The profit impact is estimated to be £203m, the rest of the reduction in spending, estimated at £1,827.4 is classed as an indirect cost. This is the only quantified indirect cost to business, all the other costs are classified as direct.
136. The proposals are expected to introduce a one-off £2.4m familiarisation cost to business.
137. The analysis uses sensitivity ranges in the underlying assumptions to convey the inherent uncertainty present in the analysis. This results in an estimated SNPV of £456.4m to £4,994.1m for Option 1. This sensitivity range is predominantly driven by the assumed range of the impact fake reviews have on consumer decision making. Empirical studies have shown that fake reviews have varying impacts on a consumer's likelihood to make a purchase based on features such as the price of the product and whether fake reviews include inflated star ratings. The assumed ranges reflect these different scenarios based on the findings of these studies. The assumed ranges in the prevalence of fake reviews also partially determine the range seen in the estimated SNPV, although to a lesser extent than the assumed impact of fake reviews on consumer decision making.

Option 2

138. Option 2, like Option 1, adds the practices related to trading fake reviews to Schedule 18. This option takes an alternative option to define what reasonable and proportionate steps mean through using a code of conduct or guidance for businesses. The monetised impacts of this option are set out in table 4 below.

⁶ <https://www.gov.uk/government/publications/rpc-case-histories-other-bit-methodology-issues-march-2019>

Table 4 - Option 2 Impacts

		Total NPV Impact – 10 years (£m)		
		Low	Central	High
Costs	Category			
Familiarisation Cost	One Off Business Cost	1.8	2.4	3.0
Algorithmic Moderation	Ongoing direct business cost	0.2	0.2	0.3
Human Moderation	Ongoing direct business cost	26.0	34.7	43.3
Foregone Spending (profit)	Ongoing direct business cost (transfer)	26.6	94.5	195.0
Foregone Spending (revenue minus profit)	Ongoing indirect business cost (transfer)	239.7	850.2	1,755.1
Benefits	Category			
Consumer savings from reduced spend	Ongoing direct benefit	266.3	944.7	1,950.2
Consumer savings from reduced issue resolution costs	Ongoing direct benefit	348.5	1,236.0	2,551.5
Net Business Impact		- 294.3	- 982.0	- 1,996.8
Net Societal Impact		320.4	1,198.7	2,504.9

139. **Option 2 is expected to offer a net benefit to society of £1,198.7m in the central scenario** over the 10-year appraisal period. Similarly, to Option 1, this benefit is driven by the consumer benefits associated with a reduction in the prevalence of fake reviews. Although this is to a lesser extent than Option 1, given the assumed reduction of fake review prevalence is lower as this option uses non-binding guidance to define reasonable and proportionate steps for businesses who host reviews to ensure they are genuine. As a result, the impacts predominantly arise from the provisions banning the trading of fake reviews.

140. For every pound of cost placed on businesses, consumers are expected to get two pounds back.

141. **Option 2 introduces a cost to business of £982.0m over the appraisal period in the central scenario.** This is predominantly formed of foregone consumer spending on products using fake reviews, albeit at a lower level than Option 1 as it is anticipated fewer fake reviews will be removed by businesses. This represents a transfer back to consumers.

142. Similarly, to Option 1, additional moderation costs are expected, albeit at much lower levels given it is expected that this option will lead to less moderation activity than under Option 1. This is because the reduction in fake review prevalence is driven by the ban on trading reviews, not additional moderation activity. For Option 2, additional human moderation and algorithmic moderation costs are estimated at £34.7m and £0.2m respectively over the appraisal period.

143. The costs to business are formed of direct and indirect costs. The direct profit impact is expected to be £94.5m, the rest of the reduction in spending, estimated at £850.2 is classed as an indirect cost.

144. Option 2 is expected to introduce a one-off £2.4m familiarisation cost to business in the central scenario.

145. The SNPV for Option 2 ranges from £320.4m to £2,504.9m in the low and high impact scenarios. Like Option 1, this sensitivity range is predominantly driven by the assumed range of the impact fake reviews have on consumer decision making.

Option 3

146. Option 3 is the preferred option. Like Options 1 and 2, Option 3 adds practices related to the trading of fake reviews to Schedule 18. However, this option uses both legislation and guidance to tackle fake reviews through outlining reasonable and proportionate steps for review hosting businesses to take to ensure reviews are genuine and not misleading to consumers. The estimated impacts for the preferred option are presented in Table 5 below.

Table 5 - Option 3 Impacts

		Total NPV Impact – 10 years (£m)		
		Low	Central	High
Costs	Category			
Familiarisation Cost	One Off Business Cost	1.8	2.4	3.0
Algorithmic Moderation	Ongoing direct business cost	1.3	1.8	2.7
Human Moderation	Ongoing direct business cost	184.7	245.8	306.9
Foregone Spending (profit)	Ongoing direct business cost (transfer)	45.0	159.6	329.4
Foregone Spending (revenue minus profit)	Ongoing indirect business cost (transfer)	404.9	1,436.1	2,964.6

Benefits	Category			
Consumer savings from reduced spend	Ongoing direct benefit	449.9	1,595.7	3,294.0
Consumer savings from reduced issue resolution costs	Ongoing direct benefit	588.6	2,087.7	4,309.6
Net Business Impact		- 637.7	- 1,845.7	- 3,606.5
Net Societal Impact		400.7	1,837.7	3,997.1

147. **Option 3 is expected to offer a net benefit to society of £1,837.7m in the central scenario** over the 10-year appraisal period. Similarly, to the prior two options, this benefit is driven by the consumer benefits associated with a reduction in the prevalence of fake reviews. Given this option is expected to lead to a greater reduction in the prevalence of fake reviews than Option 2, but less than Option 3, it has the second highest SNPV of the three options.
148. For every pound of cost placed on businesses, consumers are expected to get two pounds back.
149. **Option 3 introduces a cost to business of £1,845.7m over the appraisal period in the central scenario.** This is predominantly formed of foregone consumer spending on products using fake reviews which is estimated at a £1,595.7m cost to business. This represents a transfer back to consumers.
150. This option is also expected to lead to additional moderation costs. Given the guidance on reasonable and proportionate steps for businesses to take will be set out in conjunction with legislation it will be legally binding. Therefore, like Option 1, this option is expected to lead to a significant increase in review moderation by platforms, although to a lesser extent. For Option 3, additional human moderation and algorithmic moderation costs are estimated at £245.8m and £1.8m respectively over the appraisal period.
151. The costs to business are formed of direct and indirect costs. The direct profit impact is expected to be £159.6m, the rest of the foregone revenue, estimated at £1,436.1 is classed as an indirect cost.
152. Option 3 is expected to introduce a one-off £2.4m familiarisation cost to business in the central scenario.
153. The estimated SNPV for Option 3 ranges from £320.4m to £2,504.9m in the low and high impact scenarios. Like the prior two options, this sensitivity range is predominantly driven by the assumed range of the impact fake reviews have on consumer decision making. Please refer to para.108C for a description of the evidence used here.

Impacts Summary

154. The difference in expected impacts between the options is driven by the assumed reduction in fake review prevalence in each instance. Consequently, Option 1 has the highest estimated SNPV and Option 2 has the lowest estimated SNPV. The same ordering is also true for the expected costs each option will place on businesses. The

positive SNPVs are driven by the benefits to consumers following a reduction in the prevalence of fake reviews.

155. All three options have an approximately 2:1 consumer benefit to business cost ratio. Slight differences in this ratio are driven by differences in the balance of fake reviews which are prevented by the banning of trading them versus those removed by enhanced moderation efforts by businesses.

156. Although Option 1 has the highest estimated SNPV, wider factors which remain unquantified have been considered in determining Option 3 as the preferred option. Option 3 was selected as it ensures that undue burdens are not placed on businesses through using an overly prescriptive approach to defining reasonable and proportionate steps for businesses to take to ensure reviews are genuine whilst balancing the benefits to consumers.

157. As guidance can be updated more frequently than legislation, Option 3 also ensures the proposals can be future proofed to the changing technological landscape.

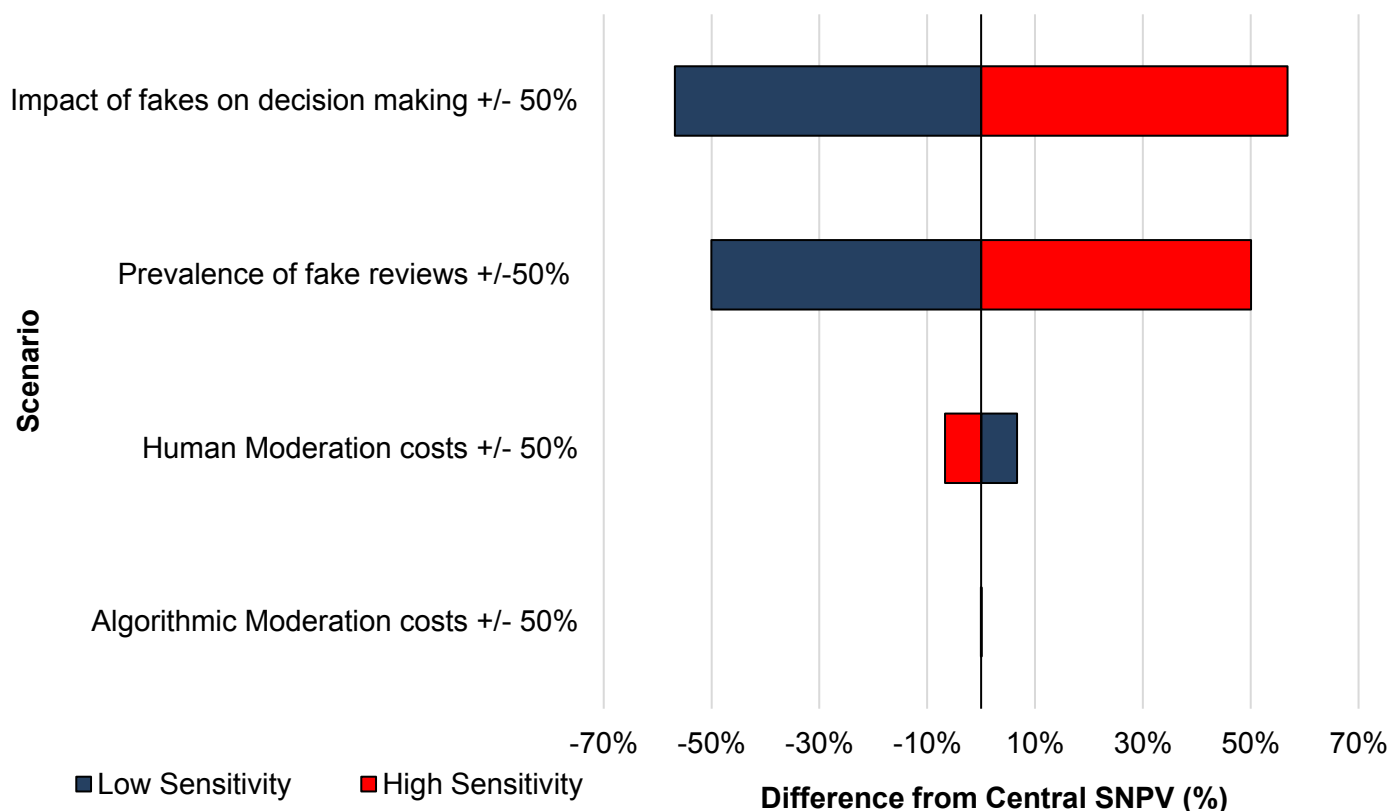
Direct costs and benefits to business calculations

158. For the preferred option, the total business NPV is estimated at -£1,845.7m (2019 prices) with a total **Equivalent Annualised Net Direct Cost to Business (EANDCB) of £47.6m** (using a 2020 base year) arising predominantly from the costs to business of additional human moderation activity and foregone profits following the removal of fake reviews.

Sensitivity analysis

159. Considering the inherent uncertainty present in the analysis, sensitivity analysis has been conducted to understand the impact varying assumptions has on the SNPV. The sensitivities tested indicate how the expected impacts would change under scenarios where current assumption levels are invalidated. This indicates the weight individual assumptions hold in the analysis as well as the assumptions which are particularly high risk.

Figure 1 - Sensitivity Analysis



160. Figure 1 shows the impact varying assumptions by 50% in either direction has on the estimated SNPV.

161. The SNPV is most sensitive to the impact of fake reviews on a consumer's likelihood to purchase a product. An increase in 50% of the initial assumption, leads to a subsequent increase of 57% in the SNPV, and vice versa for an equivalent reduction in the assumption. The impact of fake reviews on decision making is directly related to the estimated detriment caused by fake reviews as this leads to the consumer overpaying for a product as well as incurring the costs of issue resolution in some cases. As a result, if fake reviews are more effective at misleading consumers, there is a larger associated benefit with reducing the prevalence of fake reviews and the estimated SNPV will increase.

162. The second most significant assumption is the prevalence of fake reviews. Similar to varying impact of fake reviews on consumer decision making, when the number of fake reviews is 50% higher than baseline levels, the SNPV will also increase by 50%, and vice versa for an equivalent reduction in the number of reviews. The SNPV is directly proportional to the estimated prevalence of fake reviews as this determines both the level of costs placed on businesses and benefits experienced by consumers following a reduction in prevalence.

163. The analysis is less sensitive to assumptions on moderation efforts by firms to tackle fake reviews.

164. Human moderation, which is the cost of hiring labour to assess flagged reviews, does have a noticeable but small impact on the SNPV. For instance, if the cost of labour had increased by 50%, the SNPV will see a reduction of 7%, and vice versa for an equivalent decrease. This is expected as labour cost assumptions are much smaller

relative to the largest cost to business which is the foregone revenue following a reduction in the prevalence of fake reviews.

165. The SNPV is not sensitive to algorithmic moderation costs. An increase in algorithmic moderation costs of 50% results in a decline of 0.05% in the SNPV. This is expected given the relatively small magnitude of algorithmic moderation costs compared to other cost assumptions. This is expected given the scalable nature of algorithms in screening reviews given many platforms already have these systems in place.

166. Given weak existing evidence on moderation costs, government acknowledges that uncertainty is present when estimating the costs to businesses of increasing review moderation activity. Given a risk-based approach will be taken towards defining reasonable and proportionate steps, small businesses with less technical capability will not be expected to develop advanced automated moderation systems. Considering this, and the low responsiveness of the SNPV to large changes in the moderation cost assumptions, the uncertainty in the assumed levels has been deemed acceptable ahead of further evidence gathering for the final stage impact assessment.

Risks and assumptions

167. A series of assumptions have been made to enable the quantification of the expected impacts of the additions to Schedule 18 of the DMCC Bill, most notably on the impact and prevalence of fake online reviews and the costs of review moderation to businesses.

168. Given the technical nature of the issue, there is inherent uncertainty in the assumptions made. This presents the risk that the estimated impacts are not reflective of the actual impacts. However, for reasons stated in para.45 government is confident that the approach taken, and underlying evidence, has appropriately accounted for this uncertainty at this stage. Assumption ranges based on empirical evidence from primary research have been used to capture the uncertainty present. Furthermore, assumptions will be tested with stakeholders during the consultation and further evidence gathering will take place ahead of the final stage IA.

169. Sensitivity analysis has been conducted through varying key assumptions to show how changes in key variables affect the overall outcome of the analysis. This highlights the assumptions with the greatest impact as well as the best- and worst-case scenarios.

170. The area of greatest uncertainty is the costs placed on businesses arising from requirements to undertake additional review moderation activity. These assumptions will be tested with stakeholders during the consultation with any emerging evidence being incorporated into the final stage IA.

171. The following sections highlight the assumptions made to enable quantification and indicate where the highest levels of uncertainty are present.

General Assumptions

172. The analysis relies on the following general assumptions:

- i. The additionality of each option relies solely on its treatment of 'reasonable and proportionate steps to ensure reviews are genuine'.
- ii. The additional impacts on businesses of 'reasonable and proportionate' steps arises from moderating the additional fake reviews which are removed following policy intervention. The analysis does not assume that all reviews are re-evaluated (otherwise the cost would be equivalent across all options and would not provide an indicative scale of cost reflecting each options stringency). This assumes that

platforms will moderate more recent reviews, reviews from suspicious users and previously identified fringe cases.

- iii. That fake reviews are spread equally across products/services.
- iv. That the number of reviews grows in line with forecasted growth in online sales figures.
- v. The scope of the analysis includes only the largest third-party review hosting platforms.
- vi. The scope of the analysis only includes 'positive' fake reviews as opposed to 'negative' fake reviews posted to rivals. There is little available evidence on the prevalence of 'negative' fake reviews, however evidence suggests that these are rarer than positive fake reviews. The evidence suggests that it is more cost effective for traders to manipulate their own reviews than the reviews of a large market of competitors. Considering the above, and a lack of robust available evidence to assess impacts with, this scope is considered appropriate.

Quantitative Assumptions

173. Table 3 below shows the assumptions made to quantify impacts.

Table 2 - Assumptions and RAG ratings

<u>Assumption</u>	<u>Value</u>	<u>Source</u>	<u>Impact</u>	<u>Quality</u>
Annual UK review site data				
Amazon UK Gross Merchandise Volume (GMV)	£24bn	Proxied from Amazon UK net sales 2022 (converted to Pound Sterling) – <u>Amazon Annual Report 2022</u>	High	High
Ebay UK Gross Merchandise Volume (GMV)	£15bn	https://ecommercenews.eu/amazon-and-ebay-account-for-40-of-uk-ecommerce/	High	High
Etsy UK Gross Merchandise Volume (GMV)	£2.7bn	Scaled with Amazon GMV based on site visits	Low	Low
Wayfair UK Gross Merchandise Volume (GMV)	£1.8bn	Scaled with Amazon GMV based on site visits	Low	Low
E-commerce UK site visits	-	<u>Similar Web</u> UK E-Commerce site visits	Medium	Medium
Annual UK consumer expenditure on services and experience goods				
Consumer expenditure on household goods & services per annum	£57bn	<u>ONS Household expenditure based on COICOP classification</u> , Table 4.1	High	High

Consumer expenditure on recreation & culture per annum	£115bn	<u>ONS Household expenditure based on COICOP classification</u> , Table 4.1	High	High
Consumer expenditure on restaurants & hotels per annum	£81bn	<u>ONS Household expenditure based on COICOP classification</u> , Table 4.1	High	High
UK consumer expenditure growth rates				
Forecasted growth rate in UK consumer expenditure on e-commerce platforms per annum 2023 - 2035	4.1%	eCommerceDB data, eCommerce market in the UK 2022	High	Medium
Forecasted growth rate in UK consumer expenditure on services/experience goods	0.4% pa	ONS household expenditure data historic CAGR	Low	Medium
Consumer detriment assumptions				
Portion of spend influenced by fake reviews which causes detriment from overspending	12%	The Impact of Fake Reviews on Demand and Welfare, Akesson, J. et al 2022	High	Medium
Portion of spend influenced by fake reviews which causes detriment associated with issue resolution	16%	DBT internal data, based on breakdown of detriment components recorded in the Consumer Protection Study 2022	High	Medium
Review assumptions				
Yearly number of reviews on Amazon	1,500m	<u>Amazon published data</u>	High	High
Yearly number of reviews on eBay	1,100m	Derived from Amazon data by weighting on site visits	High	Low
Yearly number of reviews on Etsy	133m	Derived from Amazon data by weighting on site visits	Low	Low

Yearly number of reviews on Wayfair	89m	Derived from Amazon data by weighting on site visits	Low	Low
Yearly number of reviews on Google	926m	Google Places API sample of business ratings across the UK and scaled for UK business population	High	Medium
Yearly number of reviews on Trustpilot	47m	Number of reviews in 2021 from <u>Trustpilot Transparency Report</u>	Low	Low
Yearly number of reviews on TripAdvisor	116m	Number of reviews in 2020 published by <u>Statista Research Department</u>	Low	Low
Yearly number of reviews on Yelp	244m	<u>Number of reviews in 2021</u> , Social Media Usage in United Kingdom, S. J Dixon	Low	Low
Prevalence of fake reviews on e-commerce platforms	10 – 25%	<u>DBT research</u>	High	High
Prevalence of fake reviews for experience goods on review platforms	6 – 12%	<u>The Transparency Company research</u>	High	Medium
Moderation cost assumptions				
Algorithmic moderation cost per review	£0.0011 - £0.0014	Market implied cost derived from pricing of text moderation providers	Medium	Low
Non-wage uplift factor	1.32	<u>Derived from Eurostat data on wages and non-wage labour costs</u>		
FTE non-wage uplifted cost of human moderator per annum	£35k	Estimated from Amazon moderator salary and uprated by non-wage factor	High	Low
Reviews moderated by human per annum	38k	Indicative assumption based on qualitative data sources	High	Low
Policy option specific assumptions				
Option 1: expected reduction in fake review prevalence	77%	Assumes 31.7% reduction as a result of banning the trading of fake reviews. Derived from <u>EU Consumer Conditions</u> Data (see methodology section). Assumes reasonable and proportionate steps to ensure reviews are genuine leads to 66% reduction in remaining fakes. Derived from	High	Low

		network features of Amazon reviews from Jianmo Ni, Jiacheng Li, Julian McAuley Empirical Methods in Natural Language Processing (EMNLP), 2019		
Option 2: expected reduction in fake review prevalence	36%	Assumes 31.7% reduction as a result of banning the trading of fake reviews. Derived from <u>EU Consumer Conditions</u> Data (see methodology section). Assumes reasonable and proportionate steps to ensure reviews are genuine leads to 6% reduction in remaining fakes. Derived from network features of Amazon reviews from Jianmo Ni, Jiacheng Li, Julian McAuley Empirical Methods in Natural Language Processing (EMNLP), 2019	High	Low
Option 3: expected reduction in fake review prevalence	60%	Assumes 31.7% reduction as a result of banning the trading of fake reviews. Derived from <u>EU Consumer Conditions</u> Data (see methodology section). Assumes reasonable and proportionate steps to ensure reviews are genuine leads to 42% reduction in remaining fakes. Derived from network features of Amazon reviews from Jianmo Ni, Jiacheng Li, Julian McAuley Empirical Methods in Natural Language Processing (EMNLP), 2019	High	Low

Questions

- 2) Do you agree with the assumed algorithmic moderation cost per review? If not, please provide your reasoning and provide any supporting information you have.
- 3) Do you agree with the assumed human review moderation assumptions? If not, please provide your reasoning and provide any supporting information you have.

Risk of Unintended Consequences

174. It is considered that measures relating to the banning of buying and selling fake reviews, generally has few, if any, unintended consequences. This type of ban is implemented to protect consumers from false review information, therefore for the purposes of legality and ethics, banning these fraudulent practices is not expected to lead to significant unintended consequences. Consequently, this section focuses on the potential unintended consequences of requiring businesses who host reviews to take reasonable and proportionate steps to ensure they are genuine.
175. Increasing online review moderation will yield positive outcomes for consumers and businesses alike. However, due to the complex nature of review moderation there is a risk that policy action could lead to unintended consequences.
176. These risks have been assessed qualitatively below:

- a. **Over-moderation of reviews** – Due to the deceptive and complex nature of fake reviews, even the most advanced review moderation systems are unlikely to correctly classify whether a review is fake or not in every single instance. Some authentic reviews may appear as suspicious for unintentional reasons and subsequently be removed although they reflect someone’s genuine experience of a good or service. The misidentification of genuine reviews as fake would harm the ability of consumers to express their opinions as reviews reflecting genuine experiences would be effectively censored. Furthermore, if a significant portion of genuine reviews are removed, this will undermine the quality of review information available to users and the market failure of asymmetric information will persist.
- b. **Slower review content approval** – Placing an obligation on platforms to invest more resource into review moderation may slow down the process of review approval as review systems begin to process a greater amount of content. Some platforms currently screen reviews before they appear online, and others may only accept reviews from verified users. If platforms increase the scrutiny applied by review approval systems, or introduce new ones, it may take longer for reviews to appear online. This may lead to dissatisfaction amongst users who want their content to be visible. Furthermore, it may slow down the rate at which valuable review information is shared with prospective consumers.
- c. **Algorithmic bias** – Platforms who use algorithmic review moderation systems may be encouraged to rely more heavily on them to comply with the requirement to take reasonable and proportionate steps to ensure reviews are genuine, particularly given that evidence suggests that algorithmic systems may be more cost effective than manual means of moderation. These algorithms may be trained or fine-tuned in a way which disproportionately targets and flags certain groups or ideas in reviews which may skew the information available to consumers and lead to an uneven application of review moderation policies.
- d. **Burden on smaller platforms** – Stronger legal requirements for businesses to moderate reviews may create a cost burden which impacts their profitability. Allocating more resources to hire and train human moderators or invest in advanced automated review screening systems requires substantial financial commitments. Smaller review platforms will have less resource and technical capability to put these systems into place. As a result, if the reforms were overly stringent smaller platforms may face challenges in meeting the increasing demands of review moderation which would risk inhibiting their growth and ability to compete in the digital landscape.

177. In selecting the preferred option, careful consideration has been given to the possible unintended consequences highlighted above. While the additions to Schedule 18 aim to strengthen legal compliance and improve the quality of information available to consumers, it is essential that this is done in a proportionate manner in terms of the potential adverse consequences that may arise.

178. The preferred option uses a combination of guidance and legislation specifically to mitigate against the types of unintended consequences outlined above. Government is confident that this approach introduces a balanced policy framework to moderating fake reviews considering the flexibility enabled using guidance which will offer context-specific applications. This will accommodate a wider range of situations which cannot be appropriately defined within legislation. This will enable businesses of different sizes and review models to take a proportionate and risk-based approach towards review moderation. This will reduce the likelihood that undue costs are placed on businesses or that review screening procedures become less efficient or lead to worse outcomes.

Impact on small and micro businesses

179. At this stage, no small or micro-sized business (SMB) exemption is planned as the measures are not expected to have a disproportionate impact on SMBs given relatively simple obligations will be placed on this group. Furthermore, the preferred option utilises guidance to embed flexibility into the obligations the provisions may place on SMBs to ensure they do not face undue burden. Though the consultation will be used to gather views from stakeholders on the steps small and micro sized businesses in scope may be reasonably expected to take.
180. For the provisions concerning the banning of the trading of fake reviews, a SMB exemption is considered inappropriate as this practice is intended to mislead consumers and undermine competition. Consequently, businesses of all sizes should be held to the standard of not taking part in this clearly unfair business practice. Furthermore, most SMEs, who are not involved in trading fake reviews and are fair practicing, will be unaffected by this proposal. Therefore, there are no reasonable grounds for excluding SMBs from this measure nor is this measure expected to have a disproportionate impact on them.
181. As described earlier in this IA, the approach taken towards outlining reasonable and proportionate steps for businesses to take to ensure reviews are genuine in the preferred option has been developed with the aim of ensuring the proposal does not have a disproportionate impact on businesses of different sizes and capabilities. Through using supplementary guidance to outline reasonable and proportionate steps, a risk-based approach will be set out whereby smaller review hosting businesses will need to do less to comply relative to a large review hosting platform where there is a bigger risk associated with review manipulation. Given the provisions will directly account for the capabilities of review hosting businesses in scope, the proposal mitigates the risk that disproportionate burdens will be placed on SMBs.
182. Whilst there are smaller businesses who may host reviews on their own website, the review sector is dominated by a handful of large review hosting platforms such as Google and Amazon. For example, a local restaurant who hosts reviews on Google or Trustpilot. As a result, most online reviews are concentrated in these platforms. These platforms will also be responsible for taking reasonable and proportionate steps to ensure reviews they host are genuine.
183. As a result, for SMBs operating who host reviews on these platforms, the responsibility to ensure reviews are genuine will also be placed with the platform itself given they have access to the administrative data and technical capabilities to screen reviews at a large scale. The guidance on reasonable and proportionate steps will acknowledge the technical capability of large platforms and place further reaching obligations on them as opposed to on SMBs using the platform. As a result, whilst small businesses using these platforms may be expected to perform simple checks to ensure the reviews on them are genuine and not misleading, the platforms themselves, not the small businesses, will be responsible for
184. In the cases where a small business hosts reviews on a privately owned website, there may, subject to feedback during the consultation, be a responsibility on the small business to ensure reviews on that website are genuine and not misleading. However, given these reviews will be less visible than those on large platforms, the steps are required to take, as will be reflected in the legislation and guidance, to meet this obligation will be simpler and non-technical. The business would not be expected to implement a complex technical solution, such as implementing automated review screening software, for example. The consultation will be used to gather views from stakeholders on the steps small and micro sized businesses may be reasonably expected to take.

185. This is reflected in the scope of the cost-benefit analysis which quantifies impacts for the largest review hosting platforms only as it is anticipated that this is where most activity to ensure reviews are genuine will occur. Consequently, this is also where the significant costs to businesses are expected to arise.

186. Considering the above, a SMB exemption has not considered necessary at this stage. However, the consultation will gather views on whether and what steps are deemed proportionate for small and micro businesses to take.

Wider impacts

187. Alongside the impacts which have been quantified, wider or indirect impacts which have not been quantified are also expected as a result of the proposals in this IA. These impacts have not been quantified due to a lack of reliable evidence to enable robust quantification with. Hence, they are assessed qualitatively below:

Consumer trust

188. It is currently reported that 49% of buyers trust online reviews⁷. This indicates that there is significant room for improvement in consumer trust when it comes to their perception of online reviews.

189. It is expected that consumer trust and satisfaction will increase following a reduction in the prevalence of fake reviews as more accurate review information will make it more likely that their expectations are met following a purchase. Consequently, consumers are likely to experience higher levels of satisfaction. This is likely to lead to improvements in consumer trust when shopping online, and fostering trust is integral for businesses too, as nearly two thirds of customers are willing to pay a little more with companies they trust⁸. This demonstrates that any improvements in consumer trust and industry credibility following the reforms will have a positive impact on both businesses and consumers.

Increased competition

190. Competition between online traders is expected to increase following the removal of fake reviews. Businesses who purchase or write fake reviews gain an unfair competitive advantage over other traders who don't. As demonstrated by DBT and Which? research, well-written fake reviews directly distort consumer decision making, leading them to buy products they otherwise would not have if review information were genuine. This is because fake reviews undermine competition through product or service quality, as consumer's ability to accurately infer quality is undermined when they are exposed to fake or misleading reviews.

191. Further to this, reviews may also feed into platform algorithms and filtering systems which determine the page rank of products displayed to consumers, meaning businesses with positive reviews are more likely to be promoted or appear to consumers sooner than a business with less positive reviews. As a result, fake reviews do not just undermine competition through distorting consumer decision making, but through also influencing the visibility of businesses operating on large platforms.

192. Following the removal of fake reviews, there will be an improved "playing field" for businesses trading online, as reviews will more accurately reflect the quality of products and services. This enables consumers to make better purchase decisions following improvements in the information available to them. Consequently, with all else held equal, traders who sell higher quality products will receive feedback from reviews to reflect this and will be subsequently rewarded through improved sales. In turn, this will provide incentives to other traders to innovate and increase the quality of their own

⁷ <https://www.oberlo.com/blog/online-review-statistics>

⁸ [Press release, The Institute of Customer Service \(2022\)](#)

offering to maintain or gain market share, leading to an increase in competitive pressure in the market.

Consumer spending redirected to fair traders

193. To ensure a prudent approach towards impact quantification was taken at this stage, the CBA presented in this IA counts all foregone consumer expenditure arising from fake reviews as a cost to business. Some consumers who choose not to buy a product or service because of the absence of fake reviews will likely choose to spend their money elsewhere, particularly where consumers are seeking a specific type of product. As a result, following a reduction in the prevalence of fake reviews, expenditure on products or services with fake reviews will now be redirected towards businesses who truly offer higher quality as review authenticity improves.

194. Given weak evidence on how consumer spending will be diverted following improvements in review authenticity and in the interest of fully acknowledging the potential costs the proposals could place on businesses, this effect has not been quantified. Considering this, the true cost to business arising from foregone revenue is likely to be smaller than estimated.

Consumer utility gains

195. Through improved review authenticity consumers will be better equipped to purchase goods and services that better match their preferences and offer greater utility. For example, through purchasing higher quality products consumers may buy goods with longer lifespans or that pose less of a safety risk.

Business savings from reduced issue resolution

196. As quantified in the cost-benefit analysis, and evidenced in the CPS 2021, when consumers purchase a product or service they are ultimately unhappy with, in some cases they will seek compensation. Seeking compensation for a consumer causes detriment through the costs it imposes. These costs arise from monetary costs, such as paying for return shipping, the time spent seeking resolution and any emotional distress caused.

197. Issue resolution will also impose costs on the involved business given they must dedicate resource to handling complaints and potentially issuing compensation or refunds. A reduction in the prevalence of fake reviews will reduce this detriment whilst also offering a corresponding savings to businesses who no longer need to dedicate staff resource towards issue resolution cases.

198. Although the CPS offers evidence on the reported monetary costs imposed on consumers during instances where they sought compensation, it does not provide information on the time and resource dedicated by businesses during these cases. Therefore, considering weak evidence in this space and in the interest of not understating the costs potentially imposed on businesses, this impact has not been quantified.

A summary of the potential trade implications of measure

199. The proposals set out in this IA are not expected to have an impact on trade as they primarily focus on improving the transparency of online reviews viewed by UK consumers.

200. The new rules will apply to all businesses operating within the UK and will not impose different requirements between domestic and foreign businesses. Therefore, no trade or investment barriers are anticipated to arise because of the policies.

201. The UK market for e-commerce is primarily formed of large businesses operating domestically or large platforms acting as intermediaries (e.g., Amazon), connecting

buyers and sellers. Whilst online shopping on platforms may involve an international element where goods are shipped to and from overseas, this is typically done through the intermediary. Whilst improvements in review authenticity might affect the volume of transactions on a platform where consumers make more informed decisions, it is unlikely to alter levels of international trade. This is because changes in review quality are likely to have an effect on individual sellers or market segments within a platform rather than at the level of large trade intensive industries. As a result, this type of e-commerce activity is often classed as retail activity as opposed to trade activity.

202. Furthermore, reviews on services have a predominantly domestic influence as this type of activity is intangible and faces various regulatory, legal and cultural barriers which make operating at an international level non-viable for most UK service providers operating in review intensive industries such as entertainment and recreation. Consequently, the reforms are not expected to impact the balance of trade or investment in the services sector either.

203. As discussed in para. 190, improved review transparency will promote competition between UK businesses in the digital landscape. Where an international element is involved, improvements in quality arising from enhanced competition may make UK goods and services more appealing to international consumers. However, for the reasons given above, this effect is not expected to be significant on levels of trade given it will be contained within platforms. That said, online platforms do grant easier access to global markets, and increased transparency and competition may drive UK businesses to expand their presence in international markets which could strengthen levels of UK trade and investment.

204. In conclusion, no significant impact on trade or investment is expected to arise as a result of the policy proposals in this IA. Therefore, notifying the World Trade Organisation (WTO) is not necessary.

Monitoring and Evaluation

205. The reforms proposed in this impact assessment are expected to be reviewed following implementation to assess whether they have achieved the stated objectives, and to inform future policy making. Given these amendments are being made through secondary legislation, a review is non-statutory, however an evaluation will provide the government with an opportunity to inform any lessons learned for future interventions in this space.

206. Given the proposals are at consultation stage and subject to change, a detailed monitoring and evaluation (M&E) plan has not been developed yet. Following feedback received during the consultation and further stakeholder engagement, a M&E plan will be designed when the proposals are more developed.

207. The M&E plan will include key evaluation questions which will inform the extent to which any provisions achieved the intended objectives. Fit for purpose benefits indicators will also be developed upon the proposals being finalised which will inform the design of the planned M&E.

208. A detailed M&E plan will be included in the final stage impact assessment.

Public Sector and Equality Duty

209. **The Department is required to comply with the public-sector equality duty (PSED) set out in the Equality Act 2010 (“the Act”).** The PSED requires the Minister to have due regard to the need to advance equality of opportunity, hinder discrimination and foster good relations between those with and without certain protected characteristics. This due regard is taken to eliminate unlawful discrimination and to tackle prejudice and promote understanding. The characteristics that are protected by

the Act are: age, disability, gender reassignment, marriage or civil partnership (in employment only), pregnancy and maternity, race, religion or belief, sex and sexual orientation.

210. The proposals in this IA aim to improve the transparency of reviews hosted online. The proposals therefore only seek to improve the trustworthiness of review content as opposed to fundamentally changing the way reviews are presented to consumers. The Department acknowledges that the possibility of digital exclusion must be considered when forming regulation in the digital landscape to ensure no group is negatively impacted in a disproportionate manner. Given that reviews are hosted online, the department has explored the possibility that consumers with certain characteristics outlined in the Equality Act 2010 may be impacted in two ways. Firstly, through examining data on online shopping and secondly through primary research exploring the impact of fake reviews on consumer decision making.
211. According to ONS data⁹ on internet access, in 2020, 19% of disabled people reported that they had not shopped online within the past 12 months compared to 13% for the total population. The internet access survey also found 35% of those aged over 65 had not shopped online in the past 12 months during the same period. This demonstrates that there are discrepancies in the level of online shopping activity between different groups in the UK.
212. The proposals in this IA seek to improve the transparency of online reviews in a way which does not impact how users access reviews nor view them. Whilst the department acknowledges that some protected groups may encounter accessibility issues when shopping and viewing online reviews, this is outside the scope of the policy proposals as they do not alter any individual group's ability to use reviews relative to the status quo.
213. Additionally, the research carried out by DBT, found that "The impact of fake reviews on consumers does not vary depending on their demographic characteristics. We did not find any differences in the effect of fake reviews on characteristics such as age, sex, and ethnicity. This suggests that fake reviews have a similar impact on different groups of UK consumers."¹⁰ This provides further evidence that the measures set out in this IA will not disproportionately impact any of the protected characteristics in a negative way. Moreover, where review transparency is improved, this will benefit all users, including those within the protected groups as the information available to them becomes more trustworthy.
214. **The matters considered in this Impact Assessment do not raise any issues relevant to the public sector equality duty under section 149(1) Equality Act 2010** because the policy does not discriminate or unjustly favour any person or group of people based on their protected characteristics. Therefore, the government will proceed with consulting on the reforms as planned.

Environment Test

215. The measures covered in this IA are not expected to have an impact on the environment or the UK's carbon reduction targets. The measures are not expected to have an impact on the balance of goods imported or exported into or from the UK. Furthermore, the proposals are not expected to impact industrial production or energy consumption in the UK. Therefore, no adverse environmental impacts are expected because of the proposals.

⁹ [Internet access – households and individuals, Office for National Statistics, 2020 – Table 8](#)

¹⁰ Investigating the Prevalence and Impact of Fake Online Reviews, Department for Business and Trade, 2023

Regional Impacts

216. The measures covered in this IA are not expected to have disproportionate impacts across different regions or localities of the UK. This is because all traders in the UK must comply with consumer law and therefore will face the same obligations.

217. Furthermore, given the high rate of internet access seen in the UK, the benefits associated with improved review authenticity will benefit consumers across the UK, as opposed to groups in specific regions. Though certain groups who use online reviews more may experience greater benefits, the proposals in this IA are not expected to negatively impact any consumer group.

Appendix 1 - Technical Annex - Inferring Additionality of Policy Intervention

169. This annex covers the methodology used for estimating assumptions on the indicative proportion of fake reviews which may be prevented or removed following each policy option, i.e., the additionality of each option. The assumptions this section relates to are outlined in para.108F of this IA.

171. Assumptions on the expected reduction in fake review prevalence arising from each option are derived from two data sources.

172. Firstly, the EU Consumer Conditions Scoreboard is a biennial reporting exercise to monitor consumer sentiment across the EU, as well as in Iceland and Norway. It collects data on national conditions for consumers with regards to knowledge and trust, compliance and enforcement and complaints and dispute resolution (including fake reviews). This is used to assess the additionality of the proposals relating to the banning of the trading of fake reviews.

173. Secondly, data on the network features¹ of known fake reviews from Amazon gathered for the study 'The Market for Fake Reviews' by He et al. This information is used to assess the additionality of requirements for businesses who host reviews to ensure they are genuine and not misleading.

Categorising of fake reviews

174. The approach uses data on fake reviews compliance from the Consumer Conditions Scoreboard² to infer assumptions on the extent to which the following additions to Schedule 18 will reduce the prevalence of fake reviews:

- a. Submitting, commissioning or incentivising any person to write and/or submit a fake review of products or traders.
- b. Offering or advertising to submit, commission or facilitate a fake review.

175. For the purposes of quantification, the impact of these measures in reducing the prevalence of fake reviews is assessed together given a lack of available evidence on the balance of fake reviews which are commissioned and those exchanged for advertising.

176. Given existing evidence which suggests that many fake review brokers operate in other jurisdictions, and that the trading of fake reviews is clearly a misleading and unfair business practice, full compliance with the above bans on trading is not assumed in the CBA.

177. The Consumer Conditions Scoreboard gathered UK data on the percentage of businesses who had reported seeing fake reviews and the percentage of retailers that agree competitors comply with consumer legislation. These metrics are used to assume a 'reversion to the mean' for fake review trading compliance relative to the reported compliance with the rest of consumer law. This is expected to arise from a reduction in demand from UK businesses who commission fake reviews.

¹ Network features of reviews refer to how products and services are connected by shared reviewers, DBT research found that this was a good predictor of review authenticity

² https://ec.europa.eu/commission/presscorner/detail/en/ip_23_1891

178. In terms of assessing the impact arising from the measures to ensure businesses who host reviews take reasonable and proportionate steps to ensure they are genuine and not misleading; the dataset of known fake reviews is used to identify the proportion of additional fake reviews which could reasonably be expected to be detected following each option.
179. The following sections cover the detailed methodology and underlying data for each approach set out above.

Banning the trading of fake reviews

180. Question C4_R of the Consumer Conditions Scoreboard reports the “Percentage of retailers having come across fake reviews”. In 2020, this was reported at 26.3%. Assuming the proportion of businesses who report having come across fake reviews is indicative of the percentage of businesses who are using fake reviews, it can be inferred that approximately 73.7% are not commissioning or writing fake reviews.
181. The above assumption has been sense checked against evidence that suggests, pre-review moderation by platforms, roughly 25% of all reviews posted may be fake³. This figure is derived from data reported by Amazon⁴, which outlined that roughly 13% of reviews were proactively blocked by their review screening features. With recent research suggesting around 12% of remaining reviews (post-review moderation) are fake, this would suggest that roughly 25% of all reviews posted on products are fake.
182. Question C7_4 reports on the “Percentage of retailers that agree competitors comply with consumer legislation”. In 2020, this was reported at 77.5%. Similarly, to the above, it is assumed that this is representative of UK business compliance with consumer law.
183. Presuming that the ban on trading fake reviews brings compliance in line with the rest of UK consumer law, it is assumed that the proportion of businesses using fake reviews falls into line with the number of businesses who are generally compliant with existing UK consumer legislation.
184. Therefore, the expected decrease in the proportion of businesses using fakes following the ban is equal to the difference in the percentage of businesses who are compliant with existing UK consumer legislation (77.5%) and those who are not currently trading fake reviews (73.7%).
185. From the figures above, this leads to the assumption that there will be a 3.8% reduction in the proportion of businesses commissioning or writing fake reviews. Assuming that this is proportional to the prevalence of fake reviews, it is inferred that this results in a 3.8% reduction in prevalence.
- 186. From the outputs of DBT research, the analysis assumes roughly 12% of post-moderation reviews are fake. Therefore a 3.8% reduction in the prevalence of fake reviews represents an overall 32% decrease as a proportion of all fake reviews.**
- 187. In conclusion, it is assumed that the ban on trading fake reviews results in a 32% reduction in prevalence.**

³ This includes reviews that are detected and removed by platforms during review screening activity

⁴ [Amazon Review Blueprint, Amazon, 2022](#)

Reasonable and proportionate steps by businesses to ensure reviews are genuine and not misleading to consumers

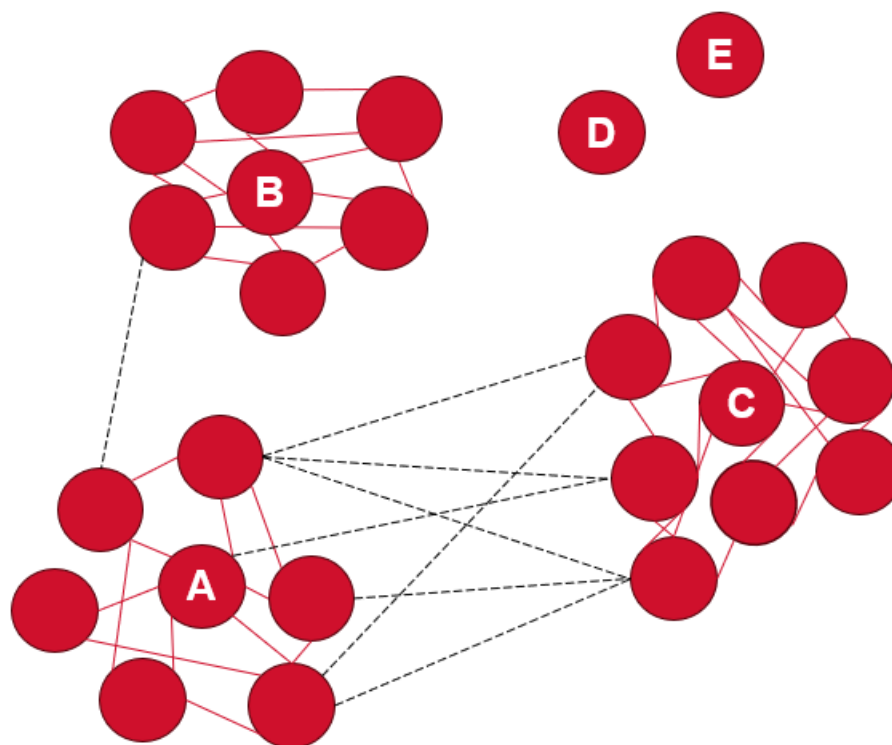
188. The impact of requiring businesses to take reasonable and proportionate steps to ensure reviews are genuine and not misleading before hosting them is expected to lead to additional moderation activity by businesses who host reviews.
189. As discussed earlier in this IA, the extent to which this leads to additional moderation activity is dependent on the stringency of the policy option taken, i.e., whether legislation or guidance is used. It is assumed that legislative approaches incentivise greater moderation activity compared to approaches incorporating non-legislative elements due to the greater penalties associated with non-compliance.
190. Although the extent to which the different approaches to outlining reasonable and proportionate steps leads to additional review moderation is inherently uncertain, indicative scenarios on the number of fake reviews which could reasonably be expected to be captured following each option can be developed based on the features of known fake reviews and how these features are distributed.
191. The features of reviews are instrumental in categorising their difficulty of detection as fake because fake reviews often exhibit telltale signs that differ from genuine reviews. These features might include linguistic patterns, reviewer activity and emotional tone. Analysing these and other features can assist in categorising the level of difficulty in detecting fake reviews, enabling platforms to fine-tune their fake review detection algorithms more efficiently.
192. If implemented, the reasonable and proportionate steps review hosting businesses of different sizes are expected to take because of the policy proposals will vary based on their capabilities. As a result, businesses will not always be expected to implement automated review screening procedures considering the technical capability and resource needed to implement these systems. However, given the high market share held by large review hosting platforms, the CBA focuses on the activity of these actors who are capable, and currently do, of screening reviews through automated means. Consequently, for the purposes of this CBA, the assessment of reasonable and proportionate steps focuses on automated means of review moderation.
193. Review network features refer to the behaviour, or network, of reviewers across multiple products. Given the randomness of shopper activity, buyers leaving genuine reviews would be expected to be evenly spread across products given the activity is organic and a strong pattern of common reviewers reviewing the same product wouldn't be expected to arise spontaneously. Where reviewer networks are highly clustered, i.e., many common reviewers shared across common products, this is indicative of suspicious and likely co-ordinated activity.
194. Given the availability of data, and the predictive power of network features in predicting review authenticity demonstrated by DBT research, it is assumed that highly clustered fake reviews are easier to detect, and less clustered fake reviews are harder to detect through automated means.
195. This assumption has been deemed appropriate given that network features have been empirically proven as a strong predictor of review authenticity, therefore where reviews are particularly clustered, indicating suspicious activity, an automated review moderating system can conclusively identify these as fake. This forgoes the need to rigorously

assess other review features commonly used, such as textual⁵ and metadata⁶ features, which have also been found to provide a less accurate prediction of authenticity. Furthermore, this would also forgo the need to employ a human moderator to assess particular review cases manually.

196. Network features can be examined empirically through specific metrics which capture different features of a review network. In the case of reviews, the network is formed by how products are connected by common reviewers. Therefore, a product in a network represents a node and common reviewers between products represent links between nodes.

197. Figure 2 below presents an illustrative diagram of a reviewer network, where each node is a product (the circles) connected by reviewers (the links).

Figure 2 - Illustrative Diagram of Fake Review Product Network



198. The clustering of the reviewer network can be used to predict review authenticity. This is because where authentic review activity has taken place naturally, high clustering (i.e. highly connected by common reviewers) of products is not expected given the chances of this happening organically are very low. In DBT research on review networks, the clustering coefficient and eigenvector centrality were found to have the highest predictive power when predicting review authenticity.

199. In figure 2, products A, B and C are in a cluster of common reviewers, indicating suspicious activity, whereas products D and E are not clustered. Product A and C's

⁵ Textual features include cosine similarity, word count and percentage of words using capitals.

⁶ Metadata features of reviews includes features such as the star rating of a review, the time between reviews and the average rating.


clusters are also highly connected to each other, whilst cluster A is also slightly connected to cluster B.

200. **The clustering coefficient measures how closely connected a product is to other products.** In figure 2, products in clusters A, B and C would have a similar clustering coefficient. Products D and E would have a clustering coefficient of zero as they have no connections.
201. **Eigenvector centrality measures the extent to which clustered products are also connected to other highly clustered products.** This measures how important they are in a network, or in other words, how easy it is to get from one cluster to another. In figure 2, cluster A would have the highest eigenvector centrality as it is connected to both clusters B and C.
202. Based on the above two network features, products in cluster A would be most likely to be identified as fake.
203. He et al collected a dataset of fake reviews from Amazon through investigating groups on social media websites which were brokering fake reviews. From this data set of ground truth, a network can be built using the unique reviewer IDs and product IDs to investigate and compare the features of fake and genuine reviews. Once this network is constructed, features such as the clustering coefficient and eigenvector centrality can be estimated.
204. Once the network features of each review have been estimated, they can be examined to categorise fake reviews into five clustering groups ranging from ‘Very low clustering’ to ‘Very high clustering’.
205. Based on the distribution of the clustering coefficients and eigenvector centrality of fake reviews seen in the Amazon dataset of known fakes, the proportion of fake reviews falling into each category is as follows:

Table 3 - Proportion of Fake Reviews by Clustering Category

<u>Network clustering category</u>	<u>Proportion</u>
Very low clustering	8%
Low clustering	26%
Moderate clustering	24%
High clustering	36%
Very high clustering	6%

Harder to detect



Easier to detect

206. In terms of mapping the above categories to each policy option based on the level of stringency, it is assumed that:
- a. **Option 1 (legislation only) captures moderate to very highly clustered reviews** – totalling 66% of remaining fake reviews.
 - b. **Option 2 (guidance) captures very highly clustered reviews only** – totalling 6% of fakes reviews.
 - c. **Option 3 (legislation & guidance) captures very high and highly clustered reviews** – totalling 42% of fakes reviews.

207. Given the difficulty in spotting fake reviews which are not particularly clustered, and the inevitable constraints faced by algorithmic methods of moderation, it is assumed that low and very lowly clustered fake reviews go uncaptured by all options.
208. The proposals banning the trading of fake reviews is expected reduce the prevalence through preventing them being bought. Consequently, seeing as these reviews would never be posted, the estimated reductions in prevalence above are applied after the impact of the ban has been accounted for. Further to this, fake reviews prevented through the ban on trading also do not impose moderation costs on platforms.
209. This analysis is only indicative of how fake reviews may be detected by the automated review detection systems used by platforms. In reality, these systems and the features they inspect will differ from platform to platform. Further to this, platforms may currently have varying levels of tolerance when it comes to suspicious review activity given the absence of legislation to date which sets out clear standards businesses should abide by in respect to review moderation.
210. Government acknowledges that platforms have administrative access to a wider range of user data other than network features (such as location data) which may also have predictive power in assessing review authenticity. Considering this, the approach taken above on network features is not meant to be representative of all review moderation systems, rather it is intended to be a simplified model of review moderation which provides indicative scenarios of how many fake reviews could reasonably be detected in response to different levels of policy stringency.

Summary

211. The assumed additional impact on the prevalence of fake reviews of each policy package is summarised below.

Table 4 - Options Impact Summary

Impact on prevalence of fake reviews			
Option	Banning the trading of fake reviews	Banning the hosting of reviews without taking reasonable and proportionate steps to ensure they are genuine and not misleading	Total ⁷
Option 1	32% reduction	66% reduction	77% reduction
Option 2	32% reduction	6% reduction	36% reduction
Option 3	32% reduction	42% reduction	42% reduction

⁷ The impact of reasonable and proportionate steps by businesses is applied after the impact of the ban on trading is accounted for (as these reviews would never be posted). Therefore, for Option 1, the total impact on the prevalence of fake reviews would be 32% + ((1- 32%) x 66%)

Appendix 2 – List of Questions

- 1) Do you agree that the costs and benefits quantified are the correct ones for the policy impact assessment? Are there any key impacts which have been overlooked?
- 2) Do you agree with the assumed algorithmic moderation cost per review? If not, please provide your reasoning and provide any supporting information you have.
- 3) Do you agree with the assumed human review moderation assumptions? If not, please provide your reasoning and provide any supporting information you have.