Holding Platforms Liable[*]

Xinyu Hua[†]   Kathryn E. Spier[‡]
HKUST   Harvard University

June 3, 2022

Abstract

Should platforms be held liable for the harms suffered by users? A two-sided platform enables interactions between firms and users. There are two types of firm: harmful and safe. Harmful firms impose larger costs on the users. If firms have deep pockets then platform liability is unnecessary. Holding the firms liable for user harms deters the harmful firms from joining the platform. If firms are judgment proof then platform liability plays an instrumental role in reducing social costs. With platform liability, the platform has an incentive to (1) raise the interaction price to deter harmful firms and (2) invest resources to detect and remove harmful firms from the platform. The residual liability assigned to the platform may be partial instead of full. The optimal level of platform liability depends on whether users are involuntary bystanders or voluntary consumers, and the intensity of platform competition.

1 Introduction

Online platforms are ubiquitous in the modern world. We connect with friends on Facebook, shop for products on Amazon, and search online for jobs, information, and entertainment. While the economic and social benefits created by platforms are undeniable, the costs and hazards for users are very real too. For example, platform users run the risk that their personal data and privacy will be compromised. Users of social networking sites may be misled by false information or harmed by cyberbullying and hate speech. Consumers who shop online run the risk of purchasing counterfeit, defective, or dangerous goods. Should internet platforms like Facebook and Amazon be liable for the harms suffered by users?

In the United States, platforms enjoy relatively broad immunity from lawsuits brought by users, although this immunity is being challenged in legislatures and the courts.[1] Section 230 of the Communications Decency Act, enacted in 1996, shields platforms from liability for the digital content created by their participants.[2] Early proponents argued that the law was necessary to allow the internet to grow and flourish, but its application is controversial and many critics question the law's merits.[3] Proposed federal legislation, including the "Health Misinformation Act of 2021," would strip platforms of Section 230 protections if the platforms facilitate the spread of misinformation about public health emergencies.[4] In 2021, Zoom reportedly agreed to pay $85 million to settle a lawsuit alleging that Zoom shared users' personal data with third parties and failed to provide appropriate security measures.[5]

Marketplace platforms have largely avoided responsibility for defective products and services sold by third-party vendors. In 2019 the Fourth Circuit held that Amazon.com is not a traditional seller and therefore not subject to strict tort liability.[6] The following year, a California court found that Amazon could be held strictly liable for a defective laptop battery that was sold by third-party vendors but "Fulfilled by Amazon."[7] Then, in

---

[1]See Buiten et al. (2020) for discussion of the European Commission's e-Commerce Directive. Hosting platforms in the EU may avoid liability for illegal content posted by users, assuming they are not aware of it, and are not responsible for monitoring the legality of the posted content.

[2]Section 230(c)(1) says that "No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider." See "Why Hate Speech on the Internet Is a Never-Ending Problem." *New York Times*, August 6, 2019.

[3]See *Force v. Facebook, Inc.*, No. 18-397 (2d Cir. 2019). The court opined that Section 230 "should be construed broadly in favor of immunity."

[4]Senators Amy Klobuchar and Ben Ray Lujan introduced this bill to combat misinformation about COVID-19. See "Bill Targets Health Misinformation." *Wall Street Journal*, July 23, 2021.

[5]See "After Issue With Hackers, Zoom Agrees to Settle Lawsuit Over 'Zoombombing'." *New York Times*, August 2, 2021.

[6]See *Erie Ins. Co. v. Amazon.com*, 2019 WL 2195146 (4th Cir. May 22, 2019). The Ninth Circuit reached a similar conclusion in *State Farm Fire and Casualty Company v. Amazon.com, Inc.*, 2020 WL 6746745 (9th Cir. Nov. 17, 2020). See also *Great Northern Insurance v. Amazon.com, Inc.*, 2021 WL 872949 (N.D. Ill. March 9, 2021).

[7]See *Bolger v. Amazon.com, LLC*, 53 Cal.App.5th 431, 267 Cal.Rptr.3d 601 (2020). The court held that Amazon "is an integral part of the overall producing and marketing enterprise that should bear the cost of injuries resulting from defective products."

2021, Amazon was held strictly liable for harms caused by a defective hoverboard that was shipped directly to the consumer by an overseas third-party vendor. Although Amazon did not fulfill the hoverboard order, the court opined that Amazon was "instrumental" in its sale and that "Amazon is well situated to take cost-effective measures to minimize the social costs of accidents."[8] In short, the law is far from settled.

This paper presents a formal model of a two-sided platform with two kinds of participants, "firms" and "users." The platform enables interactions between the firms and users, and charges the firms a fixed price per interaction.[9] There are two types of firm: harmful and safe. The harmful firms enjoy higher gross benefits per interaction but impose larger costs on the users.[10] Interactions between harmful firms and users are socially inefficient (the costs exceed the benefits). In an ideal world, the harmful firms are deterred from joining the platform. If the harmful firms remain undeterred, however, the platform plays an instrumental role in reducing social costs. The platform has the ability to prevent harmful interactions by either raising the interaction price or by investing resources to detect and remove the harmful firms from the platform.[11]

In our baseline model, the *users are bystanders* of the firms. Such settings include social and professional networking platforms such as Facebook and LinkedIn where the users enjoy same-side network benefits from sharing content with each other and the firms pay the platform to access user data or to engage in influential activities (e.g., advertising). Platform users may be harmed by the firms when their private data is breached or when they are exposed to harmful advertising or misinformation. Absent liability the harmful firms have no incentive to leave the platform, and the platform has an insufficient incentive to detect and remove them. Holding the firms and the platform jointly liable gets them to internalize the negative externalities on the user-bystanders.

If the firms have deep pockets, and must pay in full for the harms they cause, then platform liability is unnecessary. Holding just the firms liable achieves the first-best outcome. Platform liability is socially desirable when the firms are *judgment proof* and immune from liability.[12] First, if the platform is held liable, the platform will raise the interaction price for the firms to reflect the platform's future liability costs. If the harmful firms are "marginal" (i.e., the harmful firms have a lower willingness to pay than the safe firms) then the higher interaction price deters the harmful firms from joining the platform. Second, if the harmful firms are "inframarginal" and undeterrable, the platform will invest

---

[8]*Loomis v. Amazon.com LLC,* 2021 WL 1608878 (Cal. App. Ct. April 26, 2021).

[9]Consistent with the literature, we assume that the platform does not charge users. Section 3 extends the model to retail platforms where the consumers pay the firms and the firms pay the platform.

[10]The focus of this paper is cross-side harms. Similar issues arise when the injurers and victims are on the same side of the market.

[11]According to Amazon's post, "[Amazon's] proactive measures begin when a seller attempts to open an account. Our new seller account vetting includes a number of verifications and uses proprietary machine learning technology that stops bad actors before they can register or list a single product in our store." See https://www.aboutamazon.com/news/company-news/product-safety-and-compliance-in-our-store.

[12]Shavell (1986) provides the first rigorous treatment of the judgment proof problem, where injurers with limited assets tend to engage in risky activities too frequently and take too little care.

resources to detect and remove the harmful firms from the platform.[13] Interestingly, the optimal level of platform liability may be partial instead of full, as full liability could lead to excessive auditing by the platform.[14]

We then extend the baseline model to settings where *users are customers* of the firms, so interactions require the users' consent. Relevant settings include online marketplaces like eBay and Amazon where participants enjoy cross-side benefits from the sale of goods and services. As in the baseline model there are two types of seller, harmful and safe. The harmful sellers have lower production costs but cause harms more frequently. The consumers are sophisticated and their willingness-to-pay reflects their rational expectations about product risks.[15] The risk of harmful products depresses the price that consumers are willing to pay and, by extension, depresses the revenues that the platform can generate. If the harmful firms are marginal, then platform liability is unnecessary. Even without liability, the platform has a private incentive to raise the interaction price to deter the harmful firms from joining the platform. If the harmful firms are inframarginal, however, then partial platform liability gives the platform an appropriate incentive to audit and remove the harmful firms.[16] Since the platform internalizes the average harm to consumers, the socially-optimal platform liability is *lower* than in the baseline model (e.g., for social media platforms).

Next, we extend the baseline model to consider two *competing platforms*. The users are bystanders and can participate on both platforms (i.e., multi-homing), while the firms can only participate on one of the platforms (i.e., single-homing). If the harmful firms are marginal then competition reduces the platforms' incentives to deter the harmful firms by charging high prices, relative to the baseline model. Therefore the socially optimal platform liability is (weakly) *higher* than that in the baseline monopoly model. If the harmful firms are inframarginal, holding the platforms partially liable for the residual harms motivates them to make the socially efficient auditing effort. In this case, since competition reduces the price-cost margins from serving the harmful firms, the competing platforms have stronger incentives for auditing than the monopoly platform. Thus, the socially optimal platform liability is *lower* than that in the baseline model. These observations suggest that policies encouraging platform competition should be complemented by changes in platform liability.

Our paper is related to the law-and-economics literature on products liability where firms are held liable for the product-related harms suffered by consumers. Products liability may be socially desirable if consumers misperceive product risks (Spence, 1977; Epple and Raviv, 1978; Polinsky and Rogerson, 1983) or if consumers are not able to

---

[13]If the firms are very judgement proof and can evade liability, then the harmful firms are inframarginal. If the firms are moderately judgment proof, then the harmful firms are "marginal."

[14]If the firms are completely judgment proof, then the safe firms are marginal and the harmful firms get information rents. When choosing its audit intensity, the platform does not take into account the lost rents when the harmful firms are removed from the platform.

[15]Platform liability is unnecessary if the firm-sellers have sufficiently deep pockets and compensate the user-consumers for the harms.

[16]As in our baseline model, full liability would lead to excessive auditing by the platform.

observe product safety at the time of purchase (Simon, 1981; Daughety and Reinganum, 1995).[17] Building on Spence (1975), Hua and Spier (2020) emphasize the particular importance of firm liability when consumers are heterogeneous so the marginal buyer's preferences are not representative of the average consumer.

Our paper is also related to the literature about *extending* liability to parties who are not directly responsible for the victim's harms. Hay and Spier (2005) examine whether manufacturers should be held liable if a consumer, while using the product, harms somebody else (third party bystanders). If consumers are judgment proof and cannot be held accountable for the harms they cause, then extending liability to the manufacturer can help the market to internalize the harms.[18] Pitchford (1995) explores the desirability of extending liability to an injurer's lenders[19] and Dari Mattiacci and Parisi (2003) consider vicarious liability where liability is extended to the injurer's employer.[20] Our model, which has not been previously studied, investigates the design of platform liability when the platform can audit and remove harmful participants.

There is a vast literature on multi-sided platforms. The early studies (e.g., Caillaud and Jullien, 2003; Rochet and Tirole, 2003, 2006; Amstrong, 2006; and Weyl, 2010) have identified how cross-side externalities affect platform pricing schemes and users' participation incentives. The literature also examines the impact of seller competition (on a monopoly platform)[21] or the impact of platform competition on pricing.[22] Some recent studies pay attention to non-pricing strategies, including seller exclusion (Hagiu, 2009), information management (Julien and Pavan, 2019; Choi and Mukherjee, 2020), control right allocation (Hagiu and Wright, 2015, 2018), and platform governance (Teh, forthcoming). A few policy papers (Buiten, de Streel, and Peitz, 2020; Lefouili and Madio, 2022) discuss informally whether platforms should bear liability for harms caused by participants. Our paper contributes to the literature by investigating the effects of platform liability on platform pricing and auditing incentives, as well as their welfare implications.

Our paper is organized as follows. Section 2 presents the baseline model where users are bystanders to firms on a monopoly platform. This section explores the impact of liability on the platform's pricing and auditing as well as social welfare. Section 3 examines an alternative setting where the firms are sellers and the users are consumers. Section 4

---

[17]See also Simon (1981), Daughety and Reinganum (1995, 1997, 2006, 2008a and b, 2014), Arlen and Macleod (2003), Wickelgren (2006), Chen and Hua (2012, 2017), Choi and Spier (2014).

[18]Brooks (2002), and Fu, Gong, and Png (2018) investigate how legal responsibility affects firms' choice between vertical integration and outsourcing.

[19]See also Boyer and Laffont (1997) and Che and Spier (2008). Bebchuk and Fried (1996) argue informally for raising the priority of tort victims in bankruptcy above debt claims gives the debtholders an incentive to better monitor the borrower.

[20]There are related legal studies. See Kraakman (1986) for a taxonomy of gatekeeper enforcement strategies, Hamdani (2002) about liability on internet service providers and Hamdani (2003) about liability on gatekeepers such as accountants and lawyers.

[21]See Nocke, Peitz, and Stahl (2007), Galeotti and Moraga-Gonzalez (2009), Hagiu (2009), Gomes (2014), Belleflamme and Peitz (2019).

[22]See Dukes and Gal-Or (2003), Hagiu (2006), Armstrong and Wright (2007), White and Weyl (2010), Karle, Peitz, and Reisinger (2020), Tan and Zhou (2021).

extends the baseline model by considering two competing platforms. Section 5 provides concluding thoughts. The proofs are in the appendix.

2   The Model

Consider a two-sided platform (P) with two kinds of participants, firms (S) and users (B). The platform is a monopolist and necessary for interactions between firms and users.[23] Firms and users are small, have outside options of zero, and the mass of each is normalized to unity.

The platform provides two goods. First, the platform provides a quasi-public good that gives each user a private benefit $v > 0$, which we assume is the same for all users.[24] Second, the platform provides opportunities for the firms and the users to interact. The platform charges the firms a price p per interaction. Google, for example, currently enjoys a market share of more than 92% of the search engine market. There are approximately seven billion free Google searches conducted by users every day. Google monetizes the quasi-pubic good by selling online advertising to businesses through real-time auctions.[25]

We assume that interactions between firms and users do not require the users' consent and so the users are effectively "bystanders."[26] The benefits and costs of these interactions depend on the firms' type, $i \in \{H, L\}$, where λ is the mass of type H and $1 - \lambda$ is the mass of type L in the firm population.[27] The H-type firms have higher interaction benefits, $\alpha_H > \alpha_L$, but impose higher interaction losses on users, $\theta_H d > \theta_L d$ where $\theta_i \in [0, 1]$ is the probability of harm and $d > 0$ is the level of harm per firm-user interaction.[28] The firms privately observe their types.

This general specification is aligned with a variety of economic settings. First, platform users may be harmed when their personal data is compromised. Prominent examples include the breach of Facebook user data by consulting firm Cambridge Analytica.[29] Some of the firms participating in Google's auctions allegedly collect and store so-called "bidstream data" on users, which they subsequently sell to third parties (including hedge funds and political campaigns).[30] Second, users often bear direct harms from fraudulent or

---

[23]Section 4 extends the analysis to consider platform competition.

[24]For example, social platforms generate same-side network effects by attracting many users.

[25]Over 80% of Google's revenues in 2020 came from selling ads. See Google's annual report. Google's expertise in collecting and analyzing troves of user data increases the firms' willingness to participate in these auctions. Similarly, most of Facebook's revenue comes from advertising.

[26]Section 3 extends the analysis to retail platforms where interactions require the users' consent.

[27]For simplicity, λ is taken as exogenous. One may endogenize λ by allowing firms to invest resources to increase the likelihood having safe products. As shown below, if the firms are very judgment proof then the H-types are inframarginal and earn information rents. It follows that the firms' ex ante incentive to invest in product safety would be socially insufficient.

[28]If $\alpha_H < \alpha_L$ then the H-types are marginal for all liability rules and auditing is unnecessary. The threshold $\widehat{w}$ defined in (5) below is identically equal to zero, and all of our results apply.

[29]The user data was allegedly used for political purposes. Facebook paid a $5 billion fine.

[30]See "Wyden FTC Letter" https://www.wyden.senate.gov/imo/media/doc/073120 20Wyden 20Cassidy 20Led 20FTC 20Investigation 20letter.pdf

simply unwanted advertising.[31] It has been estimated that displayed advertising accounts for a large share of the data costs for mobile telephone plan users in the United States.[32] Third, our specification is also aligned with retail platforms when user-consumers consent to transactions (sales) but are unaware that the products and services are potentially dangerous.[33] Finally, although our focus is on harms to platform participants themselves, our insights also apply to harms to third parties who are external to the platform.[34]

We assume that the platform has the capability to detect and block the H-type firms. We will refer to the platform's efforts to detect the H-types as auditing. By virtue of their scale, data, and technological sophistication, platforms like Google may be in a good position to root out harmful platform participants. Specifically, by spending effort $e \in [0, 1)$ per firm, the platform can detect H-type firms with probability e and block them from interacting with users.[35] We assume that the cost of effort $c(e)$ satisfies $c(0) = 0, c'(e) > 0, c''(e) > 0, c'(0) = 0$, and $c'(e) \to \infty$ as $e \to 1$. The effort level e is neither observable nor contractible.[36] Thus, there is a potential moral hazard problem associated with auditing.[37]

Suppose that both types of firms seek to join the platform. Given audit intensity e, the number of firms that remain on the platform is $\lambda(1 - e) + (1 - \lambda)$. Since there is a unit mass of consumers, this is also the number of firm-user interactions. This may be interpreted as the volume of (infinitesimally small) interactions per consumer, assuming that each retained firm interacts with each and every consumer.[38] Alternatively, one may interpret $\lambda(1 - e) + (1 - \lambda)$ as the probability of an exclusive match between a consumer and a randomly selected firm.

The platform operates in a legal environment where harmed users may sue the platform and the firms for monetary damages. If a user suffers harm d, the court orders the firm and the platform to pay damages w_s and w_p, respectively, to the user. We will assume

---

[31] Facebook has settled lawsuits alleging that they failed to block scam advertisements. See "Facebook Hit With UK Copyright Suit Over Fraudulent Ads," Law360.com, October 8, 2021. In a lawsuit brought against Google, a user clicked on a fraudulent advertisement which took her to a website where she was unknowingly charged. See Goddard v. Google, Inc., 640 F. Supp. 2d 1193 (N.D. Cal. Jul. 30, 2009).

[32] A recent study by Enders Analysis places the share at 18% to 79%. This is higher than previous studies that estimated the costs at 10% to 50%. Other costs include those of blocking unwanted advertising. See https://www.techdirt.com/articles/20160317/09274333934/why-are-people-using-ad-blockers-ads-can-eat-up-to-79-mobile-data-allotments.shtml

[33] If consumers are naive and unaware, the harm that they suffer is effectively "externalized" on their future selves. Thus, the consumers' future selves are effectively bystanders. Section 3 extends the analysis to retail platforms with sophisticated user-consumers.

[34] Example include harms to copyright holders when illegal material is posted on Facebook or Youtube, and harms to branded products when counterfeits are sold on Amazon.

[35] If the platform takes auditing effort per interaction instead of per firm, the analysis remains the same as long as the number of users is fixed.

[36] The results in this section would be unchanged if the platform could commit to e (as users are bystanders). However, if e were observable, regulation would be a viable alternative to liability.

[37] Immune from liability, Google has a financial incentive to maximize their advertising revenue without sufficient regard to the quality (or legality) of the content or the potential harms to users.

[38] This interpretation is aligned with many platform models with non-exclusive matching including Armstrong (2006) and Weyl (2010).

that $w_s, w_p \geq 0$ and $w = w_s + w_p \leq d$ so the total damage award does not exceed the harm suffered by the user.[39] For simplicity, there are no litigation costs or other transactions costs associated with using the court system.[40] There may be practical and legal limits on firm and platform liability. Third-party vendors are often liquidity constrained or "judgment proof" and cannot be held fully accountable for the harm that they cause and platforms may enjoy immunity as well.[41] Thus, in practice, liability is often limited. Given the liability rule and the interaction price p, the net expected interaction value, $\alpha_i - \theta_i d$, is shared as follows: Platform surplus is $p - \theta_i w_p$, firm surplus is $\alpha_i - \theta_i w_s - p$, and user surplus is $-\theta_i(d - w_s - w_p)$.

In the following analysis, we assume

$$
\begin{aligned}
A0 &: \quad v - [\lambda\theta_H + (1-\lambda)\theta_L]d > 0; \\
A1 &: \quad \alpha_L - \theta_L d > 0 > \alpha_H - \theta_H d; \\
A2 &: \quad \alpha_L - (\lambda\theta_H + (1-\lambda)\theta_L)d > 0.
\end{aligned}
$$

A0 implies that the users' benefit from the quasi-public good is sufficiently high that the users would join the platform even if the H-type firms join the platform and there is no liability.[42] A1 implies that it is socially efficient (inefficient) for the L-type (H-type) firms to join the platform.[43] A2 guarantees that the platform always gets non-negative profits and implies that it is socially efficient for both types to join the platform on average. These assumptions are not essential for the main insights, but simplify the analysis.

The timing of the game is as follows.

1. The platform creates the quasi-public good for users and sets the interaction price p for the firms. The price p is publicly observed.

2. Firms privately learn their types $i \in \{H, L\}$ and decide whether to join the platform.

3. The platform chooses $e \in [0, 1)$ to audit firms on the platform and removes any detected H-type firms. The audit intensity e is not publicly observed.

4. Firms interact with the users and the interaction benefit α_i and harm $\theta_i d$ are realized.

---

[39] Our main results remain valid if punitive damage awards ($w > d$) are feasible but not too large. If the total damage award is very large, the platform would not be active.

[40] We do not consider frivolous litigation where unharmed users bring lawsuits.

[41] See for example the discussion of Section 230 of the Communications Decency Act above.

[42] Similar results would be obtained in a model where users have heterogeneous valuations and some users do not join the platform. Absent liability, the platform would pay insufficient attention to the safety of the inframarginal platform users.

[43] In our model, society is better off when the monopolist excludes the H-type firm. Given our assumptions, there is no social loss from monopoly pricing. In a more general model, platform liability could exacerbate the monopoly pricing problem (as would a Pigouvian tax).

5. Harmed users sue for monetary damages and receive compensation w_s and w_p from the responsible firm and platform, respectively.

We will maintain the assumption that the platform, firms, and users are sophisticated and understand the risks of interacting on the platform. The equilibrium concept is perfect Bayesian Nash equilibrium. Our social welfare concept is the aggregate value captured by all players: the platform, the firms (both H-types and L-types), and the users.

We now present two social welfare benchmarks.

First-Best Benchmark. The first-best outcome is achieved if the socially-harmful H-type firms do not join the platform or interact with users.[44] Auditing is unnecessary (as there are no H-types to be detected and removed). Social welfare is:

$$v + (1 - \lambda)(\alpha_L - \theta_L d). \tag{1}$$

Second-Best Benchmark. Suppose that the H-type firms join the platform. Auditing is necessary to detect and remove the H-types. Social welfare is:

$$S(e) = v + \lambda(1 - e)(\alpha_H - \theta_H d) + (1 - \lambda)(\alpha_L - \theta_L d) - c(e). \tag{2}$$

The socially optimal auditing effort $e^{**} > 0$ satisfies

$$-\lambda(\alpha_H - \theta_H d) - c'(e^{**}) = 0. \tag{3}$$

At the optimum, the marginal cost of auditing, $c'(e^{**})$, equals the marginal benefit of blocking H-type firms from interacting with users, $-\lambda(\alpha_H - \theta_H d)$. e^{**} is higher when the proportion λ of H-types is larger and when the social harm $|\alpha_H - \theta_H d|$ is larger. Note that $e^{**} \in (0, 1)$ so some H-types remain on the platform in this second-best world.

Our analysis proceeds in two steps. First, we characterize the platform's pricing and auditing strategy, p and e, given the assignment of liability to the firms and platform, w_s and w_p. Second, we explore the socially-optimal platform liability rule.

2.1   Equilibrium Analysis

A type-i firm will seek to join the platform when their expected profit per interaction is non-negative,[45]

$$\alpha_i - \theta_i w_s - p \geq 0, \tag{4}$$

where α_i is the firm's interaction benefit, $\theta_i w_s$ is the firm's expected liability, and p is the price paid to the platform. Note that depending on the level of firm liability, w_s, the H-type may have higher or lower rents than the L-type. If $w_s = 0$ then the H-type firms have

---

[44]This follows from assumption A1.

[45]By assumption A0, the users derive sufficient value from the quasi-public good that they will join the platform regardless of the probability of harm and the liability rule.

higher rents than the L-type firms (since $\alpha_H > \alpha_L$). If $w_s = d$ then the H-type firms have lower rents than the L-type firms (since Assumption A1 implies $\alpha_H - \theta_H d < \alpha_L - \theta_L d$). The rents of the two types are equal when

$$w_s = \widehat{w} = \frac{\alpha_H - \alpha_L}{\theta_H - \theta_L} < d. \tag{5}$$

The threshold $\widehat{w}$ defined in (5) is critical for understanding the impact of platform liability on the interaction price and audit intensity. If the firms are sufficiently judgment-proof, $w_s < \widehat{w}$, then the L-type firms are "marginal." If the L-types are indifferent about joining the platform then the H-types strictly prefer to join. The platform will set the interaction price p to extract all the L-types' surplus and the inframarginal H-types receive rents. In this setting, we will see that a higher level of platform liability w_p creates a stronger incentive for the platform to audit the firms and remove the harmful H-types from the platform.

If the firms are only moderately judgment proof, $w_s > \widehat{w}$, then the H-type firms are marginal. If the H-types are indifferent about joining the platform then the L-types strictly prefer to join. In this setting, the platform can easily deter the socially-harmful H-types from joining the platform by raising the interaction price p; the platform need not engage in costly auditing. A higher level of platform liability w_p gives the platform a stronger incentive to raise the interaction price to deter the harmful H-types from joining the platform.

We now characterize the equilibrium for $w_s < \widehat{w}$ and $w_s > \widehat{w}$ and present the results in two lemmas.

Case 1: $\boldsymbol{w_s < \widehat{w}}$. Suppose that firm liability is below the threshold, $w_s < \widehat{w}$, so the L-type firms are marginal. The platform sets the interaction price to extract the L-type firms' rent,[46]

$$p^* = \alpha_L - \theta_L w_s. \tag{6}$$

The H-types seek to join the platform. Using the definition of $\widehat{w}$ in (5) and the formula for p^* in (6), the H-type firms' rent per interaction is $\alpha_H - \theta_H w_s - p^* = (\theta_H - \theta_L)(\widehat{w} - w_s) > 0$. Notice that as firm liability w_s grows, the H-type's information rent falls. In the limit when $w_s \to \widehat{w}$ the H-type's rent approaches zero.

We now explore the platform's incentive to audit and remove the H-type firms. The platform's aggregate profits are:

$$\Pi(e) = (1 - e)\lambda(p^* - \theta_H w_p) + (1 - \lambda)(p^* - \theta_L w_p) - c(e). \tag{7}$$

A necessary and sufficient condition for the firm to audit, $e^* > 0$, is that the platform's profit associated with each retained H-type is negative, $p^* - \theta_H w_p < 0$. Using the formula

---

[46]The platform will choose between a low price $p_L = \alpha_L - \theta_L w_s$ where both types of firm seek to join the platform and a high price $p_H = \alpha_H - \theta_H w_s$ where only the H-type firms seek to join. Assumption A2 guarantees that the platform does not find it profitable to deter the L-types and retain the H-types. It also implies that the platform's profit margin is positive.

for $\widehat{w}$ in (5) and p^* in (6), and letting $w = w_s + w_p$ be the joint liability of the firm and platform, $e^* > 0$ if and only if

$$(\alpha_H - \theta_H d) + \theta_H(d - w) - (\theta_H - \theta_L)(\widehat{w} - w_s) < 0. \tag{8}$$

The left-hand side of (8) is the platform's profit associated with each retained H-type firm. The first term, $\alpha_H - \theta_H d < 0$, is the social loss associated with each retained H-type and the second term, $\theta_H(d - w) > 0$, is the uncompensated harm to the users. The sum of these two terms, $\alpha_H - \theta_H w$, is the joint platform-firm surplus associated with each retained H-type. The third term in (8) is the information rent captured by the H-type firm. If condition (8) holds, then the platform loses money on each retained H-type firm and so the platform invests $e^* > 0$ and removes detected H-types from the platform. If (8) does not hold then the platform makes money on each retained H-type firm and has no incentive to audit and remove the H-types from the platform, $e^* = 0$.

We now explore how the private and social incentives for auditing diverge when $e^* > 0$. Using the definition of $S(e)$ in (2), $\widehat{w}$ in (5), and p^* in (6) the platform's profit function in (7) above may be rewritten as:

$$\Pi(e) = S(e) - (1 - e)\lambda(\theta_H - \theta_L)(\widehat{w} - w_s)$$
$$+ [(1 - e)\lambda\theta_H + (1 - \lambda)\theta_L](d - w) - v. \tag{9}$$

The platform's profits $\Pi(e)$ diverge from social welfare $S(e)$ in two key respects. First, the platform does not internalize the information rents that are enjoyed by each retained H-type firm, $(\theta_H - \theta_L)(\widehat{w} - w_s)$. Second, the platform does not internalize the uncompensated losses suffered by the users, $d - w$.[47] The platform's auditing effort $e^* > 0$ satisfies the first-order condition,

$$\Pi'(e^*) = S'(e^*) + \lambda(\theta_H - \theta_L)(\widehat{w} - w_s) - \lambda\theta_H(d - w) = 0. \tag{10}$$

The first-order condition in (10) underscores that the platform's private incentive to invest in auditing may be either socially excessive or socially insufficient. First, when the platform increases e and removes H-types from the platform, the removed H-types lose their information rents, $\lambda(\theta_H - \theta_L)(\widehat{w} - w_s)$. Auditing imposes a *negative externality* on the H-type firms. Second, when the platform removes H-types from the platform, the user-bystanders get a benefit of $\lambda\theta_H(d - w)$, which is the users' uncompensated loss. Auditing confers a *positive externality* on the user-bystanders. Because there are two offsetting effects, the platform's effort, e^*, may be larger than or smaller than the socially optimal level, e^{**}.

These basic insights are summarized in the following lemma. The proof is in the appendix.

Lemma 1. *Suppose $w_s < \widehat{w}$. The platform sets $p^* = \alpha_L - \theta_L w_s$ and attracts the H-type firms. Let $r_H(w_s) \equiv (\theta_H - \theta_L)(\widehat{w} - w_s)$ denote the H-types' information rents.*

---

[47]Additionally, the platform does not internalize the benefit v for the users. This does not affect the platform's auditing incentives.

1. If $\alpha_H - \theta_H w \geq r_H(w_s)$ then the platform does not audit, $e^* = 0 < e^{**}$.

2. If $\alpha_H - \theta_H w < r_H(w_s)$ then $e^* > 0$. The platform's auditing efforts e^* increase with firm and platform liability, $de^*/dw_s > 0$ and $de^*/dw_p > 0$.

   (a) If $\theta_H(d - w) > r_H(w_s)$ then $0 < e^* < e^{**}$.

   (b) If $\theta_H(d - w) = r_H(w_s)$ then $0 < e^* = e^{**}$.

   (c) If $\theta_H(d - w) < r_H(w_s)$ then $0 < e^{**} < e^*$.

To summarize, when firm liability is below the threshold, $w_s < \widehat{w}$, the H-type firms cannot be deterred from joining the platform. Any price p that attracts the L-type firms to the platform will attract the H-type firms, too.

In case 1 of Lemma 1, the joint liability $w = w_s + w_p$ is small and the platform makes money on each H-type interaction.[48] In this case, the platform welcomes the H-types onto the platform and takes no effort to audit or remove them, $e^* = 0$. The platform is enabling the H-type firms and profiting from their socially harmful activities.

In case 2 of Lemma 1, the joint liability $w = w_s + w_p$ is larger and the platform loses money on each H-type interaction. The platform therefore has a financial incentive to audit and remove the H-types, $e^* > 0$. The platform's incentive to audit is stronger when w_p and w_s are larger. This makes intuitive sense. When platform liability w_p rises, the platform's cost of keeping H-types on the platform rises and so the platform audits more. When firm liability w_s rises, the interaction price that the platform can charge falls, reducing the platform's benefit of retaining the H-type firms.

Finally, and importantly, Lemma 1 establishes that the platform's incentive to audit and remove the H-types may be socially insufficient or socially excessive. In case 2(a) the level of joint liability is small and the platform's investment in auditing is suboptimal, $e^* = 0 < e^{**}$. The platform is not taking into account the positive impact that their investments have on the user-bystanders. In case 2(c) when the level of joint liability is large, then the platform is overly aggressive in its auditing efforts, $e^* > e^{**} > 0$. The reason is that the platform is not taking into account the negative impact that their audit imposes on the H-type firms who are removed from the platform.[49]

Case 2: $w_s > \widehat{w}$. Now suppose that firm liability is above the threshold, $w_s > \widehat{w}$, so the H-type firms are marginal.[50] The platform's profit-maximizing strategy is to either charge $p_L = \alpha_L - \theta_L w_s$ and deter the H-types from joining the platform or charge $p_H = \alpha_H - \theta_H w_s < p_L$ and attract both types. Notably, if the platform chooses the latter strategy and attracts the H-type firms then the platform will not invest in auditing, $e^* = 0$.[51]

---

[48]The platform earns a margin of $p - \theta_H w_p = \alpha_L - \theta_L w_s - \theta_H w_p = \alpha_H - \theta_H w - r_H(w_s) \geq 0$.

[49]Recall that the social welfare function includes the profits of the H-type firms.

[50]The definition of $\widehat{w}$ in (5) implies that the L-type firms obtain higher rents than the H-type firms. For any price p, $\alpha_L - \theta_L w_s - p > \alpha_H - \theta_H w_s - p$.

[51]Attracting the H-types and exerting auditing effort $e > 0$ is a dominated strategy, since the platform can deter the H-types by charging a higher price.

The platform will charge p_H and attract the H-types (instead of charging p_L and deterring the H-types) if

$$\lambda(p_H - \theta_H w_p) + (1 - \lambda)(p_H - \theta_L w_p) > (1 - \lambda)(p_L - \theta_L w_p).$$

Substituting the formulas for p_H and p_L and using the definition of $\widehat{w}$ in equation (5) this condition becomes:

$$\lambda(\alpha_H - \theta_H w) > (1 - \lambda)(\theta_H - \theta_L)(w_s - \widehat{w}). \tag{11}$$

The left-hand side is the joint value of attracting the H-type firms on the platform: the fraction λ of H-types multiplied by the interaction benefit α_H minus the joint liability $\theta_H(w_s + w_p)$. The expression on the right-hand side is the information rent captured by the inframarginal L-types.

We have the following result.

Lemma 2. *Suppose $w_s > \widehat{w}$. Let $r_L(w_s) \equiv (\theta_H - \theta_L)(w_s - \widehat{w})$ denote the L-type firm's information rents.*

1. *If $\lambda(\alpha_H - \theta_H w) > (1 - \lambda)r_L(w_s)$ then the platform sets $p^* = \alpha_H - \theta_H w_s$, attracts the H-type firms, and does not audit, $e^* = 0 < e^{**}$.*

2. *If $\lambda(\alpha_H - \theta_H w) \leq (1 - \lambda)r_L(w_s)$ then the platform sets $p^* = \alpha_L - \theta_L w_s$ and deters the H-type firms.*

If firm liability is above the threshold, $w_s > \widehat{w}$, then the H-type firms are marginal. The platform has the power to deter the H-type firms from joining the platform by raising the price from $\alpha_H - \theta_H w_s$ to $\alpha_L - \theta_L w_s$. In case 1 of Lemma 2, the joint benefit of including the H-types is larger than the information rents captured by the L-type firms. In this case, the platform charges a low price, $p^* = \alpha_H - \theta_H w_s$, welcomes the H-types on the platform and takes no steps to detect or remove them. In case 2 of Lemma 2, the joint benefit of including the H-types is smaller than the L-types' information rents. In this case, the platform has a financial incentive to raise the price to $p^* = \alpha_L - \theta_L w_s$, and deter the H-types from joining the platform.

2.2 Platform Liability

This subsection explores the social desirability and optimal design of platform liability for harm to user-bystanders, taking the level of firm liability w_s as fixed.

We begin by presenting a benchmark where the platform is not liable for the harm, $w_p = 0$. In this benchmark, the platform has no incentive to engage in costly auditing to detect and remove harmful firms from the platform. However, the H-type firms may be deterred from participating on the platform for two reasons. First, the H-type firms face expected liability $\theta_H w_s$. Second, the H-type firms need to pay the platform p per interaction with the users. Thus, the H-types are deterred from joining the platform when $\alpha_H - \theta_H w_s \leq p$.

12

Proposition 1. *(Firm-Only Liability.) Suppose that the platform is not liable for harm to users, $w_p = 0$, and firm liability is $w_s \in (0, d]$. There exists a unique threshold $\widetilde{w} = \widetilde{w}(\lambda) \in \left[\widehat{w}, \frac{\alpha_H}{\theta_H}\right)$, where $\widetilde{w}(\lambda)$ weakly increases in the number of H-types, λ.*[52]

1. *If $w_s < \widehat{w}$ then the platform sets $p^* = \alpha_L - \theta_L w_s$, attracts the H-type firms, and does not invest in auditing, $e^* = 0 < e^{**}$. The platform's auditing incentives are socially insufficient.*

2. *If $w_s \in [\widehat{w}, \widetilde{w})$ then the platform sets $p^* = \alpha_H - \theta_H w_s$, attracts the H-type firms, and does not invest in auditing, $e^* = 0 < e^{**}$. The platform's auditing incentives are socially insufficient.*

3. *If $w_s \geq \widetilde{w}$ then the platform sets $p^* = \alpha_L - \theta_L w_s$ and deters the H-type firms. The first-best outcome is achieved.*

Proposition 1 describes the market outcome when the firms, and only the firms, are liable for the harm to user-bystanders. In case 1, since $w_s < \widehat{w}$ the L-types are marginal. The platform cannot deter the H-types without excluding the L-types. So the platform attracts the H-type firms and does not invest in costly auditing to detect and remove them. This is obviously a socially undesirable outcome.

If firm liability is above the threshold, $w_s \geq \widehat{w}$, then the H-types are marginal. The platform could charge $\alpha_H - \theta_H w_s$ and welcome the H-type firms onto the platform, or charge $\alpha_L - \theta_L w_s$ to deter the H-types from joining. Increasing w_s reduces the joint value of attracting the H-types and therefore motivates the platform to deter them. In case 2, $w_s \in [\widehat{w}, \widetilde{w})$ and the platform charges $p^* = \alpha_H - \theta_H w_s$ and attracts the H-types. In case 3, $w_s \geq \widetilde{w}$ and the platform charges $p^* = \alpha_L - \theta_L w_s$ and deters the H-types.

Should platforms be held liable for the harm suffered by users? Proposition 1 establishes that platform liability is unnecessary when the firms themselves are held sufficiently liable for harm to bystanders, $w_s \geq \widetilde{w}$. In case 3, the first-best outcome is obtained without platform liability. Notice that the threshold $\widetilde{w}$ increases in the number of H-types, λ. When there are many H-type firms and the platform faces no liability, the private and social incentives diverge. The platform will choose to attract the H-type firms because they provide a welcome source of revenue. If $w_s < \widetilde{w}$, the platform makes no effort to audit and prevent socially-harmful interactions.

The next proposition characterizes the optimal platform liability rule, w_p^*.

Proposition 2. *(Optimal Platform Liability.) Suppose firm liability is $w_s \in (0, d]$. The socially-optimal platform liability for harm to users, w_p^*, is as follows:*

1. *If $w_s < \widehat{w}$ then $w_p^* = d - w_s - \left(1 - \frac{\theta_L}{\theta_H}\right)(\widehat{w} - w_s) \in (0, d - w_s)$ achieves the second-best outcome. The platform sets $p^* = \alpha_L - \theta_L w_s$ and attracts the H-type firms. The platform's auditing incentives are socially efficient, $e^* = e^{**}$.*

---

[52]If $\theta_L/\theta_H \geq \alpha_L/\alpha_H$ then $\widetilde{w}(\lambda) = \widehat{w}$ for all λ.

2. *If $w_s \in [\widehat{w}, \widetilde{w})$ then there exists a threshold $\underline{w}_p > 0$ where any $w_p^* \in [\underline{w}_p, d - w_s]$ achieves the first-best outcome. The platform sets $p^* = \alpha_L - \theta_L w_s$ and deters the H-type firms.*

3. *If $w_s \geq \widetilde{w}$ then platform liability is unnecessary. The platform sets $p^* = \alpha_L - \theta_L w_s$ and deters the H-type firms.*

Proposition 2 describes how platform liability can be designed to maximize social welfare. Recall that the platform has two possible mechanisms to reduce the harm to users: the price per interaction p and the audit intensity e. If feasible, the pricing mechanism is privately and socially more efficient than the auditing mechanism, as the pricing mechanism can deter the H-types from joining the platform without the need for costly audits. The pricing mechanism is feasible if and only if firm liability is above a threshold, $w_s \geq \widehat{w}$ (the H-type firms are marginal).

In case 1, the firms' liability is below the threshold ($w_s < \widehat{w}$) and the L-type firms are marginal. From Proposition 1 we know that firm-only liability fails to deter the H-types and gives the platform no incentive to audit and remove the H-types. Imposing liability on the platform motivates the platform to take the socially efficient auditing effort. Notice that the socially-optimal platform liability is positive but less-than-full, $w_p^* \in (0, d - w_s)$. If the platform bears no liability, $w_p = 0$, it would underinvest in auditing; if the platform was held responsible for the full residual harm, $w_p = d - w_s$, then the platform would *overinvest* in auditing. Therefore the second-best outcome is achieved when the platform bears some but not all of the residual damage.

Note that, in case 1, the optimal platform liability, w_p^*, decreases in w_s. From the social planner's perspective, platform liability and firm liability are *substitutes*. Intuitively, when firm liability (w_s) is larger, the H-type firms get less rent, which reduces the platform's auditing incentives; at the same time, the uncompensated harm for users becomes lower and the firms are less willing to pay, which raises the platform's auditing incentives. In equilibrium, the second effect dominates, so the increase in w_s leads to more auditing. To prevent excessive auditing, it is efficient to reduce platform liability. Similarly, when the risk caused by the H-type firms is much larger relative to the risk caused by the L-types (i.e., smaller $\frac{\theta_L}{\theta_H}$), the platform is more likely to invest in auditing. To prevent excessive auditing, it is efficient to reduce platform liability w_p.

In case 2, the firms' liability is in an intermediate range and the H-type firms are marginal. According to Proposition 1, without platform liability, the platform would charge p_H and attract the H-type firms since the joint value of including the H-types (for the platform and the firms) is larger than the L-type firms' rents. Since the firms' rent is independent of w_p while the joint value of keeping the H-types decreases in w_p, the social planner can motivate the platform to raise the price and thus deter the H-types by imposing residual liability on the platform, $w_p^* = d - w_s$. The first-best outcome is obtained.

Finally, in case 3, platform liability is unnecessary when firm liability is sufficiently high. For example, if firms can be forced to pay in full for the harm that they cause, $w_s = d$, then the H-types would be deterred from joining the platform and interacting

14

with users. The first-best outcome is obtained without platform liability. If the firms are judgment proof or immune from liability, however, then platform liability is critical for aligning the private and social incentives.[53]

2.3   Discussion

This section investigated the need for platform liability when the firms that participate on the platform cause harm to user-bystanders. The analysis applies to a variety of settings, such as social-networking platforms where firms pay the platform to access user data or engage in influence activities or spread misinformation. The users enjoy the network benefits associated with the platform for free, but are also subject to the resulting harm. Absent any liability, the firms and the platforms have insufficient incentives to reduce the negative externalities imposed on the user-bystanders.

Our analysis has important implications for the design of liability rules. First, if firms have deep pockets and can compensate the user-bystanders for the harm that they cause, then platform liability is unnecessary. Placing liability on the firms themselves is socially optimal, as it solves the problem of negative externalities. Firms that pose excessive risks to users are deterred from participating on the platform by the threat of future litigation.

Second, if firms are judgment proof and have inadequate resources to compensate victims for their losses or can evade liability in other ways, then platform liability is socially desirable. Holding the platform liable for some or all of the residual harm has two potential benefits. First, the platform may raise the price that it charges to the firms, which will help to deter firms that pose excessive risks to users. Second, the platform will invest resources to detect and remove risky firms from the platform. Interestingly, we show that the socially-optimal level of platform liability may be less than full. When the firms have very limited resources, then holding the platform fully responsible for the residual harm would lead the firm to overinvest in auditing.

Pricing Structure. Our analysis assumed a very simple pricing structure where the platform monetized its activities through an interaction price paid by the firms. Alternatively, we could have assumed that the firms pay a lump-sum membership fee. Our results would be unaffected if the membership fee is paid by the firms that are retained by the platform. With additional instruments, such as a non-refundable application fee or bond, the platform's ability to deter risky firms would be enhanced and the platform could save resources on auditing.[54] The platform's incentives to detect and remove risky firms would

---

[53]In Proposition 2, the second-best outcome is obtained if $w_s < \widehat{w}$ and the first-best outcome is obtained if $w_s \geq \widehat{w}$. Therefore social welfare is weakly increasing in the level of firm liability (w_s).

[54]In practice, many firms have budget constraints so that they could not make a large upfront payment when joining platforms. Theoretically, if the platform can charge a non-refundable lump-sum payment, the H-types may still join. To see this, suppose that the H-types do not join. Then the platform would not take any auditing effort. But anticipating this, the H-types would deviate to join. Therefore, in Case 1, even if the platform can use two-part tariffs, there is no equilibrium where the H-types are fully deterred. Depending on platform liability, there can be two possible equilibria: One where the platform attracts H-types as in the baseline model; the other (a mixed-strategy equilibrium) where the platform

also be enhanced if the users were not bystanders and side payments to the users were possible.[55] The next section considers retail platforms where the users are consumers and shows that our key insights are robust.

False Positives. Our analysis assumed that there were no "false positives." The auditing efforts of the platform did not erroneously detect or remove the L-type firms. Several new insights emerge when the analysis is extended to include false positives. First, the second-best auditing effort is lower than in our baseline model (since it is socially efficient for L-types to remain on the platform). Second, the platform has weaker incentives to invest in auditing than in the baseline model (since the platform loses revenue when it excludes the L-types). Third, the platform's incentives are even weaker relative to the social incentives. When choosing its audit intensity, the platform does not account for the positive externality that excluding the L-types confers on the platform users.[56] It follows that the optimal platform liability w_p^* is larger when there are false positives, compared to our baseline model.

Litigation Costs. Our baseline model assumed that litigation is free, so any user who experiences harm simply files a lawsuit and collects damages from the responsible parties. In reality, bringing a lawsuit is expensive and requires the services of a lawyer. The implications of litigation costs for the design of optimal platform liability is nuanced. On the one hand, when the H-type firms are inframarginal, litigation costs reduce these firms' information rent and raise the users' uncompensated harm, as compared to the baseline model. These effects make the platform's auditing incentives even weaker relative to the social incentives.[57] Moreover, litigation costs may discourage victims from bringing meritorious claims. Without a meaningful threat of litigation, the platform has little incentive to deter and remove harmful firms. Thus, a higher level of liability may be necessary to encourage plaintiffs (and their lawyers) to sue and raise the platform's auditing incentives. On the other hand, insofar as the costs of litigation exceed the benefits of improved platform incentives, a lower level of liability, or indeed the elimination of liability altogether, may be warranted.

Same-Side Network Externalities. Our analysis assumed that the value of the quasi-public good v was both fixed and sufficiently high so that all of the users joined the

---

randomizes on auditing and the H-types randomize on participation.

[55]The socially-optimal platform liability rule maximizes the joint value of the platform, firm, and users. Thus, Coasian bargaining would serve to align the interests of the parties.

[56]Assume that the L-type firms are removed with probability δe where $\delta < 1$. The platform's auditing effort satisfies: $\Pi'(e^*) = S'(e^*) + \lambda(\theta_H - \theta_L)(\widehat{w} - w_s) - [\lambda\theta_H + (1 - \lambda)\delta\theta_L](d - w) = 0$. Comparing this expression to equation (10) reveals that the platform's incentives are weaker relative to the social incentives.

[57]Assume that the litigation costs are k_p, k_s, k_b, respectively for the platform, firms, and users (victims). As long as the victims will bring claims, the platform's auditing effort satisfies: $\Pi'(e^*) = S'(e^*) + \lambda(\theta_H - \theta_L)(\widehat{w} - w_s - k_s) - \lambda\theta_H(d + k_b - w) = 0$. Comparing this expression to equation (10) reveals that the platform's incentives are weaker relative to the social incentives.

platform, regardless of the users' beliefs about platform safety. Our model may be extended to include the user participation decisions and same-side network externalities. If users expect that the platform will allow more H-type firms to join the platform, then fewer users would be willing to join the platform and the value of the quasi-public good would fall. Compared to the baseline model, the second-best auditing effort e^{**} (given the level of platform liability) is weakly higher than before, since society benefits when there are more users (and therefore more interactions). The divergence between the private and social incentives is higher, too. When choosing e^*, the platform does not take into account the impact that the marginal user's participation decision has on the value of the quasi-public good (which is captured by the infra-marginal users). Thus, the optimal platform liability w_p^* may be larger than in the baseline model.

3   Extension: Retail Platforms

We now extend the analysis to consider a retail platform where the firms are the sellers of a product or service and the users are sophisticated consumers. The model will be the same as the baseline model with one important difference: Interactions between the firms and the users are *market transactions* that require the users' consent.

This extension has many practical applications. Most of the products that are bought and sold through the Amazon platform are manufactured and distributed by third-party vendors. Even relatively straightforward products like computer chargers and lightbulbs are of varying quality and safety. The third-party vendors, especially those without existing reputations, would have an incentive to sell products that have low costs but may harm consumers. This problem is particularly severe when the third-party vendors are judgment-proof, and cannot be held accountable for the injuries that their products cause. Extending liability to Amazon gives the platform the incentive to monitor third-party vendors and block dangerous products from reaching the marketplace.

As in the baseline model, there are two types of firm, H and L. The type-i firm produces a good or service at cost c_i which causes accidents with probability θ_i. The unsafe products are cheaper to produce, $c_H < c_L$, and cause harm more frequently, $\theta_H > \theta_L$. The consumer's gross value from the good is α_0. Letting $\alpha_i = \alpha_0 - c_i$, the net interaction value is $\alpha_i - \theta_i d$ as in the baseline model. The timing is the same as the baseline model with one important difference: In stage 4, the firm-sellers are randomly matched with the user-consumers and propose price t. If the consumer accepts the price offer t then the consumer pays t to the firm, and the firm pays p to the platform.[58]

The consumers' willingness to transact with the firms in stage 4 depends on their beliefs about the likelihood that they will suffer harm. The consumers do not observe the safety of the product directly, or the auditing efforts of the platform, but are sophisticated and form beliefs that are, in equilibrium, correct.[59] If the H-type firms seek to join the

---

[58] The results would be the same if the firms pay the platform a percentage of their gross revenue rather than a fixed amount.

[59] One could relax the assumption of full user rationality. If the consumers are unsophisticated, and do

platform and the platform invests e in auditing, the conditional probability of harm per interaction is:

$$E(\theta|e) = \frac{(1-e)\lambda\theta_H + (1-\lambda)\theta_L}{(1-e)\lambda + (1-\lambda)}, \tag{12}$$

which is a decreasing function of e. We let $\theta^{**} = E(\theta|e^{**})$ be the probability of harm when auditing is socially optimal ($e = e^{**}$) and let $\theta^0 = E(\theta|0) = \lambda\theta_H + (1-\lambda)\theta_L$ be the probability of harm when the platform does not audit ($e = 0$).[60] If a user believes that the platform invested $\widehat{e}$ in auditing, then the expected probability of harm from an "average" transaction is $\widehat{\theta} = E(\theta|\widehat{e})$.

If a transaction between a type-i firm and a consumer is consummated, the platform's surplus is $p - \theta_i w_p$, the consumer's surplus is $\alpha_0 - t - \theta_i(d - w_s - w_p)$, and the firm's surplus is $t - (\theta_i w_s + c_i) - p$. Notice that the surplus of the two firm types $i = H, L$ are equal when

$$w_s = \frac{c_L - c_H}{\theta_H - \theta_L} = \frac{\alpha_H - \alpha_L}{\theta_H - \theta_L} = \widehat{w}. \tag{13}$$

The threshold $\widehat{w}$ is exactly the same threshold as in (5) above.

Assumption A2 implies that even if the platform does not audit at all, the gross profit for the L-type firms (before paying p to the platform) is positive. Thus, this assumption guarantees that an equilibrium exists for all assignments of liability, w_s and w_p.

3.1   Equilibrium Analysis

We will now characterize the prices t and p paid by users and firms and the audit intensity e of the platform. The price t that the firms can charge to the user-consumers depends on the users' beliefs about the probability of harm, $\widehat{\theta}$.

There is no separating equilibrium where the H-types and L-types charge different prices and have positive sales. If a separating equilibrium existed, the users would have the correct belief of the firms' types. But given $\alpha_H - \theta_H d < 0$, the users and the H-types would not have interactions.[61]

In any pooling (or semi-pooling) equilibrium, the equilibrium price t^* paid by users cannot be larger than the users' maximum expected willingness to pay. We will construct equilibria with

$$t^* = \alpha_0 - \widehat{\theta}(d - w), \tag{14}$$

so consumer surplus is zero. The consumers believe that any firm charging a different price would have at least the average probability of harm, $\widehat{\theta}$. In this equilibrium, no firm

---

not anticipate future harm from using the products sold through the platform, the model would resemble the model about bystanders in the previous section.

[60]e^{**} is defined in equation (3).

[61]It is possible to have a separating equilibrium where the platform deters all the H-types through the pricing mechanism.

has incentives to raise its price, as otherwise the users would not buy from the firm.[62] Assumption A2 guarantees that $t^* > 0$.

Case 1: $w_s < \widehat{w}$. Suppose $w_s < \widehat{w}$ so the L-type is marginal. The platform sets the price to extract rents from the L-type firm, $p^* = t^* - (\theta_L w_s + c_L)$.[63] Using the expression for t^* in (14) above and the fact that $\alpha_L = \alpha_0 - c_L$ gives

$$p^* = \alpha_L - \theta_L w_s - \widehat{\theta}(d - w). \tag{15}$$

Comparing this expression to the baseline model with bystanders in (6) reveals an important difference: the price (15) reflects the user-consumers' beliefs about the uncompensated harm, $\widehat{\theta}(d - w)$. As before, the platform's profits are

$$\Pi(e) = (1 - e)\lambda(p^* - \theta_H w_p) + (1 - \lambda)(p^* - \theta_L w_p) - c(e). \tag{16}$$

We now explore the platform's equilibrium choice of auditing effort, e^*. A necessary condition for an equilibrium with $e^* > 0$ is that the platform loses money on each retained H-type. Using the formulas for p^* from (15) and $\widehat{w}$ from (5), the platform loses money on each retained H-type, $p^* - \theta_H w_p < 0$, if and only if

$$(\alpha_H - \theta_H d) + (\theta_H - \widehat{\theta})(d - w) - (\theta_H - \theta_L)(\widehat{w} - w_s) < 0. \tag{17}$$

It is interesting to compare (17) to the analogous condition in the baseline model with bystanders (8). The first and the third terms are the same but the second terms are different. In (8), $\theta_H(d - w)$ is the user's uncompensated harm. In (17), $(\theta_H - \widehat{\theta})(d - w)$ is the user's *uncompensated harm beyond their expectations*. Since the users are consumers instead of bystanders, the price that the user-consumers pay, t^* (and accordingly the platform's price p^*), reflects the user-consumers' expected uncompensated harm, $\widehat{\theta}(d - w)$.

We now explore how the private and social incentives for auditing diverge when $e^* > 0$. Substituting p^* from (15), $S(e)$ from (2), and $\widehat{w}$ from (5), allows us to rewrite the platform profits as

$$\Pi(e) = S(e) - (1 - e)\lambda(\theta_H - \theta_L)(\widehat{w} - w_s)$$
$$+ [(1 - e)\lambda(\theta_H - \theta^*) + (1 - \lambda)(\theta_L - \widehat{\theta})](d - w) - v. \tag{18}$$

The platform's profits $\Pi(e)$ diverge from social welfare $S(e)$ for two reasons. First, the platform does not internalize the information rents that are enjoyed by each retained H-type firm, $(\theta_H - \theta_L)(\widehat{w} - w_s)$. Second, the platform does not internalize the users'

---

[62]This equilibrium maximizes the platform's profits. See the proof of Lemma 3. Other equilibria may exist. For example, any price $t \in (\alpha_0 - \theta_H(d - w), \alpha_0 - \widehat{\theta}(d - w))$ can be an equilibrium if the users hold the off-equilibrium belief that any firm charging a different price would be the H-type. However, in such equilibria, firms are playing a dominated strategy: their profits would be higher if they raise the prices.

[63]See the proof of Lemma 3.

unanticipated losses or gains (i.e., the term in the second line of (18)).[64] Since the users cannot observe e, the platform's off-the-equilibrium-path choice of auditing may diverge from the users' expectations.[65]

Differentiating (18) with respect to e, the firm's equilibrium auditing effort e^* satisfies

$$\Pi'(e^*) = S'(e^*) + \lambda(\theta_H - \theta_L)(\widehat{w} - w_s) - \lambda(\theta_H - \theta^*)(d - w) = 0, \tag{19}$$

where $\theta^* = E(\theta|e^*)$ are the equilibrium beliefs. Equation (19) shows that the platform's private incentive to invest in auditing may be either socially excessive or socially insufficient. First, when the platform increases e and removes H-types from the platform, the removed H-types lose their information rents, a loss of $\lambda(\theta_H - \theta_L)(\widehat{w} - w_s)$. Auditing imposes a negative externality on the H-types. Second, when the platform audits more and removes H-types from the platform, the users get a benefit of $\lambda(\theta_H - \theta^*)(d - w)$. Auditing confers a positive externality on the users. Because there are two offsetting effects, the platform's choice of effort, e^*, may be smaller or larger than the socially optimal level, e^{**}. The platform's incentive to invest in auditing is socially insufficient ($e^* < e^{**}$) if and only if the H-type firms' rent, $\lambda(\theta_H - \theta_L)(\widehat{w} - w_s)$, is smaller than the loss to the users $\lambda(\theta_H - \theta^{**})(d - w)$ where θ^{**} are the posterior market beliefs if $e = e^{**}$.

We then have the following result:

Lemma 3. *Suppose $w_s < \widehat{w}$. The platform sets $p^* = \alpha_L - \theta_L w_s - \theta^*(d - w)$ and attracts the H-type firms where $\theta^* = E(\theta|e^*)$ are the equilibrium posterior beliefs. Let $\theta^{**} = E(\theta|e^{**})$, $\theta^0 = E(\theta|0)$, and $r_H(w_s) = (\theta_H - \theta_L)(\widehat{w} - w_s)$.*

1. *If $(\alpha_H - \theta_H d) + (\theta_H - \theta^0)(d - w) \geq r_H(w_s)$ then the platform does not audit, $e^* = 0 < e^{**}$.*

2. *If $(\alpha_H - \theta_H d) + (\theta_H - \theta^0)(d - w) < r_H(w_s)$ then $e^* > 0$. The platform's auditing effort decreases in firm liability $de^*/dw_s < 0$ and increases in platform liability $de^*/dw_p > 0$.*

   (a) *If $(\theta_H - \theta^{**})(d - w) > r_H(w_s)$ then $0 < e^* < e^{**}$.*
   (b) *If $(\theta_H - \theta^{**})(d - w) = r_H(w_s)$ then $0 < e^* = e^{**}$.*
   (c) *If $(\theta_H - \theta^{**})(d - w) < r_H(w_s)$ then $0 < e^{**} < e^*$.*

When firm liability is below the threshold, $w_s < \widehat{w}$, the H-type firms are infra-marginal and cannot be deterred from joining the platform. Since the price paid by the users reflects their beliefs of the average probability of harm, which is lower than the probability of harm caused by the H-type firms, the joint benefit (for the platform and the H-type firms) of

---

[64]If the platform could commit to e then they would internalize the users' losses and gains. With commitment, the platform's auditing incentives may be socially excessive but not socially insufficient.

[65]If $e < e^*$ ($e > e^*$) then the users experience an unanticipated loss (gain) and expression in the second line of (18) is negative (positive).

keeping the H-types can be positive. The platform makes profits on each H-type firm as long as this joint benefit is larger than the firm's rent.

In case 1 of Lemma 3, the joint liability $w = w_s + w_p$ is small and the platform is making money on each H-type firm. In this case, the platform welcomes the H-types onto the platform and makes no attempt to audit or remove them, $e^* = 0$. The platform is enabling the H-type firms and profiting from their socially-harmful activities. In case 2 of Lemma 3, the joint liability $w = w_s + w_p$ is larger and the platform would lose money on each H-type firm. The platform would therefore have a financial incentive to audit and remove the detected H-types, and so $e^* > 0$. Importantly, the platform's incentive to audit and remove the H-types may be socially insufficient or socially excessive. Finally, in case 2, the platform's incentive to audit is stronger when w_p is larger but weaker when w_s is larger.

This latter result, that the platform audits less when w_s is larger, is different from the result in the baseline model where users are bystanders. In Lemma 1, when users are bystanders, the platform's auditing incentive is larger if the uncompensated harm to users, $\theta_H(d - w)$ is lower and/or the H-types' information rent $r_H(w_s)$ is higher. Given a marginal increase in w_s, the drop in *the uncompensated harm* is greater than the drop in the firms' rent. Accordingly, the platform takes more auditing effort. By contrast, in Lemma 3 when users are consumers, the price paid to the firms reflects the users' beliefs of the probability of harm. Thus, the platform's auditing incentive now depends on the uncompensated harm beyond the users' expectation, $(\theta_H - \widehat{\theta})(d - w)$, and the H-types' information rent $r_H(w_s)$. Given a marginal increase in w_s, the drop in *the uncompensated harm beyond the users' expectation* is less than the drop in the firms' rent. Accordingly, the platform takes less auditing effort.

Case 2: $w_s > \widehat{w}$. Now suppose that $w_s > \widehat{w}$. In this case, the H-types are marginal. The platform will choose between a high price, under which only the L-types join the platform, and a lower price under which both types join the platform.

Suppose that the platform sets a high price and deters the H-type firms from joining the platform. Consumers understand that the H-types are deterred, $\widehat{\theta} = \theta_L$, and so the firms charge the consumers $t = \alpha_0 - \theta_L(d - w)$. The platform charges the firms a transaction price $p = t - (\theta_L w_s + c_L)$ or $p = \alpha_L - \theta_L(d - w_p)$. The platform's profit when it deters the H-types and attracts the L-types is $(1 - \lambda)(p - \theta_L w_p)$ or

$$(1 - \lambda)(\alpha_L - \theta_L d). \tag{20}$$

In other words, the platform extracts all of the social surplus associated with the transactions between users and the L-type firms.

Suppose instead that the platform sets a low price and attracts the H-type firms. As in the previous section where users were bystanders, note that the platform has no incentive to audit and remove the H-types from the platform.[66] The platform's profits

---

[66]This is by revealed preference, as it could deter the H-types without any auditing costs by raising the price.

will be strictly lower if they attract the H-types. To see why, observe that the incremental social benefit of attracting the H-type firms is negative, $\lambda(\alpha_H - \theta_H d) < 0$. If the platform attracts the H-types, then the consumers, firms, and platform are jointly worse off. In equilibrium, the consumers are compensated for purchasing the less safe products and the L-type firms capture rents. Therefore the platform's incremental profit from attracting the H-types is unambiguously negative.[67]

Lemma 4. *Suppose $w_s > \widehat{w}$. The platform sets $p^* = \alpha_L - \theta_L(d - w_p)$ and deters the H-type firms.*

In the baseline model of Section 2 where the users are bystanders, given $w_s > \widehat{w}$, the platform may (inefficiently) attract the H-type firms if the joint value for the platform and firms is larger than the firms' rent. By contrast, when the users are consumers, the platform always deters the H-types, because the firms' price (and accordingly the platform's charge over firms) incorporates the consumers' willingness to pay.

3.2   Platform Liability

Now we explore the social desirability and optimal design of platform liability for harm to user-consumers. We begin by presenting a benchmark where the platform is not liable for harm to users, $w_p = 0$. Although the platform does not take auditing effort, $e^* = 0$, the platform may deter the H-types from the platform through the price charged to the firms.

Proposition 3. *(Firm-Only Liability.) Suppose that the platform is not liable for harm to consumers, $w_p = 0$, and firm liability is $w_s \in (0, d]$. Let $\theta^0 = E(\theta|0)$.*

1. *If $w_s < \widehat{w}$ the platform sets $p^* = \alpha_L - \theta_L w_s - \theta^0(d - w_s)$, attracts the H-type firms, and does not invest in auditing, $e^* = 0 < e^{**}$. The platform's auditing incentives are socially insufficient.*

2. *If $w_s \geq \widehat{w}$ then the platform sets $p^* = \alpha_L - \theta_L d$ and deters the H-type firms. The first-best outcome is obtained.*

This result is intuitive. When $w_s < \widehat{w}$ and $w_p = 0$, the platform's auditing incentive is socially insufficient. When $\widehat{w} \geq w_s$, Lemma 4 implies that the platform deters the H-types.

The next proposition characterizes the socially optimal platform liability, w_p^*.

Proposition 4. *(Optimal Platform Liability.) Suppose firm liability is $w_s \in (0, d]$. Let $\theta^{**} = E(\theta|e^{**})$. The socially-optimal platform liability for harm to consumers, w_p^*, is as follows:*

---

[67]See the proof of Lemma 4.

1. *If $w_s < \widehat{w}$ then $w_p^* = d - w_s - \left(\frac{\theta_H - \theta_L}{\theta_H - \theta^{**}}\right)(\widehat{w} - w_s) \in (0, d - w_s)$ achieves the second-best outcome. The platform sets $p^* = \alpha_L - \theta_L w_s - \theta^{**}(d - w_s)$ and attracts the H-type firms. The platform's auditing incentives are socially efficient, $e^* = e^{**}$.*

2. *If $w_s \geq \widehat{w}$ then platform liability is unnecessary. The platform sets $p^* = \alpha_L - \theta_L(d - w_p)$ and deters the H-type firms.*

If $w_s < \widehat{w}$ then the L-types are marginal. The platform cannot deter the H-types directly through the price, but can remove them through auditing. If the platform was responsible for the residual harm, $w_p = d - w_s$, then the platform would overinvest in auditing. It is socially efficient to have the platform bear some but not all the residual damage.

Given $w_s < \widehat{w}$, the optimal platform liability, w_p^*, increases in w_s. From the social planner's perspective, platform liability and firm liability are *complements*. Recall that in Proposition 2 where the users are bystanders, platform liability and firm liability were substitutes. To see the intuition for this difference, note that when users are bystanders, the optimal platform liability satisfies

$$(\theta_H - \theta_L)(\widehat{w} - w_s) = \theta_H(d - w_s - w_p^*). \tag{21}$$

A marginal increase in w_s reduces the firms' rent (the left-hand side) but also reduces *the uncompensated harm for users* (the right-hand side). The latter effect dominates, so the increase in w_s raises the platform's auditing incentives. To prevent excessive auditing, the optimal platform liability should be adjusted lower.

In Proposition 4, when users are consumers, the optimal platform liability satisfies

$$(\theta_H - \theta_L)(\widehat{w} - w_s) = (\theta_H - \theta^{**})(d - w_s - w_p^*). \tag{22}$$

A marginal increase in w_s reduces the firms' rent (the left hand side) but also reduces *the user-consumers' uncompensated harm beyond their expectation* (the right-hand side). The first effect dominates. So the increase in w_s reduces the platform's auditing incentives. To prevent insufficient auditing, the optimal platform liability should be larger.

Corollary 1. *Suppose $w_s < \widehat{w}$. When the users are bystanders, the optimal platform liability decreases in w_s; when the users are consumers, the optimal platform liability increases in w_s.*

3.3 Discussion

This section explored the role of platform liability in retail settings, where sellers interact with sophisticated consumers through consensual market transactions. Although the consumers cannot observe product safety directly, they form rational inferences which, in equilibrium, are correct. Holding all else equal, a consumer's willingness to pay for a product is higher if the consumer believes that the product is safer. Thus, even absent

liability, the platform has private incentives to assure higher product safety to stimulate consumer demand. As in our baseline model with user-bystanders, the platform's incentive to audit and remove the H-type firms may be socially insufficient or excessive. The private and social incentives are in greater alignment, however, so the optimal level of platform liability is smaller for retail platforms.

Consumer Precautions. In some settings, the consumers of potentially harmful products can take pre- and post-sale precautions to mitigate the harm. Before even adding a product to their Amazon shopping cart, a shopper can read the product reviews posted by others and can check the Consumer Product Safety Commission website for warnings and recalls.[68] Consumers can take further precautions after receiving the item to reduce the risk of harm. For example, a hoverboard buyer can take care when charging the device to reduce the likelihood of an electrical fire. Placing strict products liability on the platform, regardless of the care taken by consumers, could create a perverse incentive for the consumers to be careless. The optimal design of platform liability must strike a balance between creating incentives for the platform to detect and remove harmful products and creating incentives for consumers to be prudent.

Naive Consumers. The analysis in this section assumed that the consumers were fully rational and forward looking. In our equilibrium, the consumers correctly anticipate the audit intensity e and view the expected damage award $E(\theta|e)w$ as a rebate of the purchase price. In reality, consumers may not be fully aware of the safety hazards posed by products or the opportunities for future litigation. Our model may be easily adapted to consider naive consumers. Indeed, if the consumers are totally unaware of product risks, then the analysis closely follows our baseline model where users are bystanders. If a consumer is unaware of product risks, then each consensual transaction imposes a *negative externality on the consumer's future self.* Since the consumer's future self is essentially a non-consenting "bystander" to the transaction, the analysis of the baseline model and all of its implications apply.[69] The case for holding retail platforms is stronger when consumers are naive and unaware of product risks.

Information Disclosure. The analysis above assumed that the platform removed discovered H-types from the platform. What would happen if the platform is required to disclose the audit results to the consumers, and the consumers decided for themselves whether to interact with the known H-types? Absent platform liability ($w_p = 0$), a rational consumer would decline to interact with a known H-type ex post.[70] Although ex post

---

[68]See for example https://www.cpsc.gov/Recalls/2018/Amazon-Recalls-Portable-Power-Banks-Due-to-Fire-and-Chemical-Burn-Hazards-Recall-Alert

[69]A naive buyer would be willing to pay $t = \alpha_0$ for the good (regardless of their beliefs about audit intensity e). The firm's surplus is $t - (\theta_i w_s + c_i) - p = \alpha_0 - (\theta_i w_s + c_i) - p = \alpha_i - \theta_i w_s - p$, exactly as in equation (4) in the baseline model.

[70]The consumer's net surplus is $\alpha_0 - t^* - \theta_H(d - w_p - w_s)$ and the H-type's net surplus $t^* - p^* - c_H - \theta_H w_s$. Their joint surplus is $\alpha_H - \theta_H(d - w_p) - p^*$. If $w_p = 0$ then the joint surplus is negative, $\alpha_H - \theta_H d - p^* < 0$.

24

efficiency would be obtained without platform liability, the platform would have insufficient incentives to audit the sellers ex ante.[71] At the other extreme, with full platform liability ($w_p = d$), a rational consumer would interact with a known H-type.[72] Since disclosure would not deter harmful interactions, the platform's incentive to exert effort to identify the H-types is insufficient in this case, too. These observations underscore the importance of granting retail platforms the discretion to remove bad actors rather than relying on disclosure alone.

4   Extension: Platform Competition

The baseline model considers a monopoly platform, while competing platforms exist in some industries. Regulators in the US and the EU are concerned about anti-competitive strategies used by platforms. How would platform competition change the social desirability and optimal design of platform liability? In this section, we extend our baseline model (where users are bystanders) by considering two competing platforms, Platform 1 and Platform 2. We will show that platform competition may raise or reduce social welfare and the optimal platform liability may be higher or lower as compared to that in the baseline model.

Denote the platforms' prices and auditing effort as p_j and e_j, $j = 1, 2$. Users can join both platforms, but each firm can only join one platform.[73] Thus, the platforms compete for firms but not for users. In stage 1, the platforms set their prices simultaneously. The assumptions and timing are otherwise identical to the baseline model.

4.1   Equilibrium Analysis

Case 1: $w_s < \widehat{w}$. Suppose that $w_s < \widehat{w}$, so the L-type firms are marginal. The platforms cannot deter the H-types from joining the platform without also deterring the L-types. We show in the appendix that, as long as both platforms are active, they have zero profits, charge the same price, and take the same auditing effort in any (symmetric or asymmetric) equilibrium. Without loss of generality, we focus on the symmetric equilibrium where each platform attracts half of the firms. The equilibrium price $p^c = p_1 = p_2$ and auditing effort $e^c = e_1 = e_2$ (if it is an interior solution) satisfy

$$\tfrac{1}{2}\Pi(e^c) = \tfrac{1}{2}\{(1 - e^c)\lambda(p^c - \theta_H w_p) + (1 - \lambda)(p^c - \theta_L w_p) - c(e^c)\} = 0, \qquad (23)$$

$$-\lambda(p^c - \theta_H w_p) - c'(e^c) = 0. \qquad (24)$$

---

[71]If consumers are naive and underestimate product risks then the platform's incentive to audit and disclose negative information would be further diluted. Recent empirical work by Culotta et al. (2022) shows that Airbnb may limit the flow of negative safety reviews.

[72]If $w_p = d$ then the consumer and seller's joint surplus is positive, $\alpha_H - p^* > 0$. The accident losses are externalized on the platform.

[73]In practice, many firms choose single-homing due to fixed costs or reputation concerns.

If there is no platform liability, $w_p = 0$, then (23) and (24) imply that the platforms charge $p^c = 0$ and do not waste resources auditing the firms, $e^c = 0$. If there is platform liability, $w_p > 0$, then the platforms will engage in costly auditing, $e^c > 0$. To see why, suppose to the contrary that $w_p > 0$ and the platforms do not audit, $e^c = 0$. Then $c(e^c) = 0$ and the zero-profit condition (23) implies $p^c - \theta_H w_p < 0 < p^c - \theta_L w_p$. Since the platforms are losing money on each retained H-type, condition (24) implies that the platforms would invest $e^c > 0$ to detect and remove the H-types, a contradiction. So, if platforms are liable, $w_p > 0$, the platforms invest in auditing, $e^c > 0$.

Interestingly, platform competition increases the platforms' auditing incentives. The reason is that since the platforms compete to serve the L-type firms, the equilibrium price is lower than in the baseline model of monopoly, $p^c < p^* = \alpha_L - \theta_L w_s$. This implies that the price-cost margins from serving the H-type firms is lower, too. Therefore the platforms have a greater incentive to detect and remove the H-types. Condition (24) confirms that $de^c/dp^c < 0$. Therefore, since $p^c < p^*$ we have $e^c > e^*$. One can also show that e^c increases in w_p: if platform liability increases, the platforms take greater effort to detect and remove the harmful firms.[74]

The following lemma summarizes these results. In the lemma, e^* are the monopoly auditing incentives as defined in Lemma 1 and w_p^* is the optimal platform liability for a monopoly as defined in Proposition 2.

Lemma 5. *Suppose $w_s < \widehat{w}$. The platforms set $p^c < \alpha_L - \theta_L w_s = p^*$ and attract the H-type firms. If $w_p = 0$ then the platforms do not audit, $e^c = e^* = 0$. If $w_p > 0$ then $e^c > 0$, and $de^c/dw_p > 0$. There exists a unique threshold $\overline{w}_p \in (0, w_p^*)$.*

1. *If $0 < w_p \leq \overline{w}_p$ then $e^* < e^c \leq e^{**}$.*

2. *If $w_p \in (\overline{w}_p, w_p^*)$ then $e^* < e^{**} < e^c$.*

3. *If $w_p \geq w_p^*$ then $e^{**} \leq e^* < e^c$.*

Lemma 5 implies that the competing platforms' incentives to audit and remove the harmful H-type firms may be socially insufficient or socially excessive. When platform liability is lower than a threshold, $w_p < \overline{w}_p$ then the platforms' incentives are socially insufficient, $e^c < e^{**}$. When platform liability is above the threshold, $w_p > \overline{w}_p$ then the platforms' incentives are socially excessive, $e^c > e^{**}$.

Lemma 5 also implies that competition between platforms may raise or lower social welfare. As noted above, competing platforms spend more resources to detect and remove the harmful H-type firms than a monopoly platform: if $w_p > 0$ then $e^c > e^*$. This can be a very good thing if platform liability w_p is low.[75] When platform liability is sufficiently high, $w_p > w_p^*$, then the monopoly incentives are excessive (see Lemma 1). In this case, competition compounds the excessive auditing incentives and social welfare falls.

---

[74] See the proof in the Appendix.

[75] If platform liability w_p is very low, then the monopoly incentives are socially insufficient and platform competition raises social welfare relative to monopoly.

Case 2: $w_s > \widehat{w}$ Now suppose that $w_s > \widehat{w}$. In this case, the H-type firms are marginal.

Lemma 6 characterizes the unique equilibrium of the game. The formal proof is presented in the appendix, and we summarize the main ideas here. In the equilibrium, the platforms earn zero profits. When the H-type firms' interaction surplus is above a threshold, $\alpha_H - \theta_H w_s \geq p^c$, then the platforms charge p^c and invest e^c, the same values defined in (23) and (24) above. In the unique competitive equilibrium, the H-types participate and earn positive surplus, and the platforms break even. This is case 1 of Lemma 6.

If the H-type's interaction surplus is below a threshold, $\alpha_H - \theta_H w_s < p^c$, then there does not exist an equilibrium with full H-type participation. It is easy to see why. If all of the H-types participated, then the platforms would need to charge p^c just to break even, violating the H-types' participation constraint. In case 2 of Lemma 6, some but not all of the H-types participate and the equilibrium price and auditing efforts render the H-types indifferent between participating and not. In case 3 of Lemma 6, when the interaction surplus is sufficiently small, the platforms will set their prices to break even on the L-types, $p_1 = p_2 = \theta_L w_p$, and the H-types are deterred from participating.

Lemma 6. *Suppose $w_s > \widehat{w}$.*

1. *If $\alpha_H - \theta_H w_s \geq p^c$ then the platforms set $p_1 = p_2 = p^c$, attract the H-types, and take auditing effort e^c. If $w_p = 0$ then $e^c = 0$; If $w_p > 0$ then $e^c > 0$, and $de^c/dw_p > 0$.*

2. *If $\alpha_H - \theta_H w_s \in [\theta_L w_p, p^c)$ then the platforms set $p_1 = p_2 = \alpha_H - \theta_H w_s$, attract some but not all of the H-types, and take auditing effort $e < e^c$.*

3. *If $\alpha_H - \theta_H w_s < \theta_L w_p$, the platforms set $p_1 = p_2 = \theta_L w_p$ and deter the H-type firms.*

4.2   Platform Liability

Given Lemma 5 and Lemma 6, we have the following observations for the setting where the platform is not liable for harm to the users, $w_p = 0$.

Proposition 5. *(Firm-Only Liability with Platform Competition) Suppose that the platforms are not liable for harm to bystanders, $w_p = 0$, and firm liability is $w_s \in (0, d]$.*

1. *If $w_s \leq \frac{\alpha_H}{\theta_H}$ then the platforms set $p_1 = p_2 = p^c = 0$, attract the H-type firms, and do not invest in auditing, $e^c = 0 < e^{**}$. The platforms' auditing incentives are socially insufficient.*

2. *If $w_s > \frac{\alpha_H}{\theta_H}$ then the platforms set $p_1 = p_2 = \theta_L w_p = 0$ and deter the H-type firms. The first-best outcome is obtained.*

Comparing this result to our baseline model with a monopoly platform reveals some important differences. Under platform competition, Proposition 5 holds that if $w_p = 0$ then the platforms attract the H-type firms and invest nothing in auditing for all $w_s < \frac{\alpha_H}{\theta_H}$.

27

In the baseline model with a monopoly platform, Proposition 1 showed that the platform deters the H-type firms for all $w_s \in [\widetilde{w}, \frac{\alpha_H}{\theta_H})$. The implication is that social welfare is (weakly) lower when platforms compete with each other. The superiority of the monopoly platform arises because the monopoly has a financial incentive to raise the price and exclude the H-types from the market. So, in this case, the monopolist's incentives are more closely aligned with social welfare.

Following Lemma 5 and Lemma 6, the next proposition characterizes the socially-optimal platform liability when there is competition.

Proposition 6. *(Optimal Platform Liability with Platform Competition) Suppose that firm liability is $w_s \in (0, d]$. The socially-optimal liability for the competing platforms, w_p^c, is as follows:*

1. *If $w_s < \widehat{w}$ then there exists a unique $w_p^c < w_p^* < d - w_s$ that achieves the second-best outcome. The platforms set $p_1 = p_2 = p^c < p^*$ and attract the H-type firms. The platforms' auditing incentives are socially efficient, $e^* = e^{**}$.*

2. *If $w_s \in [\widehat{w}, \frac{\alpha_H}{\theta_H}]$ then any $w_p^c \in (\frac{\alpha_H - \theta_H w_s}{\theta_L}, d - w_s]$ achieves the first-best outcome. The platforms set $p_1 = p_2 = \theta_L w_p < p^*$ and deter the H-type firms.*

3. *If $w_s > \frac{\alpha_H}{\theta_H}$ then platform liability is unnecessary. The platforms set $p_1 = p_2 = \theta_L w_p$ and deter the H-type firms.*

Comparing Proposition 6 to Proposition 2, we can observe how competition changes the optimal platform liability.

If the level of firm liability is small, $w_s < \widehat{w}$, the H-type firms cannot be deterred from joining the platform. Platform liability can raise social welfare by encouraging the platforms to detect and remove the H-type firms from the platform. Interestingly, the socially-optimal level of platform liability with platform competition is lower than that in the baseline model of monopoly, $w_p^c < w_p^*$. The reason for this result is that the prices are lower with competition, $p^c < p^*$, and so the platform's surplus from retaining the harmful H-types, $p^c - \theta_H w_p$ is lower too. Therefore the platforms' incentives to detect and remove the harmful H-type firms is stronger. To prevent excessive auditing, platform liability should be lower with platform competition.

If firm liability is in an intermediate-range, $w_s \in [\widehat{w}, \frac{\alpha_H}{\theta_H}]$, then the H-type firms are marginal. In this case, platform liability can induce the platforms to raise their prices and thus deter the H-types from participating. Comparing Proposition 6 to Proposition 2 reveals that the lowest platform liability that implements the first-best outcome under competition is $\frac{\alpha_H - \theta_H w_s}{\theta_L}$, while the lowest liability that induces a monopoly platform to deter the H-type firms is either $\underline{w}_p \leq \frac{\alpha_H - \theta_H w_s}{\theta_L}$ or 0.[76] In this case, competition reduces the platforms' incentives to deter the risky firms. So, platform liability should be (weakly) larger with platform competition.

---

[76] As defined in the proof of Proposition 2 $\underline{w}_p = \frac{\alpha_H}{\theta_H} - w_s - \frac{1-\lambda}{\lambda}\left(1 - \frac{\theta_L}{\theta_H}\right)(w_s - \widehat{w})$, which is strictly lower than $\frac{\alpha_H - \theta_H w_s}{\theta_L}$ as long as $\alpha_H - \theta_H w_s > 0$.

Finally, if firm liability is high, $w_s > \frac{\alpha_H}{\theta_H}$, the H-type firms' gross interaction value from participation is negative and they would never choose to join the platform. In other words, firm-only liability is sufficient to achieve full deterrence. Platform liability is unnecessary no matter whether there is competition or not.

4.3 Discussion

This section examined how platform competition would change the social desirability and optimal design of platform liability. Recall that platforms have two possible mechanisms to deter or remove the H-type firms: the price per interaction p and the audit intensity e. Competition reduces the equilibrium prices, so that the pricing mechanism becomes less effective. Thus, when the firms are modestly judgment proof, as compared to the baseline model, greater platform liability is needed to motivate the platforms to raise prices and deter the H-types. However, when the firms are very judgment proof (so the pricing mechanism is not effective at all), competition reduces the price-cost margin and therefore enhances the platforms' auditing incentives. In this case, the optimal platform liability should be lower than in the baseline model of monopoly.

These observations suggest that policies encouraging platform competition should be complemented by changes in platform liability. A report written by Cremer, et al. and published by the European Commission (2019) raised concerns about increased concentration in platform markets.[77] The anti-trust authories in both the U.S. and EU have initiated investigations and lawsuits against platforms. For example, the Federal Trade Commission in the U.S. filed a lawsuit against Facebook, asking the court to force it to sell WhatsApp and Instagram.[78] The potential changes in market competition would affect platforms' incentives to deter or remove harmful firms, which would call for changes in platform liability.

In fact, the Digital Services Act and Digital Markets Act proposed in 2020 by the European Commission try to achieve the two goals together: creating a safer digital space and establishing a level playing field (to foster innovation and competitiveness).[79] Holding platforms liable for user harm can improve safety in the digital space. However, our analysis implies that, if these policies increase platform competition, the socially optimal platform liability could be higher or lower (depending on the extent to which the harmful firms are judgment proof).

Single-Homing Users. Our analysis assumed that users could join both platforms. In some applications, users may join only one platform due to switching costs. Suppose that a certain proportion of users are single-homing. Then the platforms would compete for these users as well, which would raise their incentives to deter or remove the H-type firms. Accordingly, the optimal platform liability can be lower than in the earlier analysis.

---

[77]See https://ec.europa.eu/competition/publications/reports/kd0419345enn.pdf

[78]See https://www.reuters.com/technology/us-ftc-says-court-should-allow-antitrust-lawsuit-against-facebook-go-forward-2021-11-17/

[79]See https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package

However, if the platforms could differentiate their quasi-public goods, more users would be single-homing. Given the reduced competition, the optimal liability would be higher.

5 Conclusion

Should platforms be held liable for the harms suffered by platform participants? This question is of practical as well as academic interest. Platforms like Amazon, Google, and Facebook create considerable social value for their users but may also expose them to considerable risk. These and other platforms claim that they value their users' privacy and safety, are careful to protect their users' sensitive personal information, and spend considerable sums of money to monitor platform activity and block harmful actors from participating. But in reality, platforms in the United States and abroad face lax regulatory oversight from public enforcement agencies and are largely immune from private litigation.

We explored the social desirability of platform liability in a two-sided platform model where firms impose cross-side harms on users. The model, while very simple, underscores several key insights. First, if firms have sufficiently deep pockets, and are held fully accountable for the harms they cause, then platform liability is unnecessary. Holding the firms (and only the firms) liable deters the harmful firms from joining the platform and interacting with users. If firms are judgment proof and immune from liability, however, then platform liability is socially desirable. With platform liability, the platform has an incentive to (1) raise the interaction price to deter the harmful firms and (2) invest resources to detect and remove the harmful firms from the platform. The optimal level of platform liability depends on whether users are involuntary bystanders or voluntary consumers of the firms and on the intensity of platform competition. With appropriate incentives, platforms can play an important role in reducing social costs.

Our model abstracted from other salient factors. First, we did not consider ex ante incentives for innovation. Section 230 of the Communications Decency Act was adopted to allow the internet to grow and flourish, and has been referred to as "the one line of federal code that has created more economic value in this country than any other."[80] Do Facebook, Google, and Amazon still require the protection of Section 230? Second, our model abstracted from reputation building and peer-to-peer reviews. Is platform liability a substitute or a complement for decentralized market mechanisms?[81] Third, we assumed that the platform could audit and remove participants from the platform. Should a platform that maintains tight control be held to a higher legal standard? What

---

[80]This quote is attributed to Michael Beckerman with the Internet Association, a lobbying organization that represents some of the largest Internet companies. See https://www.npr.org/sections/alltechconsidered/2018/03/21/591622450/section-230-a-key-legal-shield-for-facebook-google-is-about-to-change.

[81]Many platforms rely on a combination of screening and peer-to-peer feedback mechanisms. For example, Uber runs various background checks on its drivers, eliminates drivers based on negative reviews, and shares reviews with users. See Einav et al. (2016). See Tadelis (2016) for a thoughtful discussion of the limits and biases in peer-to-peer feedback mechanisms.

if the allocation of control rights is endogenous?[82]

Although internet platforms provided the motivation for this paper, our insights apply more broadly. Our analysis provides a strong economic rationale for holding traditional newspapers liable for harmful advertising content[83] and for holding bricks-and-mortar retailers liable for the harm caused by defective products.[84] Although our model is broadly applicable, we believe that the insights are particularly salient for online platforms including Facebook, Google, and Amazon. First, the harmful participants on these platforms are frequently small and judgment proof with insufficient incentives to curtail their harmful activities. Second, the big tech giants have the data and technology to detect and block participants that are more likely to harm others. It is therefore ironic that the big internet platforms enjoy legal protections that are unavailable to traditional business models.

---

[82]Hagiu and Wright (2015 and 2018) examine the allocation of control rights between intermediaries and firms over transferable decisions such as marketing activities. Platform liability is not addressed.

[83]See *Braun v. Soldier of Fortune Magazine*, Inc., 968 F.2d 1110 (1992). The court opined: "[T]he first Amendment permits a state to impose upon a publisher liability for compensatory damages for negligently publishing a commercial advertisement where the ad on its face, and without the need for investigation, makes it apparent that there is substantial danger of harm to the public."

[84]See In re Mattel, Inc., 588 F. Supp. 2d 1111 (C.D. Cal. 2008). Some toy buyers brought suit against manufacturers and retailers (including Wal-Mart) for unsafe toys. See also Restatement (Third) of Torts (1998). "One engaged in the business of selling or otherwise distributing products who sells or distributes a defective product is subject to liability for harm to persons or property caused by the defect."

References

[1] Armstrong, Mark, "Competition in Two-sided Markets," *RAND Journal of Economics*, Vol. 37 (2006), pp. 668-691.

[2] Armstrong, Mark and Julian Wright, "Two-Sided Markets, Competitive Bottlenecks, and Exclusive Contracts," *Economic Theory*, Vol. 32 (2007), pp. 353-380.

[3] Arlen, Jennifer, and W. Bentley MacLeod, "Malpractice Liability for Physicians and Managed Care Organizations," *New York University Law Review*, Vol. 78 (2003), pp. 1929-2006.

[4] Bebchuk, Lucian and Jesse Fried, "The Uneasy Case for the Priority of Secured Claims in Bankruptcy," *Yale Law Journal*, Vol. 105 (1996), pp. 857-934.

[5] Belleflamme, Paul and Martin Peitz, "Managing Competition on a Two-Sided Platform," *Journal of Economics & Management Strategy*, Vol. 28 (2019), pp. 5-22.

[6] Ben-Shahar, Omri, "Should Products Liability Be Based on Hindsight?" *Journal of Law, Economics, and Organization*, Vol.14 (1998), pp. 325-357.

[7] Boyer, Marcel, and Jean-Jacques Laffont, "Environmental Risk and Bank Liability," *European Economic Review*, Vol. 41 (1997), pp. 1427-1459.

[8] Buiten, Miriam C., Alexandre de Streel, and Martin Peitz, "Rethinking Liability Rules for Online Hosting Platforms," *International Journal of Law and Information Technology*, Vol. 28 (2020), pp. 139-166.

[9] Caillaud, Bernard, and Bruno Jullien, "Chicken & Egg: Competition among Intermediation Service Providers," *RAND Journal of Economics*, Vol. 34 (2003), pp. 309-328.

[10] Che, Yeon-Koo, and Kathryn E. Spier, "Strategic Judgment Proofing," *RAND Journal of Economics*, Vol. 39 (2008), pp. 926-948.

[11] Chen, Yongmin and Xinyu Hua, "Ex ante Investment, Ex post Remedies, and Product Liability," *International Economic Review*, Vol 53 (2012), pp. 845-866.

[12] Chen, Yongmin and Xinyu Hua, "Competition, Product Safety, and Product Liability," *Journal of Law, Economics, & Organization*, Vol. 33 (2017), pp. 237-267.

[13] Choi, Albert, and Kathryn E. Spier, "Should Consumers Be Permitted to Waive Products Liability? Product Safety, Private Contracts, and Adverse Selection," *Journal of Law, Economics, & Organization*, Vol. 30 (2014), pp. 734-766.

[14] Choi, Jay Pil, and Arijit Mukherjee, "Optimal Certification Policy, Entry, and Investment in the Presence of Public Signals," *RAND Journal of Economics*, Vol. 51 (2020), pp. 989-1013.

[15] Culotta, Aron, Ginger Zhe Jin, Yidan Sun, and Liad Wagman, "Safety Reviews on Airbnb: An Information Tale," (2022), working paper.

[16] Daughety, Andrew F., and Jennifer F. Reinganum, "Product Safety: Liability, R&D, and Signaling," *American Economic Review,* Vol. 85 (1995), pp. 1187-1206.

[17] Daughety, Andrew F., and Jennifer F. Reinganum, "Everybody Out of the Pool: Products Liability, Punitive Damages, and Competition," *Journal of Law, Economics and Organization*, Vol. 13 (1997), pp. 410-432.

[18] Daughety, Andrew F. and Jennifer F. Reinganum, "Market, Torts, and Social Inefficiency," *RAND Journal of Economics*, Vol. 37 (2006), pp. 300-323.

[19] Daughety, Andrew F., and Jennifer F. Reinganum, "Communicating Quality: a Unified Model of Disclosure and Signaling," *RAND Journal of Economics*, Vol. 39 (2008b), pp. 973-989.

[20] Daughety, Andrew F., and Jennifer F. Reinganum, "Imperfect Competition and Quality Signaling," *RAND Journal of Economics,* Vol. 39 (2008a), pp. 163-183.

[21] Daughety, Andrew F. and Jennifer F. Reinganum, "Cumulative Harm and Resilient Liability Rules for Product Markets," *Journal of Law, Economics, and Organization,* Vol. 30 (2014), pp. 371-400.

[22] Dari Mattiacci, Giuseppe, and Francesco Parisi, "The Cost of Delegated Control: Vicarious Liability, Secondary Liability and Mandatory Insurance," *International Review of Law and Economics*, Vol. 23 (2003), pp. 453-475.

[23] Dukes, Anthony, and Esther Gal-Or, "Negotiations and Exclusivity Contracts for Advertising," *Management Science*, Vol. 22 (2003), pp. 222-245.

[24] Einav, Liran, Chiara Farronato, and Jonathan Levin, "Peer-to-Peer Markets," *Annual Review of Economics*, Vol. 8 (2016), pp. 615-635.

[25] Epple, Dennis, and Artur Raviv, "Product Safety: Liability Rules, Market Structure, and Imperfect Information," *American Economic Review*, Vol. 68 (1978), pp. 80-95.

[26] Fu, Qiang, Jie Gong, and Ivan Png, "Law, Social Responsibility, and Outsourcing," *International Journal of Industrial Organization*, Vol. 57 (2018), pp. 114-146.

[27] Galeotti, Andrea, and Jose Luis Moraga-Gonzalez, "Platform Intermediation in a Market for Differentiated Products," *European Economic Review*, Vol. 53 (2009), pp. 417-428.

[28] Gomes, Renato, "Optimal Auction Design in Two-Sided Markets," *RAND Journal of Economics*, Vol. 45 (2014), pp. 248-272.

[29] Hagiu, Andrei, "Pricing and Commitment by Two-Sided Platforms," *RAND Journal of Economics*, Vol. 37 (2006), pp. 720-737.

[30] Hagiu, Andrei, "Quantity vs. Quality and Exclusion by Two-Sided Platforms," (2009), working paper.

[31] Hagiu, Andrei, and Julian Wright, "Marketplace or Reseller?" *Management Science*, Vol. 61 (2015), pp. 184-203.

[32] Hagiu, Andrei, and Julian Wright, "Controlling vs. Enabling," *Management Science*, Vol. 65 (2018), pp. 577-595.

[33] Hamdani, Assaf, "Who is Liable for Cyberwrongs?" *Cornell Law Review*, Vol. 87 (2002), pp. 901-957.

[34] Hamdani, Assaf, "Gatekeeper Liability," *Southern California Law Review*, Vol. 77 (2003), pp. 53-122.

[35] Hay, Bruce, and Kathryn E. Spier, "Manufacturer Liability for Harms Caused by Consumers to Others," *American Economic Review*, Vol.95 (2005), pp. 1700-1711.

[36] Hua, Xinyu and Kathryn E. Spier, "Product Safety, Contracts, and Liability," *RAND Journal of Economics*, Vol. 51 (2020), pp. 233-259.

[37] Julien, Bruno, and Alessandro Pavan, "Information Management and Pricing in Platform Markets," *Review of Economic Studies*, Vol. 86 (2019), pp. 1666-1703.

[38] Kraakman, Reinier H., "Gatekeepers: The Anatomy of a Third-Party Enforcement Strategy," *Journal of Law, Economics, & Organization*, Vol. 2 (1986), pp. 53-104.

[39] Karle, Heiko, Martin Peitz, and Markus Reisinger, "Segmentation versus Agglomeration: Competition between Platforms with Competitive Sellers," *Journal of Political Economy*, Vol. 128 (2020), pp. 2329-2374.

[40] Lefouili, Yassine, and Leonardo Madio, "The Economics of Platform Liability," *European Journal of Law and Economics*, Vol. 53 (2022), pp. 319-351.

[41] Nocke, Volker, Martin Peitz, and Konrad Stahl, "Platform Ownership," *Journal of the European Economic Association*, Vol. 5 (2007), pp. 1130-1160.

[42] Pitchford, Rohan, "How Liable Should a Lender Be? The Case of Judgment-Proof Firms and Environment Risk," *American Economic Review*, Vol. 85 (1995), pp. 1171-1186.

[43] Polinsky, A. Mitchell, "Strict Liability vs. Negligence in a Market Setting," *American Economic Review*, Vol. 70 (1980), pp. 363-367.

[44] Polinsky, A. Mitchell and William P. Rogerson, "Products Liability, Consumer Misperceptions, and Market Power." Bell Journal of Economics, Vol. 14 (1983), pp. 581-589.

[45] Rochet, Jean-Charles, and Jean Tirole, "Platform Competition in Two-Sided Markets," *Journal of the European Economic Association*, Vol. 1 (2003), pp. 990-1029.

[46] Rochet, Jean-Charles, and Jean Tirole, "Two-Sided Markets: A Progress Report," *RAND Journal of Economics*, Vol. 37 (2006), pp. 645-667.

[47] Simon, Marilyn J, "Imperfect Information, Costly Litigation, and Product Quality," *Bell Journal of Economics*, Vol. 12 (1981), pp. 171-184.

[48] Shavell, Steven, "The Judgment Proof Problem," *International Review of Law and Economics*, Vol. 6 (1986), pp. 45-58.

[49] Shavell, Steven, "Liability for Accidents," Chapter 2, Handbook of Law and Economics (2007), Elsevier.

[50] Spence, A. Michael, "Consumer Misperceptions, Product Failure, and Producer Liability," *Review of Economic Studies*, Vol. 44 (1977), pp. 561-572.

[51] Spence, A. Michael, "Monopoly, Quality, and Reputation," *Bell Journal of Economics*, Vol. 6 (1975), pp. 417-429.

[52] Tadelis, Steven, "Reputation and Feedback Systems in Online Platform Markets," *Annual Review of Economics*, Vol. 8 (2016), pp. 321-340.

[53] Tan, Guofu and Junjie Zhou, "The Effects of Competition and Entry in Multi-sided Markets," *Review of Economic Studies*, Vol. 88 (2021), pp. 1002-1030.

[54] Teh, Tat-How, "Platform Governance," *American Economic Journal: Microeconomics*, (Forthcoming)

[55] Weyl, Glen, "A Price Theory of Multi-Sided Platforms," *American Economic Review*, Vol. 100 (2010), pp. 1642-1672.

[56] White, Alexander, and Glen Weyl, "Imperfect Platform Competition: A General Framework," (2010), working paper.

[57] Wickelgren, Abraham L., "The Inefficiency of Contractually Based Liability with Rational Consumers," *Journal of Law, Economics, & Organization*, Vol. 22 (2006), pp. 168-183.

Appendix: Proofs

Proof of Lemma 1. We first show that the platform does not find it profitable to deter the L-types and retain the H-types. If the platform deters the L-types by setting a high price $p_H = \alpha_H - \theta_H w_s$, its profit is

$$\Pi_H(e) = \lambda(1-e)(\alpha_H - \theta_H w) - c(e),$$

where $w = w_s + w_p$. As defined in the text, $\Pi(e)$ is the platform's profit when it charges $p_L = \alpha_L - \theta_L w_s$. Consider two scenarios.

First, suppose $w > \frac{\alpha_H}{\theta_H}$. Then $\Pi_H(e) < 0$ for any e. Assumption A2 implies $\Pi(0) > 0$, that is, the profit from attracting both types is larger than the profit from deterring the L-types.

Second, suppose $w \leq \frac{\alpha_H}{\theta_H}$. Since $\alpha_H - \theta_H w \geq 0$, the platform would not take any auditing effort and the optimal profit is $\Pi_H(0) = \lambda(\alpha_H - \theta_H w)$. We have

$$
\begin{aligned}
\Pi(0) - \Pi_H(0) &= \lambda(\alpha_L - \theta_L w_S - \theta_H w_p) + (1-\lambda)(\alpha_L - \theta_L w_S - \theta_L w_p) \\
&\quad -\lambda(\alpha_H - \theta_H w) \\
&= \alpha_L - \lambda \alpha_H - (1-\lambda)\theta_L w + \lambda(\theta_H - \theta_L)w_s \\
&\geq \alpha_L - \lambda \alpha_H - (1-\lambda)\theta_L \frac{\alpha_H}{\theta_H} \\
&= \alpha_L - (\lambda \theta_H + (1-\lambda)\theta_L)\frac{\alpha_H}{\theta_H} \\
&> 0,
\end{aligned}
$$

where the first inequality holds given $w \leq \frac{\alpha_H}{\theta_H}$ and the second inequality follows from Assumption A2. Therefore, the platform would not deter the L-types.

Now we prove the remaining results in the lemma. Using the definition of $r_H(w_s)$ in the lemma, (8) implies $e^* > 0$ if and only if $(\alpha_H - \theta_H w) - (\theta_H - \theta_L)(\widehat{w} - w_s) < 0$. This gives the condition for cases 1 and 2. Totally differentiating (10), and using the fact the social welfare function is concave, gives $de^*/dw_s = -\lambda\theta_L/S''(e) > 0$ and $de^*/dw_p = -\lambda\theta_H/S''(e) > 0$. When $e^* > 0$ (an interior solution), increasing the level of liability for either the firm or the platform increases the platform's auditing effort. Equation (10) implies $e^* > e^{**}$ if and only if $\lambda r_H(w_s) - \lambda\theta_H(d - w) > 0$. This gives the condition for subcases 2(a), 2(b) and 2(c).

Proof of Proposition 1. Note that $\widehat{w} < d < \frac{\alpha_L}{\theta_L}$ by Assumption A1. Suppose $w_p = 0$ and $w_s < \widehat{w}$. From Lemma 1, a necessary and sufficient condition for $e^* = 0$ is (8) or

$$\alpha_H - \theta_H w_s > (\theta_H - \theta_L)(\widehat{w} - w_s).$$

Substituting for $\widehat{w}$ from (5),

$$\alpha_H - \theta_H w_s > (\alpha_H - \alpha_L) - (\theta_H - \theta_L)w_s,$$

which is equivalent to $w_s < \frac{\alpha_L}{\theta_L}$. Since $w_s < \widehat{w} < \frac{\alpha_L}{\theta_L}$ we have $e^* = 0$.

Suppose $w_s \geq \widehat{w}$. There are two possible scenarios. First, if $\theta_L/\theta_H < \alpha_L/\alpha_H$, then setting $w_p = 0$ in Lemma 2 and rearranging terms gives a threshold value $\widetilde{w}(\lambda) = \frac{\alpha_H - \alpha_L + \lambda\alpha_L}{\theta_H - \theta_L + \lambda\theta_L} \in \left(\widehat{w}, \frac{\alpha_H}{\theta_H}\right)$. Moreover, $\frac{d\widetilde{w}(\lambda)}{d\lambda} > 0$ given $\theta_L/\theta_H < \alpha_L/\alpha_H$. When $w_s < \widetilde{w}(\lambda)$, the platform sets $p^* = \alpha_H - \theta_H w_s$, and attracts the H-types; when $w_s > \widetilde{w}(\lambda)$, the platform sets $p^* = \alpha_L - \theta_L w_s$ and deters the H-types. Second, if $\theta_L/\theta_H \geq \alpha_L/\alpha_H$, then $\frac{\alpha_H - \alpha_L + \lambda\alpha_L}{\theta_H - \theta_L + \lambda\theta_L} \leq \widehat{w} \leq w_s$. In this scenario, Lemma 2 implies that the platform always sets $p^* = \alpha_L - \theta_L w_s$ and deters the H-types. The two scenarios can be combined by defining $\widetilde{w}(\lambda) = \max\{\frac{\alpha_H - \alpha_L + \lambda\alpha_L}{\theta_H - \theta_L + \lambda\theta_L}, \widehat{w}\}$.

Proof of Proposition 2. Suppose $w_s < \widehat{w}$, so the L-type is marginal. The platform cannot deter the H-types directly through the price, but can remove them through auditing. From equation (10) we have $e^* = e^{**}$ if and only if $w_p = w_p^* = d - w_s - \left(1 - \frac{\theta_L}{\theta_H}\right)(\widehat{w} - w_s)$. Note that $w_p^* \in (0, d - w_s)$.

Suppose $w_s \in [\widehat{w}, \widetilde{w})$. From Proposition 1, if $w_p = 0$, the platform sets $p = \alpha_H - \theta_H w_s$, and attracts the H-type firms. This is socially inefficient. Lemma 2 implies that the platform would deter the H-type if $\lambda(\alpha_H - \theta_H w) \leq (1 - \lambda)r_L(w_s)$. $\lambda(\alpha_H - \theta_H w)$ decreases in w_p and the firms' rent $(1 - \lambda)r_L(w_s)$ is independent of w_p. Setting $\lambda(\alpha_H - \theta_H w) = (1 - \lambda)r_L(w_s)$ gives the lower bound $\underline{w}_p > 0$:

$$\underline{w}_p = \frac{\alpha_H}{\theta_H} - w_s - \frac{1-\lambda}{\lambda}\left(1 - \frac{\theta_L}{\theta_H}\right)(w_s - \widehat{w}).$$

For any $w_p^* \geq \underline{w}_p$, the platform deters the H-types and the first-best outcome is obtained.

Suppose $w_s \geq \widetilde{w}$. Proposition 1 implies that even if $w_p = 0$ the platform sets $p^* = \alpha_L - \theta_L w_s$, deters H-type firms, and the first-best outcome is obtained. Platform liability is unnecessary. Any $w_p^* \in [0, d - w_s]$ achieves the first-best outcome.

Proof of Lemma 3. Since $w_s < \widehat{w}$, it is not possible for the platform to deter the H-types without deterring the L-types, too. If the L-type is willing to participate, then the H-type strictly prefers to participate.

To begin, we construct values $\{e^*, p^*, t^*\}$ that maximize the platform's profits subject to the platform's incentive compatibility constraint and the participation constraints of the consumers and the L-type firms (as the L-type firm is marginal). Then, we will verify that these values are an equilibrium of the game.

$$\max_{\{e,p,t\}} \quad \Phi(e, p) = (1 - e)\lambda(p - \theta_H w_p) + (1 - \lambda)(p - \theta_L w_p) - c(e) \tag{25}$$

subject to

$$e = \arg\max_{e' \geq 0} \Phi(e', p) \tag{26}$$

$$\alpha_0 - t - E(\theta|e)(d - w_s - w_p) \geq 0 \tag{27}$$

$$t - (\theta_L w_s + c_L) - p \geq 0. \tag{28}$$

(26) is the platform's incentive compatibility constraint, (27) is the consumer's participation constraint, and (28) is the L-type firm's participation constraint.[85]

The L-type's participation constraint (28) must bind. To see this, consider two cases. First, suppose that neither (27) nor (28) binds. Then the platform would increase the price p which would increase the platform's profits in (25) and maintain the consumer's participation constraint (27). Second, suppose that (27) binds while (28) does not. Again, the platform would increase the price p marginally. The direct effect of increasing p is that the platform's profits in (25) increase. Since $\partial^2 \Phi(e,p)/\partial e \partial p = -\lambda < 0$, increasing p also (weakly) decreases the platform's effort e in (26), which in turn raises $E(\theta|e)$ and, since (27) binds, reduces t. However, since t is not in (25), the platform's profits still increase.

Since the L-type's constraint (28) binds, $p = t - (\theta_L w_s + c_L)$ and we can rewrite the optimand (25) as a function of e and t:

$$(1-e)\lambda(t - (\theta_L w_s + c_L) - \theta_H w_p)$$
$$+ (1-\lambda)(t - (\theta_L w_s + c_L) - \theta_L w_p) - c(e). \quad (29)$$

Next, we show that the consumer's participation constraint (27) binds. Suppose not. Then, the platform would increase t and its profits would rise. Since both participation constraints (27) and (28) bind, we have

$$p = \alpha_0 - E(\theta|e)(d - w_s - w_p) - (\theta_L w_s + c_L). \quad (30)$$

Since $\alpha_L = \alpha_0 - c_L$ and $w = w_s + w_p$ the solution to the platform's optimization problem is:

$$e^* = \arg\max_{e \geq 0} \Phi(e, p^*) \quad (31)$$

$$t^* = \alpha_0 - E(\theta|e^*)(d - w) \quad (32)$$

$$p^* = \alpha_L - \theta_L w_s - E(\theta|e^*)(d - w). \quad (33)$$

We now verify that the values $\{e^*, p^*, t^*\}$ defined in (31), (32), and (33) are an equilibrium of the game. Suppose that the platform charges p^* in (33), and that the firms and consumers believe that the probability of harm is $\theta^* = E(\theta|e^*)$ where e^* defined in (31). The consumers are (just) willing to pay t^* in (32) and the L-type firms are (just) willing to pay p^* in (33). If the consumers and the firms all participate, the platform exerts effort e^* in (31). Therefore the equilibrium beliefs $\theta^* = E(\theta|e^*)$ are consistent.

Next, we verify that Assumption A2 guarantees that the platform's profits are positive. To do this, we will show that the platform's profits are positive even if consumers and the firms believed that the platform is not auditing at all, so $\widehat{\theta} = E(\theta|0) = \theta^0$.[86] In this

[85]The H-type's participation constraint is satisfied if (28) holds, and is therefore not included in the program.

[86]The platform is better off if the consumers believe that the product is safer (i.e., that $\widehat{\theta}$ is smaller). If consumers perceive the product to be safer, they will pay a higher price t for the product which means that the platform can charge the firms a higher price p.

scenario, the most that consumers would be willing to pay is $t = \alpha_0 - \theta^0(d - w)$ from (27). The most that the L-type firms would be willing to pay is $p = \alpha_L - \theta_L w_s - \theta^0(d - w)$ from (28). The platform's profits can be rewritten as

$$\Pi(0) = \alpha_L - \theta^0 d + \lambda(\theta_H - \theta_L)w_s.$$

Therefore, $\Pi(0) > 0$ for any $w_s \geq 0$ if Assumption $A2$ holds.[87]

We now show that the algebraic condition in case 1 is necessary and sufficient for a corner solution, $e^* = 0$. We first show the condition is necessary. If $e^* = 0$ then $E(\theta|0) = \theta^0$. Since the consumer's participation constraint (27) binds we have $t^* = \alpha_0 - \theta^0(d - w)$; since the L-type firm's participation constraint (28) binds we have $p^* = \alpha_L - \theta_L w_s - \theta^0(d - w)$. Finally, for $e^* = 0$ to satisfy the platform's IC constraint (26) we need $\partial\Phi(e, p)/\partial e \leq 0$ or equivalently $p^* - \theta_H w_p \geq 0$. Substituting p^*, this condition becomes

$$\alpha_L - \theta_L w_s - \theta^0(d - w) - \theta_H w_p \geq 0. \tag{34}$$

Adding and subtracting terms this becomes

$$(\alpha_H - \theta_H d) - (\alpha_H - \alpha_L) - \theta_L w_s - \theta_H w_p + \theta_H w$$
$$+ (\theta_H - \theta^0)(d - w) \geq 0, \tag{35}$$

and rearranging this expression gives

$$(\alpha_H - \theta_H d) + (\theta_H - \theta^0)(d - w) \geq (\alpha_H - \alpha_L) - (\theta_H - \theta_L)w_s. \tag{36}$$

The right-hand side is $r_H(w_s)$. This confirms that the condition in case 1 is necessary.

Next, we show that the condition in case 1 is sufficient. Suppose the condition holds and $e^* > 0$. Since $E(\theta|e^*) < \theta^0$, $t^* > \alpha_0 - \theta^0(d - w)$ and $p^* > \alpha_L - \theta_L w_s - \theta^0(d - w)$. Assumption A2 implies $p^* - \theta_H w_p > 0$, so the platform does not audit, $e^* = 0$.

Now consider case 2 of Lemma 3. The condition implies $p^* - \theta_H w_p < 0$ so the platform is losing money from each H-type transaction. The equilibrium effort $e^* > 0$ and consumers' equilibrium beliefs $\theta^* = E(\theta|e^*)$ satisfy equation (19). the platform charges $p^* = \alpha_L - \theta_L w_s - \theta^*(d - w)$ and consumers believe that the platform will exert effort e^* and are willing to pay $t^* = \alpha_0 - \theta^*(d - w)$. Condition (19) implies that $e^{**} < e^*$ if and only if $(\theta_H - \theta^{**})(d - w) < (\theta_H - \theta_L)(\widehat{w} - w_s)$. Totally differentiating condition (19) and using the fact that the welfare function is concave, we have $de^*/dw_s < 0$ and $de^*/dw_p > 0$.

Proof of Lemma 4. Since $w_s > \widehat{w}$ the H-type firms are marginal. The platform can deter the H-types by charging a price that only the L-types would accept. The users'

---

[87]If $e = 1$ then $\widehat{\theta} = \theta_L$. One can verify that $\Pi(1) > 0$ if and only if $\alpha_L - \theta_L d > \frac{c(1)}{1-\lambda}$. This condition is independent of w_s and w_p. It may hold even if $A2$ is not satisfied (that is, $\alpha_L - \theta_L d \leq \lambda(\theta_H - \theta_L)d$). When this condition holds, even if $A2$ is not satisfied, the platform may still be active. That is, $A2$ is a sufficient but not necessary condition for the platform to be active.

posterior beliefs are $\widehat{\theta} = \theta_L$, and so the firms charge the consumers $t^* = \alpha_0 - \theta_L(d - w)$. The platform's price extracts the L-type firm's surplus, $p^* = t^* - (\theta_L w_s + c_L)$. Therefore

$$p^* = \alpha_L - \theta_L w_s - \theta_L(d - w) = \alpha_L - \theta_L(d - w_p) \tag{37}$$

and the platform's profits are

$$(1 - \lambda)(p^* - \theta_L w_p) = (1 - \lambda)(\alpha_L - \theta_L d). \tag{38}$$

In other words, the platform extracts the full social surplus from the L-types.

If the platform chooses to attract the H-type firms, then the platform will not audit them. (The firm's profits would be higher if they charged a high price that deterred the H-types.) If the platform charges a price p that attracts the marginal H types, the users' posterior beliefs are the same as their priors, $\widehat{\theta} = \theta^0 = \lambda\theta_H + (1 - \lambda)\theta_L$, and the firms charge the consumers $t^* = \alpha_0 - \theta^0(d - w)$. The platform's price extracts the marginal H-type firm's surplus, that is, $p^* = t^* - (\theta_H w_s + c_H)$ or

$$p^* = \alpha_H - \theta_H w_s - \theta^0(d - w). \tag{39}$$

The platform's profits are

$$\begin{aligned}
p^* - \theta^0 w_p &= (1 - \lambda)(\alpha_L - \theta_L d) + \lambda(\alpha_H - \theta_H d) + (1 - \lambda)[\alpha_H - \alpha_L - (\theta_H - \theta_L)w_s] \\
&= (1 - \lambda)(\alpha_L - \theta_L d) + \lambda(\alpha_H - \theta_H d) + (1 - \lambda)(\theta_H - \theta_L)(\widehat{w} - w_s) \\
&< (1 - \lambda)(\alpha_L - \theta_L d)
\end{aligned}$$

where the inequality follows from Assumption A1 and $w_s > \widehat{w}$. Therefore, if $w_s > \widehat{w}$, the platform charges $p^* = \alpha_L - \theta_L(d - w_p)$ and deters the H-types.

Proof of Proposition 3. Suppose $w_s \geq \widehat{w}$. The result follows immediately from Lemma 4. Suppose $w_s < \widehat{w}$. Setting $w_p = 0$ in Lemma 3, the platform attracts the H-types and does not invest in auditing ($e^* = 0$) if

$$\begin{aligned}
0 &\leq (\alpha_H - \theta_H d) + (\theta_H - \theta^0)(d - w_s) - r_H(w_s) \\
&= (\alpha_H - \theta_H d) + (\theta_H - \theta^0)(d - w_s) - (\theta_H - \theta_L)(\widehat{w} - w_s) \\
&= (\alpha_H - \theta_H d) + (\theta_H - \theta^0)(d - w_s) - (\alpha_H - \alpha_L) + (\theta_H - \theta_L)w_s \\
&= \alpha_L - \theta_L w_s - \theta^0(d - w_s)
\end{aligned}$$

where $\theta^0 = E(\theta|0)$. Replacing $\theta^0 = E(\theta|0) = \lambda\theta_H + (1 - \lambda)\theta_L$, we can rewrite this as

$$\begin{aligned}
\alpha_L - \theta_L w_s &- (\lambda\theta_H + (1 - \lambda)\theta_L)(d - w_s) \\
&= \alpha_L - (\lambda\theta_H + (1 - \lambda)\theta_L)d + \lambda(\theta_H - \theta_L)w_s \geq 0. \tag{40}
\end{aligned}$$

Assumption A2 guarantees that this is true for all $w_s \geq 0$.

Proof of Proposition 4. Suppose $w_s < \widehat{w}$, so the L-type is marginal. From equation (19) we have $e^* = e^{**}$ if and only if

$$(\theta_H - \theta_L)(\widehat{w} - w_s) - (\theta_H - \theta^{**})(d - w) = 0. \tag{41}$$

Substituting that $w = w_p + w_s$ and isolating w_p on the left-hand side establishes the result. Suppose $w_s \geq \widehat{w}$. The results follow from Lemma 4.

Proof of Lemma 5. We first show that the platforms receive zero profits in equilibrium. If one platform received positive profits while the other got no profit, the second platform would deviate and imitate the first one's strategies.

Suppose that both platforms got positive profits and Platform 2 attracted weakly more H-type firms than Platform 1. Since $w_s < \widehat{w}$, the L-type firms get lower rents than the H-types. Thus, the L-type firms must be indifferent between joining the two platforms. But then Platform 2 would reduce its price marginally but keep its auditing effort, which would steal all the L-types (and possibly the H-types) from Platform 1 and therefore weakly reduce the proportion of H-types on Platform 2. Since Platform 2 got positive profits when having more H-types, attracting more firms with a larger proportion of L-types would strictly raise its profits. Therefore, both platforms should receive zero profits in equilibrium.

Next, we show that, as long as both platforms are active, they charge the same price and take the same auditing effort in equilibrium. Note that, if only one platform got the L-type firms, then the H-types would join this platform too because they get higher rents than the L-types. Therefore, as long as both platforms are active, they should get some L-types. That is, the L-types are indifferent between joining the two platforms. Since the L-types would never be removed, the platforms' prices must be the same. Furthermore, if the two platforms chose different auditing levels, the one with less auditing would attract all the H-types. However, this platform could reduce its price marginally and steal the L-types from the other platform, which would reduce the proportion of H-types and raise its profits.

To summarize, the above analysis suggests that in any equilibrium the platforms get zero profits, charge the same price, and take the same auditing effort.

Now we show that e^c increases in w_p. Define $Z = p^c - \theta_H w_p$. Condition (23) can be re-written as

$$\Pi = (1 - e^c)\lambda Z + (1 - \lambda)[Z + (\theta_H - \theta_L)w_p] - c(e^c) = 0.$$

Differentiating with respect to w_p, and recognizing that e^c and Z are functions of w_p, this implies

$$\frac{d\Pi}{de^c}\frac{de^c}{dw_p} + \frac{d\Pi}{dZ}\frac{dZ}{dw_p} + (1 - \lambda)(\theta_H - \theta_L) = 0.$$

Since $\frac{d\Pi}{de^c} = 0$ (the first-order condition), $\frac{d\Pi}{dZ} > 0$, and $(1 - \lambda)(\theta_H - \theta_L) > 0$, we have $\frac{dZ}{dw_p} < 0$. Condition (24) may be written as $-\lambda Z - c'(e^c) = 0$. Differentiating this with

respect to w_p gives $-\lambda \frac{dZ}{dw_p} - c''(e^c)\frac{de^c}{dw_p} = 0$. Finally, $\frac{dZ}{dw_p} < 0$ and $c''(e^c) > 0$ imply that $\frac{de^c}{dw_p} > 0$.

Finally, note that $p^c < p^* = \alpha_L - \theta_L w_s$. To see this, suppose that $e = 0$ and $p \geq \alpha_L - \theta_L w_s$. Then

$$
\begin{aligned}
\Pi(0) &\geq \alpha_L - \theta_L w_s - [\lambda \theta_H + (1-\lambda)\theta_L]w_p \\
&\geq \alpha_L - \theta_L w_s - [\lambda \theta_H + (1-\lambda)\theta_L](d - w_s) \\
&> 0
\end{aligned}
$$

where the second inequality follows from $w_p \leq d - w_s$ and the last inequality holds given Assumption A2. Thus, condition (23) implies that $p^c < p^* = \alpha_L - \theta_L w_s$. And condition (24) then implies $e^c > e^*$ as long as $w_p > 0$.

Lemma 1 implies that, when there is a monopoly platform, $0 < e^* < e^{**}$ if $w_p < w_p^*$ and $0 < e^{**} < e^*$ if $w_p > w_p^*$. Note that $w_p^* \in (0, d - w_s)$. Since e^c increases in w_p and $e^c > e^*$ if $w_p > 0$, there exists a unique value $\overline{w}_p \in (0, w_p^*)$ such that $e^c = e^{**}$ if and only if $w_p = \overline{w}_p$.

If $0 < w_p \leq \overline{w}_p$, then $e^c \leq e^{**}$, while $e^* < e^c$ as shown by Lemma 1. Therefore, under competition, the auditing intensity is closer to the socially efficient level, which raises welfare.

If $w_p \geq w_p^*$ then $e^{**} \leq e^* < e^c$ (given Lemma 1). Therefore, competition exacerbates the distortion in auditing and reduces welfare.

Proof of Lemma 6. We show that there is no equilibrium where the platforms get positive profits. Consider four scenarios.

First, if one platform received positive profits while the other platform got no profit, the second platform would deviate and imitate the first one's strategies.

Second, suppose that both platforms got positive profits, and they both got some L-type firms while Platform 2 attracted weakly more H-type firms. The L-types must be indifferent between joining the two platforms. But then Platform 2 would reduce its price marginally and keep its auditing effort, which would steal firms from Platform 1 and weakly reduce the proportion of H-types on Platform 2. Since Platform 2 got positive profits when having more H-types, attracting more firms with a larger proportion of L-types would strictly raise its profits.

Third, suppose that both platforms got positive profits and only Platform 1 attracted the L-type firms. Note that Platform 2 would not take any auditing effort in this case. Also, since $w_s \geq \widehat{w}$, the H-type firms get (weakly) lower rents than the L-types. We can show that in equilibrium the L-type firms must be indifferent between joining the two platforms. If Platform 1 got non-negative profits from the H-types, it would not take any auditing effort and the platforms' equilibrium prices must be the same, which implies that the L-types would be indifferent between joining the two platforms. If Platform 1 got negative profits from H-type firms, its price should deter the H-types and make the L-types indifferent between joining the two platforms (as otherwise Platform 1 could raise the price). However, since the L-types are indifferent, Platform 2 could reduce its price

marginally, which would weakly reduce the proportion of H-types on Platform 2 and raise its profit.

Finally, suppose that both platforms deterred the H-types and still got positive profits. Platform j would receive positive profits if and only if $p_j > \theta_L w_p$; and the H-type firms would be deterred if and only if $p_j > \alpha_H - \theta_H w_s$, for $j = 1, 2$. So, the price set by platform j would satisfy $p_j > \max\{\alpha_H - \theta_H w_s, \theta_L w_p\}$. In this case, each platform could deviate and reduce the price marginally, which would steal the L-type firms from the competitor but not attract the H-type firms.

To conclude, in any equilibrium, the platforms get zero profits. Without loss of generality, we focus on the symmetric equilibrium.

In any equilibrium where all the H-type firms are deterred (if it exists), the zero-profit condition implies $p_1 = p_2 = \theta_L w_p$. And the H-types would not join the platform if $\alpha_H - \theta_L w_p - \theta_H w_s < 0$, or equivalently,

$$w_p > \frac{\alpha_H - \theta_H w_s}{\theta_L}. \tag{42}$$

Given $w_p > \frac{\alpha_H - \theta_H w_s}{\theta_L}$ and $p_1 = p_2 = \theta_L w_p$, each platform has no incentive to reduce or raise its price; the H-type firms do not join any platform. Therefore, condition (42) is sufficient and necessary for the existence of the equilibrium where all the H-type firms are deterred. Moreover, under this condition, there is no equilibrium where the H-type firms would participate, as the platforms would never reduce their prices below $\theta_L w_p$, which is higher than $\alpha_H - \theta_H w_s$.

Now, suppose $w_p \le \frac{\alpha_H - \theta_H w_s}{\theta_L}$. Recall that p^c and e^c are defined by (23) and (24).

If $\alpha_H - \theta_H w_s \ge p^c$, then there exists an equilibrium where the platforms attract the H-types and get zero profits. In such an equilibrium, the equilibrium price is p^c and the equilibrium auditing effort (if it is an interior solution) is e^c. Similar to the analysis in Lemma 5, if $w_p = 0$ then the platforms do not audit, $e^c = e^* = 0$; if $w_p > 0$ then $e^c > 0$, and $de^c/dw_p > 0$. Since $p^c \le \alpha_H - \theta_H w_s$, the H-type firms would join the platforms.

If $\alpha_H - \theta_H w_s < p^c$, then the H-types would not participate if the platforms charge p^c. However, we will show that there exists an equilibrium where the platforms charge $p_1 = p_2 = \alpha_H - \theta_H w_s$ and attracts a mass $\widehat{\lambda} \in [0, \lambda)$ of the H-types, while the mass of participating L-types remains at $1 - \lambda$. To see this, note that the H-types are indifferent between participating and not participating. Given the reduced participation of the H-types, the posterior fraction of H-types is $\lambda' = \widehat{\lambda}/(\widehat{\lambda} + (1 - \lambda))$ and the (posterior) fraction of L-types is $1 - \lambda'$. Each platform's profit can be written as

$$\Pi(e'; \lambda') = \frac{1}{2}[(1 - e')\lambda'(\alpha_H - \theta_H w_s - \theta_H w_p) + (1 - \lambda')(\alpha_H - \theta_H w_s - \theta_L w_p) - c(e')], \tag{43}$$

where e' satisfies $-\lambda'(\alpha_H - \theta_H w_s - \theta_H w_p) - c'(e') = 0$ and $\Pi(e'; \lambda')$ decreases in λ'. If $\lambda' = \lambda$, then $\Pi(e'; \lambda) < 0$ given condition (23) and $\alpha_H - \theta_H w_s < p^c$. If $\lambda' = 0$, then $e' = 0$ and $\Pi(0; 0) \ge 0$ given $\alpha_H - \theta_H w_s \ge \theta_L w_p$. So, there exists a unique $\lambda' \in [0, \lambda)$ such that $\Pi(e'; \lambda') = 0$. It follows that the mass of participating H-types is $\widehat{\lambda} = \lambda'\left(\frac{1-\lambda}{1-\lambda'}\right)$.