Regulating platform fees*

Chengsi Wang† and Julian Wright‡

June 17, 2024

Abstract

We consider platforms that help consumers discover and transact with suppliers. Such platforms have come to dominate many sectors of the economy, raising issues about the high fees they charge suppliers, especially since they tend to commoditize the suppliers they aggregate. We show that in a baseline setting, the efficient platform fee is determined by a simple formula: it equals the platform's marginal cost plus the difference between suppliers' markups on the direct channel and suppliers' markups on the platform. We explore the extent to which this simple formula provides a robust cap for regulating the platform's fee more generally.

Keywords: platforms, marketplaces, commoditization, regulation

*We thank Jay-Pil Choi, Jacques Cremer, Hanna Halaburda, Andrea Mantovani, Volker Nocke, Martin Peitz, Patrick Rey, David Ronayne, Nicolas Schutz, Tat-How Teh, Jean Tirole, as well as participants in talks at 5th Asia-Pacific Industrial Organization Conference, EPoS224 MaCCI Summer School on Platform Economics, Workshop on Platform Economics at National University of Singapore, Oxera, and the TSE Workshop: Regulating the Digital Economy for valuable comments and suggestions. We thank Tian Chuyue, Li Danhou and Zheng Zhongxi for excellent research assistance. Chengsi Wang gratefully acknowledges financial support from the Australian Research Council Discovery Project DP210102015. Julian Wright gratefully acknowledges research funding from the Singapore Ministry of Education Social Science Research Thematic Grant, MOE2017-SSRTG-023. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not reflect the views of the Singapore Ministry of Education or the Singapore Government. Julian Wright was retained on behalf of Epic Games as an independent expert in a litigation in Australia (Epic Games v. Apple) which overlapped with the period during which the article was written. All views and any errors are strictly our own.

†Department of Economics, Monash University; chengsi.wang@monash.edu.

‡Department of Economics, National University of Singapore; jwright@nus.edu.sg.

1 Introduction

Regulators are struggling with the right way to address market power concerns arising from large digital platforms that act as gatekeepers for third-party suppliers, app developers, online sellers, and other small businesses to access end consumers. In Europe, the Digital Markets Act (DMA), which was recently passed, seeks to do this primarily by prohibiting various types of platform behavior: e.g. self-preferencing, price parity clauses and bundling/tying, while obliging platforms to make certain changes that are supposed to promote easier user choice and switching. It is unclear, however, the extent to which these prohibitions and obligations will really limit platforms' ability to exercise market power. This motivates our interest in another, possibly complementary, policy solution, which is the regulation of the fees charged by platforms to the suppliers that use these platforms to access consumers.

The issue of high fees charged by online platforms to third-party suppliers arises for big-tech platforms such as Amazon's marketplace, Apple's App Store, Booking's and Expedia's hotel booking sites, and Google's Play Store. More generally, large digital marketplaces have emerged across almost every sector of the economy: beauty salons (Booksy), dog walking (Rover), fashion (Zalando), handmade products (Etsy), home design (Houzz), local contractors (Task Rabbit), restaurant booking (OpenTable), and so on. As these marketplaces become the main place consumers discover and transact with third-party suppliers in a particular vertical and in a particular geography, they are able to exercise their market power by charging higher and higher fees to suppliers.[1]

To analyze whether and how one might regulate the fees set by a dominant platform, we develop a simple framework of a monopoly platform that helps consumers discover and match with listed suppliers, and charges a fee to suppliers for transactions made on the platform. The framework captures three key features of many such marketplace platforms: (i) the platform intensifies competition between listed suppliers by facilitating consumers' choice of their preferred suppliers; (ii) suppliers pass the platform's fees back to consumers via higher prices on the platform; and (iii) the platform has to attract consumers in the first place who can alternatively buy directly from suppliers. The framework is developed in a setting in which suppliers are free to set different prices on different channels, and in the baseline version of the model, consumers cannot switch to buy in the direct channel after searching

[1] Mitchell (2021) claims that after accounting for all fees in the U.S. "Amazon pockets an average of 34 percent of each sale made by independent businesses on its site. That's up from 30 percent in 2018, and 19 percent in 2014."

on the platform.

In such a setting, the reduction in suppliers' markups from intensified competition as a result of the platform is not in itself a social benefit. Consumers who use the platform do benefit from lower supplier markups, but this comes at the expense of lower supplier profit. To the extent the platform can capture increased consumer demand for using the platform through a higher fee on suppliers, the primary effect is a transfer of surplus from suppliers to the platform. The platform's fee is constrained by the fact that as it is increased further and further above cost, prices on the platform will increase, and more consumers will prefer to use the direct channel to find and transact with suppliers. But using the direct channel may be inefficient for many consumers, so this raises the question, what is the efficient fee in this setting? Perhaps surprisingly, setting the platform fee at its marginal cost does not maximize total welfare even if the suppliers on the platform fully pass through the fee into their prices and we ignore the possibility that the platform needs to set a higher fee to cover its fixed costs. Rather, we show that the efficient fee exceeds the platform's marginal cost by the extent to which the platform lowers suppliers' markups. Specifically, it equals the platform's marginal cost plus the markup differential between the two channels (i.e., the difference between suppliers' markups on the direct channel and their markups on the platform).

To understand why the efficient fee takes this form, note that intensified supplier competition on the platform comes at a cost to these suppliers, which consumers ignore. As a result, too many consumers will use the platform if it just charges a fee equal to its marginal cost. If the platform fee is increased above marginal cost by the amount to which the platform decreases supplier markups, then provided that suppliers pass this fee increase through into their prices on the platform, consumers will internalize the negative externality that their choice to use the platform has on suppliers, leading them to choose between channels efficiently. This works like a normal Pigouvian tax but has the novel feature that it is levied on the party it is designed to help (i.e., suppliers).

The idea that suppliers' interests need to be taken into account accords well with the fact that it is often the concerns of suppliers (third-party sellers on Amazon, hotels on Booking and Expedia, app developers on the App Store or the Play Store) and not consumers that are front and center in discussions around the need for regulatory intervention on these platforms. A key concern is that suppliers are being commoditized by these large marketplaces, which are built on the back of their supply. Lu (2020) puts it in stark terms: "In talking to merchants, it becomes apparent that it isn't so much that they dislike marketplaces inherently; it's that they dislike the symptoms that marketplaces typically engender. Every brand would love exposure to huge amounts of passing traffic — just not if that means they have to be a serf to the feudal lord of the marketplace."

The efficient fee we propose is relatively straightforward to implement. Other than the platform's marginal cost per transaction, it can be inferred simply by observing the current prices suppliers charge on the platform and in their direct channel, as well as the platform's current fee. We propose this as a cap that the regulator would enforce on platform fees. We show that it is possible the monopoly platform's unregulated fee already satisfies this cap. This can happen when the platform lowers suppliers' markups a lot, but the platform does not create much additional surplus in terms of improved matches and reduced search costs. Otherwise, and arguably more realistically, the cap will be binding in the case of a monopoly platform. Indeed, for a quite general class of demand, we show this is true provided the platform creates positive welfare when its fee is set at its marginal cost. Moreover, we show that whenever the efficient fee cap is binding, the planner should shift down the regulated fee (towards the platform's marginal cost) when the planner puts less weight on the platform's profit and/or more weight on the surplus of consumers.

We then explore three extensions of our framework. The first extends our baseline model to allow for final demand for the suppliers' goods or services to be elastic and the pass-through of fees to be incomplete. The same simple efficient fee cap can still work in the case of incomplete pass-through, appropriately adjusted for differences in demands on the two channels if necessary, provided it is updated over time to reflect the changes in markups that will be induced by the regulation of fees. Incomplete pass-through helps rationalize why suppliers may want to lobby for lower rather than higher platform fees, even though a higher fee helps steer consumers to buy in the direct market in which firms earn a higher markup. Accounting for elastic demand is more challenging but we note that the welfare-maximizing fee is still capped by our efficient fee formula (appropriately measured to allow for elastic demand), and is lower than this cap to the extent that the regulator also values lowering suppliers' prices on the platform towards their costs to expand demand for their products.

The second extension allows for showrooming, so consumers can switch to buy directly from a supplier after discovering the supplier on the platform. We show that the same simple rule to set the efficient fee applies, although the markup differential is lower due to showrooming, thus implying showrooming lowers the level of the efficient fee. Using this showrooming extension, we also compare fee regulation with an alternative policy which has come into focus in recent years. This policy requires platforms to remove rules which stop suppliers informing their customers on the platform of their own direct channels. Removing these rules makes showrooming easier. Finally, in our third extension, we allow for platform competition, showing how platform competition pushes the equilibrium fee towards cost, thus making it less likely the regulatory cap on fees is needed.

As will be explained in more detail in the next section, our paper complements the work of Gomes and Mantovani (2024), who also look at how to regulate platform fees but in a setting in which price parity clauses apply (so suppliers cannot set different prices on the platform than in their direct channel). Our focus on the case without price parity clauses is motivated by the fact for a range of platforms, price parity clauses have not been imposed, or in some cases, have been banned by regulators (see Baker and Scott Morton, 2018). Neither Apple nor Google imposes price parity clauses in their app stores, and app developers remain free to set lower prices for purchases in other channels (e.g. via their own websites). Price parity clauses have been removed in Europe in the case of Amazon, and in most of Europe in the case of Booking and Expedia. Under the DMA, price parity clauses will indeed not be allowed to be used by designated platforms. More generally, we have in mind understanding how to evaluate settings where price parity (either imposed directly or indirectly via steering or self-preferencing) has been addressed by regulations like the DMA. And we ask what fees platforms would then set. Would they still be set too high? If so, how could we best regulate them?

1.1 Related literature

There is surprisingly little prior research on the question of the right level at which to regulate prices set by digital platforms. There is an earlier literature that considers efficient pricing in a generic monopoly two-sided platform in which transactions between the two sides are not modelled explicitly (e.g., Rochet and Tirole, 2003, Armstrong, 2006, and Weyl, 2010). However, this literature fundamentally differs because it focuses on the efficient price structure across the two sides in order to get the right balance of participation on each side. Price structure is not an issue in our setting given the platform only charges fees to the seller-side.[2] Instead, we

[2]This is consistent with the fact that we model transactions between the two sides whereas the earlier literature treated the benefits to each side as exogenously given. Endogenizing the pricing of

focus on a completely different margin: whether consumers will use the platform or purchase from suppliers directly.

One exception to considering only generic two-sided platforms when characterizing efficient fees is the literature on payment card platforms, where the transactions between buyers and sellers are modelled explicitly and the focus is on choosing the structure of fees to induce the right level of card usage versus cash usage by consumers (Rochet and Tirole, 2002 and 2011, and Wright, 2004). Indeed, our paper is in part inspired by the work of Rochet and Tirole (2011) who propose a simple rule that can be used to regulate interchange fees (the so-called "Merchant Indifferent Test"), one that has been adopted by regulators in Europe, among other places. Their setting is different, however, for two main reasons: (i) unlike the types of marketplace platforms we're focused on in this paper, card platforms do not help intensify competition between suppliers given they are not primarily used to discover suppliers; (ii) a no-surcharge rule applies, so suppliers are not allowed to set a higher price to consumers who purchase using the card platform than those who pay with cash.

Gomes and Mantovani (2024) is the first paper to consider regulation of platform fees in marketplace settings. But they maintain the assumption of a single price across the platform channel and the direct channel (i.e., price parity) like the payment cards literature. That is, they relax (i) but not (ii). In their setting, the platform expands the consideration set of consumers and in so doing also intensifies competition between suppliers. Under price parity, they show the platform's unregulated fee to suppliers is excessive.[3] Given they focus on price parity holding, not surprisingly, their characterization of the socially optimal fee is different from ours. Under price parity there is no role for consumers' channel choice to be influenced by fees, which is what drives our results without price parity. This is why our framework is not applicable when price parity holds, and a setting like theirs is more appropriate. In their setting, it is the extensive margin between whether the platform invests or not given randomness in its fixed cost of investment that pins down the efficient fee. Specifically, their efficient fee is determined by the extent to which the platform expands consumers' consideration set as well as any

suppliers in the context of a platform charging transaction fees on both sides leads to a neutral fee structure. One could then normalize the buyer fee to zero, so that only the level of the seller-side fee would matter.

[3]Other papers also look at settings in which price parity clauses hold (e.g. see Boik and Corts, 2016, Edelman and Wright, 2015, Johnson, 2017, Ronayne and Taylor, 2020, and Wang and Wright, 2020) but they differ in not exploring how to regulate the platform's fee.

convenience benefits it provides to suppliers. We will directly compare Gomes and Mantovani's characterization of the efficient fee with ours under the same model of supplier competition.

Two other recent papers explore caps on the fees platforms charge suppliers, but in contrast to our paper, they take into account the possibility that the platforms also charge fees on the consumer side. Bisceglia and Tirole (2023) views the lack of a platform fee to consumers as an indication that the platform would like to set negative prices to consumers if such fees were feasible, and explores the consequences of this missing price (as well as a zero lower bound for the supplier's own prices) for the efficient cap. Their setting is quite different from ours in that the platform operates a hybrid marketplace and there is no direct channel. Their main focus is on the interplay between the platform fee to suppliers and whether the platform steers consumers to its own apps or squeezes (or forecloses) third-party apps. Sullivan (2022) empirically studies commission caps on food delivery platforms, and after taking into account that such caps increase the platforms' delivery fees to consumers, he finds they lower consumer surplus and total welfare.

Our analysis of a policy which requires platforms to remove their rules which try to prevent showrooming obviously relates to the recent literature studying showrooming on platforms (Bergemann and Bonatti, 2023; Hagiu and Wright, 2024; Wang and Wright, 2020, 2023). That literature focuses on the implications of showrooming on platforms, but does not analyze a policy which seeks to enable it. Such an analysis is timely, given as we will detail, such policies have been recently proposed by regulators, and indeed implemented following judgements in several competition cases.

Finally, our paper relates to the literature modelling price comparison websites. The seminal paper is Baye and Morgan (2001), in which consumers can use the platform to find the lowest priced supplier (which are homogenous) or instead go to their local monopolist. They maintain price parity. Galeotti and Moraga-González (2011) extend their work to the case with differentiated firms, as well as allowing suppliers to set different prices across channels, like in our paper. A key difference in these papers is that they assume the platform can set a fixed fee to each side (both consumers and suppliers), and consumers all face the same fixed benefit of shopping via the platform relative to shopping in the direct channel. Thus, they shut down the smooth channel choice that drives our results, and the efficient fees are just set so all consumers and suppliers participate on the platform. The models of price comparison websites by Ronayne (2021) and Ronayne and Taylor (2022) are closer to our setting, since they assume, more realistically, that such platforms charge firms a per-transaction fee and nothing to consumers directly. They also allow for differential prices across channels. However, in their setting the platform fee does not affect total welfare, and their interest lies rather in whether the existence of such platforms is good for consumers, which it always is in our setting.

2 Baseline model

Suppose there are multiple suppliers (either a finite number or a continuum) producing horizontally differentiated products. For brevity, we will refer to suppliers as "firms", but the reader should keep in mind these can sometimes be individuals (e.g. a dog walker on Rover or a local contractor on Task Rabbit). There is a unit mass of consumers, each with unit demand and wanting to make one (and only one) transaction. There is an outside option, with surplus normalized to zero. The firms' costs are normalized to zero.

Firms and consumers can trade directly. In addition, a marketplace platform M can facilitate the trades between firms and consumers at a marginal cost $c \geq 0$, and for doing so it charges firms a per-transaction fee f, a commonly used form of fee charged by such marketplaces.[4] Consumers are heterogenous in an additive benefit (if positive) or cost (if negative) associated with shopping on M, which is denoted b and is distributed across consumers according to H on $[\underline{b}, \overline{b}]$. In our baseline model, we can interpret b as either the additional benefit (or cost) of consumers joining M or the additional transaction benefits or costs of using M for a transaction. We assume a strictly positive density h and a weakly increasing hazard rate for H (which implies that demand for M as defined by $1 - H(\cdot)$ is weakly log-concave). Corresponding to this, we define $\lambda(x) = (1 - H(x))/h(x)$ as the inverse hazard rate, which is weakly decreasing.

We will adopt a fairly general reduced form approach to model transactions between firms and consumers. The game has three stages. In stage 1, M sets the fee f, which is publicly observed. In stage 2, each firm simultaneously and independently chooses a direct price, and whether to list on M, and if so, their price on M. At the same time, based on their draw of b, consumers simultaneously and

[4]As detailed in https://www.theverge.com/21445923/platform-fees-apps-games-business-marketplace-apple-google, marketplaces typically charge firms (sellers and developers) on a per transaction basis. Often such a fee is written as a percentage of the value of a transaction rather than a fixed amount per transaction. We adopt the latter type of fee for tractability. In the Online Appendix we show how our analysis can be modified to handle percentage fees.

independently choose which channel they use to shop. They can choose only one channel (something we will relax in our showrooming extension in Section 4.2). In stage 3, consumers make search and purchase decisions on their chosen channel.

The key assumption in this timing is that an individual firm cannot influence a consumer's choice of channel by whether to list on M or what price to set. Yet M's choice of fee f ultimately impacts which channel consumers want to use since it affects the prices they expect to face on the platform.[5] This set of assumptions implies firms will always list on M to obtain incremental revenue. An individual firm that does not list on M does not attract more consumers to its direct channel since consumers do not observe whether firms are listed on M until after they have chosen which channel to go to. This is also consistent with the possibility that each individual firm is too small to influence whether consumers use the platform or not.

We next lay out the key reduced-form properties that we assume capture what happens in stages 2 and 3, properties that we will show can be derived from each of the three different models of firm competition that we will detail below. Since firms are free to price discriminate across channels, and consumers choose a channel without observing firms' actual pricing and participation choices, on-platform prices are determined independently of direct prices. In our baseline model, we focus on settings with symmetric firms, inelastic aggregate demand and full pass-through, so the symmetric equilibrium price is always equal to the sum of the fee (equal to zero on the direct channel) and a constant markup. Let $p_D = \mu_D$ be the symmetric equilibrium direct price, where μ_D represents firms' symmetric markup in the direct channel, and $p_M = f + \mu_M$ be the equilibrium price on M, where μ_M is firms' symmetric markup on M. Second, as long as the platform attracts all firms, the gross expected surplus to consumers on the platform ϕ_M is greater than that on the direct channel ϕ_D, i.e., we assume $\phi_M > \phi_D$. The platform offers some efficiencies (e.g. reduced search costs or more firms to choose from). Third, the competition between firms is assumed to be more intense on the platform than the direct channel, i.e., we assume $\mu_M < \mu_D$. Naturally, we assume $\phi_D \geq \mu_D$, so that positive expected net surplus arises from transactions on the direct channel, which together with the other assumptions implies $\phi_M > \mu_M$. Finally, we assume that for the relevant fees under consideration, all consumers who go to M always make a single transaction on M and get non-negative net surplus from doing so. Each of these reduced-form

[5] We could alternatively assume that consumers cannot observe f, but they can observe the modal price set by firms on M, which would allow them to infer f provided there are three or more symmetric firms.

properties will be endogenously derived, and this last property will be confirmed, in the three microfounded applications detailed below.

Firms often obtain extra transaction benefits (or equivalently, face lower marginal costs) when selling through the platform. We do not explicitly allow for such benefits in our baseline model given any homogenous firm-side transaction benefits can be implicitly captured by shifting up the level of consumer-side benefits b. This reflects that any firm-side transaction benefits would be fully passed on to consumers via firms setting lower prices on the platform given our assumption of full pass-through above.

Some examples of micro-founded settings that fit this baseline model include[6]:

- *Sequential search model* such as in Wolinsky (1986) and Anderson and Renault (1999). There is a continuum of firms with consumers' match values drawn iid from a distribution $G(\cdot)$ which is assumed to have an increasing hazard rate. All firms will be available on either channel given the result above that all firms will want to list on M in equilibrium. After choosing a channel in stage 2, in stage 3 consumers search sequentially in their chosen channel to discover match values and prices of individual firms, but search costs are lower on M than the direct channel as in Wang and Wright (2020). Formally, $s_M < s_D$, where s_i is each consumer's search cost for sampling a firm on channel i. If x_i represents the corresponding equilibrium reservation utility for searching in channel i, then $\phi_M = x_M$, $\phi_D = x_D$, $\mu_M = \frac{1-G(x_M)}{g(x_M)}$ and $\mu_D = \frac{1-G(x_D)}{g(x_D)}$, with $x_M > x_D$ given $s_M < s_D$ (consumers have a higher reservation utility on M due to lower search costs on M), which implies $\phi_M > \phi_D$, and also implies $\mu_D > \mu_M$ given G has an increasing hazard rate. We assume the search cost in the direct channel is sufficiently small such that $\phi_D \geq \mu_D$ as otherwise no one will ever search directly.

- *Random-utility model* with $n \geq 3$ firms (Perloff and Salop, 1985). This is the competition model adopted by Gomes and Mantovani (2024). We will use it to compare our results with those in Gomes and Mantovani in Section 3.2.3. The utility a consumer can get from buying at firm i is $u^i = v - p^i + \beta\xi^i$, where ξ^i, which is drawn by consumers in stage 3, is iid across firms and consumers from a distribution G, and $\beta > 0$ measures the importance of the

[6] The full details of each model and the assumptions required to fit our general setting are given in the Online Appendix. A key assumption in each case is inelastic aggregate demand and full pass-through. Thus, for instance, the standard representative consumer model with differentiated demand does not fit. In Section 4.1 we generalize the framework to handle such settings.

match value. In stage 3, consumers also randomly draw a set of n_D firms ($2 \leq n_D < n$) if they are in the direct channel, whereas if they go on M they can choose from all the firms that list on M, which as noted above, will be all n firms in equilibrium. The difference between n and n_D then drives the differences between match values and markups. For example, if G is a uniform distribution on $[0,1]$, $\phi_M = v + \frac{\beta n}{n+1}$ and $\phi_D = v + \frac{\beta n_D}{n_D+1}$, with $\phi_M > \phi_D$, and $\mu_M = \frac{\beta}{n}$ and $\mu_D = \frac{\beta}{n_D}$ with $\mu_D > \mu_M$. The more general expressions for the differences in these expressions across the two channels are given in (7)-(8).

- *Circular-city model* with $n \geq 2$ firms that are located equidistant from each other along the circumference of a circle (Salop, 1979). Each firm offers a good of value v, and consumers face a standard linear mismatch cost parameter t, with their location (relative to the firms) drawn in stage 3. In stage 3, consumers also randomly draw a single firm if they are in the direct channel, whereas if they go on M they can choose from all the firms that list on M, which as noted above, will be all n firms in equilibrium. The larger number of firms on the platform drives the differences between match values and markups. In equilibrium, we have $\phi_M = v - \frac{t}{4n}$ and $\phi_D = v - \frac{t}{4}$, with $\phi_M > \phi_D$, and $\mu_M = \frac{t}{n}$ and $\mu_D = v - \frac{t}{2}$, with $\mu_D > \mu_M$ given $v > t$ and $n \geq 2$. Two special cases of this setting are: (i) the Hotelling model when $n = 2$; (ii) Bertrand competition if we take $t \to 0$, so $\phi_D \to \phi_M$ and $\mu_D - \mu_M \to v$.

3 Analysis of the baseline model

Consumers' expected net utility of shopping directly is $\phi_D - p_D = \phi_D - \mu_D$, and their expected net utility of shopping on M is $\phi_M - p_M + b = \phi_M - (f + \mu_M) + b$. Consumers choose M if and only if they draw b such that

$$\phi_M - (f + \mu_M) + b \geq \phi_D - \mu_D \ \Leftrightarrow \ b \geq f - (\phi_M - \phi_D) - (\mu_D - \mu_M).$$

Define $\Delta_s \equiv \phi_M - \phi_D$ as the surplus differential and $\Delta_m \equiv \mu_D - \mu_M$ as the markup differential, both of which are positive given our assumptions. The condition for consumers to choose M becomes

$$b \geq f - \Delta_s - \Delta_m,$$

which makes clear the only reason a consumer uses M is if the transaction benefit b they obtain of doing so plus the surplus and markup differential that M creates more than covers the fee it charges.

Throughout the paper we adopt the very weak condition that

$$\overline{b} + \Delta_s + \Delta_m > c. \tag{1}$$

This just requires that the platform can create positive expected net surplus for the consumer who values using it the most. For some propositions, we will make use of a somewhat stronger but still reasonable condition:

$$\int_{c-(\Delta_s+\Delta_m)}^{\overline{b}} (\Delta_s + b - c) dH(b) > 0. \tag{2}$$

This requires that the existence of the platform increases total welfare when its fee is set at marginal cost.

Taking these preliminaries into account, we proceed to characterize (i) the profit-maximizing platform fee, (ii) the welfare-maximizing platform fee as well as the fee maximizing other welfare benchmarks, (iii) the comparison between the profit-maximizing and welfare-maximizing fees, and (iv) comparative statics on these fees.

3.1 Profit-maximizing platform fee

As an unregulated monopolist, M chooses f to maximize

$$(f - c)(1 - H(f - \Delta_s - \Delta_m)).$$

Note that the resulting fee would be equal to cost c only if the demand $1 - H(f - \Delta_s - \Delta_m) = 0$ for all $f > c$. Otherwise, M would be better off by setting some fee strictly above its marginal cost. The condition that $1 - H(f - \Delta_s - \Delta_m) = 0$, or equivalently, $H(f - \Delta_s - \Delta_m) = 1$, for all $f > c$ is equivalent to $\overline{b} + \Delta_s + \Delta_m \leq c$, which (1) rules out. Thus, we obtain the following characterization of M's optimal fee (as with other results not proven in the text, the proof is given in the Appendix):

Proposition 1. (The platform's profit maximizing fee)
The platform sets f^*, where f^* is the unique solution to

$$f^* = c + \lambda (f^* - \Delta_s - \Delta_m), \tag{3}$$

and satisfies $c < f^* < \bar{b} + \Delta_s + \Delta_m$*, where* λ *is the inverse hazard rate of the distribution function* H*.*

M trades off the higher margin it gets from setting a higher fee on each transaction it enables with the cost of fewer consumers coming to it to make transactions given prices will be higher on M. We can illustrate this result when H takes the generalized Pareto distribution (GPD), which covers several well-known distributions such as uniform, exponential, normal, logistic, type I extreme value, and Weibull.

Example 1 (Generalized Pareto distribution)**.** *Given our assumption of a weakly increasing hazard rate, when* H *takes the generalized Pareto distribution (GPD) form, it can be written as*[7]

$$H(b) = \begin{cases} 1 - \left(1 - \frac{\epsilon(b - \underline{b})}{\sigma}\right)^{\frac{1}{\epsilon}} & \text{if } \epsilon > 0 \\ 1 - e^{-\frac{b - \underline{b}}{\sigma}} & \text{if } \epsilon = 0 \end{cases}$$

over the support $\underline{b} \leq b \leq \bar{b}$*, where* $\bar{b} = \underline{b} + \frac{\sigma}{\epsilon}$*. Note* $\lambda(b) = \sigma - \epsilon(b - \underline{b})$*. Then (3) implies*

$$f^* = \frac{c + \sigma + \epsilon(\underline{b} + \Delta_s + \Delta_m)}{1 + \epsilon}$$

so $f^* = \frac{c + \underline{b} + \sigma + \Delta_s + \Delta_m}{2}$ *when* $\epsilon = 1$ *(uniform distribution) and* $f^* = c + \sigma$ *when* $\epsilon = 0$ *(exponential distribution).*

3.2 Welfare analysis

In this section we characterize how a planner optimally chooses the platform's fee f assuming firms are still free to set their final prices on each channel. To keep things general initially, consider the planner's objective being a weighted average of the different surplus components making up total welfare, which can be written as

[7]To make our results easier to interpret, we define the shape parameter ϵ so it is non-negative, with a higher value of ϵ representing a more concave distribution for H.

$$
\begin{aligned}
W^w = & w_c\left(\int_{f-(\Delta_s+\Delta_m)}^{\bar{b}}(\phi_M+b-f-\mu_M)dH(b)+\int_{\underline{b}}^{f-(\Delta_s+\Delta_m)}(\phi_D-\mu_D)\,dH(b)\right) \\
& + w_f\left(\int_{f-(\Delta_s+\Delta_m)}^{\bar{b}}\mu_M dH(b)+\int_{\underline{b}}^{f-(\Delta_s+\Delta_m)}\mu_D dH(b)\right) \\
& + w_m\int_{f-(\Delta_s+\Delta_m)}^{\bar{b}}(f-c)\,dH(b),
\end{aligned}
$$

where the terms in the expression are consumer surplus (the first line), firms' total profit (the second line), and platform profit (the third line), with the respective weights satisfying $w_c + w_f + w_m = 1$. This allows us to see how the interests of the three different constituents (consumers, firms, and the platform) diverge. After simplifying, the derivative of W^w with respect to f is

$$
\begin{aligned}
& w_f\Delta_m h\left(f-(\Delta_s+\Delta_m)\right)-w_c\left(1-H\left(f-(\Delta_s+\Delta_m)\right)\right) \\
& +w_m\left(1-H\left(f-(\Delta_s+\Delta_m)\right)-(f-c)\,h\left(f-(\Delta_s+\Delta_m)\right)\right).
\end{aligned} \tag{4}
$$

3.2.1 Consumer surplus only

Clearly, when considering consumers' surplus only ($w_f = 0$ and $w_m = 0$), the expression in (4) is always negative, so consumer surplus is maximized by setting f as low as possible. This lowers the firms' prices on M while not affecting the firms' prices in the direct channel. But the planner would not want to set $f < c$ since then M wouldn't operate, which is even worse for consumers. Consumer surplus is always higher when M operates (regardless of the fee) because consumers are always free to choose the channel which makes them better off, which for some consumers can be M. Thus, maximizing consumer surplus involves setting the fee as low as possible while ensuring M still wants to operate, which in our setting (without any fixed costs to be recovered via the transaction fee f) involves setting M's fee at marginal cost.

Proposition 2. (Consumer surplus standard) *The fee that maximizes consumer surplus is* $f^{cs} = c$.

3.2.2 Consumer and platform surplus only

As a benchmark to understand our more general welfare result, next consider welfare absent firms' surplus (i.e., $w_f = 0$, with $w_c = w_m$). Then the derivative of W^w in (4) with respect to f simplifies to $-(f-c)\,h\,(f-(\Delta_s+\Delta_m))$, so W^w is single-peaked around $f=c$ in this case, and we have:

Proposition 3. *(Ignoring firms' surplus) The fee that maximizes consumer surplus and platform profit is $f^{cp}=c$.*

This result highlights that it is the interests of third-party suppliers (i.e., the firms) that causes the efficient fee to deviate from simply being set based on M's marginal costs.

3.2.3 Total welfare standard

Next we add back in firms' surplus to weighted welfare, and consider the welfare standard that is most commonly adopted in economics, total welfare, which corresponds to $w_m=w_f=w_c$. Then (4) simplifies to $-(f-c-\Delta_m)h\,(f-(\Delta_s+\Delta_m))$, so W^w is single-peaked around $f=c+\Delta_m$ in this case, and we get our main welfare result[8]

Proposition 4. (The planner's welfare maximizing fee)
The fee that maximizes total welfare is

$$f^e=c+\Delta_m. \tag{5}$$

Proposition 4 says from an efficiency perspective, M's fee should be set above its marginal cost by the extent to which the platform lowers firms' markups. The result is simple yet surprising at first glance. Why should the efficient fee be anything other than M's marginal cost? Indeed, from an efficiency perspective, the only difference in the price consumers should face across the two channels is the marginal cost c that M faces to provide its intermediation service. However, given markups are lower on M (i.e., $\Delta_m>0$), in equilibrium the difference in the price consumers face across the two channels will be less than c if M's fee is regulated to c, which is why the efficient fee is higher than c in order to restore the correct price differential. Put differently, the markup differential makes consumers favor M, and as a result too many consumers choose M. The planner uses a fee above cost to correct for

[8] In case $\underline{b}\geq c-\Delta_s$, efficiency requires all consumers use the platform. The fee proposed in (5) induces such an outcome, even though a range of fees around that level would also induce the same outcome. A parallel argument holds when $\overline{b}<c-\Delta_s$ such that efficiency requires no consumers use the platform.

this distortion. Formally, if $f = c$, consumers will choose M if $\phi_M - c - \mu_M + b \geq \phi_D - \mu_D$, whereas in the efficient outcome, consumers should choose M if and only if $\phi_M - c + b \geq \phi_D$ (i.e., the difference in markups is removed in the efficient solution).[9]

Another way to understand why the efficient fee exceeds c is in terms of externalities. When consumers decide to use M, they do not take into account the negative effect of their choice on firms who earn lower margins on this channel. If they did, there would be no need for the fee to be set above c. By setting a higher fee, consumers pay a Pigouvian tax for using M that equals the loss in firms' markups that results from their choice, thereby getting them to internalize the full effects of their choice. A novel feature of this tax is that it is imposed on the side it is designed to help, i.e., firms, reflecting that the platform only charges firms and not consumers directly.

The efficient fee f^e that we identify is relatively easy to implement. Note that the formula in (5) can be re-written as

$$f^e = c + \Delta_m = c + p_D - (p_M - f^*) = c + f^* + p_D - p_M. \tag{6}$$

Each of p_D, p_M, and f^* should be directly observable, so the regulation only requires working out the platform's marginal cost of providing the relevant service, which may be considered negligible for some digital platforms.

In practice, consumers might sometimes not find any product they would want to purchase in the direct market, in which case p_D may overstate the margin that firms earn in the direct market. For example, suppose consumers, if shopping directly, can only find a desired product category with probability $\gamma \in (0, 1)$; they subsequently obtain a utility $\phi_D - \mu_D$ in this case. With probability $1 - \gamma$, the desired category does not exist in the direct market and the consumer obtains a zero utility. So a consumer only gets expected utility $\gamma(\phi_D - \mu_D)$ on the direct channel. In this case we can redefine $\Delta_s = \phi_M - \gamma\phi_D$ and $\Delta_m = \gamma\mu_D - \mu_M$. The efficient fee is still $f^e = c + \Delta_m$, but now Δ_m is lower, which implies the efficient fee is lower than when $\gamma = 1$.

We can directly compare our characterization of the efficient fee to that in Gomes and Mantovani (2024) by assuming, as they do, $c = 0$. In their mature market setting in which there is no positive latent demand, and assuming competition is determined by the random-utility framework in which consumers get to see n_D firms

[9] Even if consumers are homogenous with respect to b, so M would set $f = \Delta_s + \Delta_m + b$ and all consumers would use M, the efficient fee formula in (5) remains relevant. Efficiency requires consumers use M if and only if $\Delta_s + b \geq c$, a condition that is ensured by (5).

(with i.i.d. match value ξ and cumulative distribution function $G(\xi)$) in the direct market and n such firms on the platform, they find the efficient fee equals $\Delta_s + b_f$, where

$$\Delta_s = \beta \left(\int_{\underline{\xi}}^{\overline{\xi}} \xi dG(\xi)^n - \int_{\underline{\xi}}^{\overline{\xi}} \xi dG(\xi)^{n_D} \right) \tag{7}$$

and b_f is the convenience benefit they assume firms get from on-platform transactions.[10] This compares to the efficient fee in our setting for the same competition model, which is Δ_m, where

$$\Delta_m = \beta \left(\frac{1}{n_D \int_{\underline{\xi}}^{\overline{\xi}} g(\xi) dG(\xi)^{n_D - 1}} - \frac{1}{n \int_{\underline{\xi}}^{\overline{\xi}} g(\xi) dG(\xi)^{n-1}} \right). \tag{8}$$

The parameters β, n_D and n have the same qualitative effects on the efficient fee across both settings, although for very different reasons. In Gomes and Mantovani, the effect of these parameters is via the surplus differential, which provides a reason to make sure the platform will want to invest in providing its service when it is efficient to do so, which in the face of uncertain investment costs is achieved by setting a sufficiently positive fee. In our setting, the effect of these parameters is via the markup differential, which requires a sufficiently positive fee to offset excessive use of the platform by consumers. Another difference in our characterization of the efficient fee is that allowing firms to enjoy a convenience benefit from using the platform of b_f would not change our efficient fee formula (5) since such a benefit would be like a negative marginal cost for firms — it would lower their equilibrium prices on M since it would be fully passed through by firms, which would induce consumers to correctly take it into account when deciding which channel to choose.

3.2.4 Total user surplus

Often policymakers will want to put more weight on the surplus of consumers than that of a monopoly firm selling to those consumers. In a two-sided platform setting, the platform's customers consist of the users on both sides of the platform (here both final consumers and the competing firms that want to reach them). The interests of the firm side of the platform may be particularly relevant in a setting where the firms involved are often individuals or small businesses. Thus, it is natural to explore how fees should be set when less weight is put on the platform's profit than that of its users (consumers and firms).

[10] The derivation of (7)-(8) is given in the Online Appendix.

In the extreme case in which no weight is put on the platform's profit (i.e., $w_m = 0$) and equal weight is put on consumers' and firms' interests (i.e., $w_c = w_f$), weighted welfare becomes equal to total user surplus. In this case we find:

Proposition 5. (Total user surplus standard) *Suppose (2) holds. The fee that maximizes total user surplus is $f^{tus} = c$.*

In the proof of Proposition 5 we show that total user surplus is U-shaped (or in case b is exponentially distributed, it is weakly monotone) over the range of f for which the platform would operate. Given this property, the planner prefers either the lowest feasible fee, which just allows M to cover its costs, or a fee that causes there to be no transactions on the platform. The latter is any fee high enough that even the consumer with the highest possible b has no reason to go to M. Equivalently, the latter can be induced by setting $f < c$, so the platform would not operate. However, given our assumption (2) which requires the platform's existence to increases total welfare when its fee is set at marginal cost, total user surplus (which equals total welfare when $f = c$, or when M does not operate) must be maximized by setting the lowest fee at which the platform operates (i.e., $f^{tus} = c$).[11]

3.2.5 Weighted welfare

Finally, let us return to the original problem we started with of choosing a platform fee to maximize weighted welfare W^w. Suppose, the planner puts at least as much weight on consumer surplus as it does on the firms' profit or the platform's profit. We adopt the GPD form for H which rules out W^w having multiple turning points.

Proposition 6. (Weighted average of total surplus) *Assume $w_c \geq \max\{w_f, w_m\}$ so the planner puts at least as much weight on the interests of end-consumers as on the interests of firms or the platform. Also assume (2) holds. When H takes the GPD form, the fee that maximizes a weighted average of consumer surplus, firms' profit and the platform's profit lies between c and f^e.*

The proposition further justifies using f^e as a cap for regulating the platform's fee. To the extent the planner may put less weight on the platform's profit and/or

[11] A separate question is whether total user surplus is enhanced by the existence of M when the fee is set at the efficient level f^e. In the Online Appendix we explore this question. We show in contrast to Gomes and Mantovani (2024), this is not necessarily the case in our setting. Instead we provide a lower cap on the regulated fee (between c and f^e) such that M's existence would always increase total user surplus.

the firms' profits compared to that of consumers, its preferred fee is always above the platform's marginal cost but below the efficient fee as characterized in (5).

We can illustrate this result in the special case that H follows the exponential distribution (i.e., $\epsilon = 0$), in which case because λ is constant, W^w is single-peaked in f around f^w over the relevant range of f, where

$$f^w = c + \frac{w_f}{w_m}\Delta_m - \sigma\left(\frac{w_c - w_m}{w_m}\right).$$

There are two cases of particular interest:

- Set $w_c = \frac{1}{3} + \frac{\alpha}{3} > w_m = w_f = \frac{1}{3} - \frac{\alpha}{6}$ so consumer surplus is weighted more heavily than profit (firms' and M's). Then provided $\alpha < 2$ so $w_m > 0$ and $w_f > 0$,

$$\begin{aligned} f^w &= c + \Delta_m - \frac{3\alpha}{2-\alpha}\sigma \\ &= f^e - \frac{3\alpha}{2-\alpha}\left(f^* - c\right). \end{aligned}$$

- Set $w_c = w_f = \frac{1}{3} + \frac{\alpha}{6} > w_m = \frac{1}{3} - \frac{\alpha}{3}$ so the surplus of users (consumers and firms) is weighted more heavily than M's profit. Then provided $\alpha < 1$ so $w_m > 0$,

$$\begin{aligned} f^w &= c + \Delta_m - \frac{3\alpha}{2\left(1-\alpha\right)}\left(\sigma - \Delta_m\right) \\ &= f^e - \frac{3\alpha}{2\left(1-\alpha\right)}\left(f^* - f^e\right). \end{aligned}$$

These results show that when consumer surplus is weighted more than profits (either of firms or M's), the weighted-welfare maximizing fee is necessarily less than f^e, and when the surplus of users (consumers and firms) is weighted more than the profit of M, this remains true provided f^* exceeds f^e which will be true if (2) holds. Moreover, the informational requirements of these two solutions are no more demanding than for regulating the efficient fee. The only additional requirement is the planner has to pick the weights it wants to put on the different types of participants (i.e., α).

3.3 Comparing the unregulated and efficient fees

The next proposition compares the platform's unregulated (i.e., profit-maximizing) fee in (3) and the efficient (i.e., welfare-maximizing) fee in (5).

Proposition 7. (Comparison of fees) *The profit-maximizing fee exceeds the efficient fee iff* $\lambda (c - \Delta_s) \geq \Delta_m$.

Given firms' markups are higher in the direct channel than on the platform (i.e., $\Delta_m > 0$), Proposition 7 does not rule out the possibility that M's unregulated fee is lower than the efficient fee. A case where M's unregulated fee is always too low is when M's added social value does not cover its marginal cost so $\Delta_s \leq c$ and it is always costly for consumers to go to the platform so $\overline{b} = 0$. The planner would then prefer a higher fee to reduce the number of consumers going to M. This also implies a platform that drastically reduces markups while producing little or no gains in gross surplus, may in fact decrease welfare, as consumers over-use it (in the sense that for many transactions the social gain is less than the marginal cost). Underlying this result is that individual firms cannot circumvent M by delisting and redirecting consumers to their direct channel. This reflects our assumption that each individual firm's delisting decisions are not observable, and which more generally captures the idea that each individual firm is small. If firms were large and established, a platform that destroys welfare would unlikely be able to exist because each such firm could redirect its consumers to its direct channel via delisting.

Given the marginal cost of digital platforms c is likely to be small, and marketplace platforms do typically provide significant benefits to help consumers better search for firms as well as providing other types of transactions benefits for at least some consumers (e.g. payment services, official receipts, insurance, dispute resolution etc), we do not think the above case where the platform sets its fee too low is very relevant for the types of platforms that regulators are concerned about. The possibility that the unregulated platform fee is too low is even less likely to be relevant when less weight is put on the platform's profit (as shown in Section 3.2). Moreover, by imposing the efficient fee formula (5) as a cap on the fee the platform is allowed to set, the regulation only binds in case the platform indeed sets its fee too high.

When H takes the GPD form, we can show M's unregulated fee exceeds the efficient fee under a fairly weak condition.

Proposition 8. (Comparison of fees under Generalized Pareto distribution) *When*

H takes the GPD form, M's profit-maximizing fee f^* exceeds the efficient fee f^e if and only if (2) holds, or equivalently if and only if

$$\sigma + \epsilon \left(\underline{b} + \Delta_s - c \right) \geq \Delta_m . \tag{9}$$

Recall (2) just says the platform creates positive welfare when its fee is set at its marginal cost. To the extent (2) is expected to be true, Proposition 8 lends support to the view that we would expect such platforms to set their fees too high.

3.4 Comparative statics

As noted earlier, f^* is increasing in Δ_m and Δ_s. Moreover, f^e as defined in (5) is increasing in Δ_m. Since both Δ_m and Δ_s depend on the underlying primitives of competition arising in each channel, we can ask how changes in the primitives of competition affect the choices of f^* and f^e. We also define the tendency of M to set its fee too high as

$$L = f^* - f^e,$$

and consider how L changes in the primitives for each of our three competition applications from Section 2. An increase in L means the tradeoff shifts towards the unregulated monopoly fee becoming more excessive.

Suppose the primitive of interest is called x. From (3), we have

$$\frac{df^*}{dx} = \frac{-\lambda'}{1-\lambda'} \left(\frac{d\Delta_s}{dx} + \frac{d\Delta_m}{dx} \right).$$

Also, $\frac{df^e}{dx} = \frac{d\Delta_m}{dx}$. Thus,

$$\frac{dL}{dx} = -\left(\frac{1}{1-\lambda'} \right) \frac{d\Delta_m}{dx} - \left(\frac{\lambda'}{1-\lambda'} \right) \frac{d\Delta_s}{dx},$$

and since $\lambda' \leq 0$, we have that

$$\frac{dL}{dx} \geq 0 \quad \Leftrightarrow \quad \frac{d\Delta_m}{dx} + \lambda' \frac{d\Delta_s}{dx} \leq 0. \tag{10}$$

In the proposition that follows, λ' is evaluated at $f = f^*$.[12]

[12] For brevity, we have only stated shortened results here. The full statement of comparative static results along with the proof is given in the Online Appendix.

Proposition 9. (Comparative statics) *The level of f^*, f^e and the tendency for the platform to set its fee too high (L) change with the primitives of the respective competition models in the following ways:*

- *Sequential search model: Let λ_G be the inverse hazard rate of the distribution G on the match value ξ, which is decreasing. A decrease in search costs s_M on the platform or an increase in search costs s_D in the direct channel increases both f^* and f^e, and decreases L if and only if $|\lambda'| \leq |\lambda'_G|$.*

- *Random-utility model: An increase in the importance of match value (β), an increase in the number of firms n that consumers can evaluate on the platform channel, or a decrease in the number of firms n_D that consumers can evaluate in the direct channel increases both f^* and f^e, and decreases L if and only if $|\lambda'|$ is sufficiently small.*

- *Circular-city model: An increase in the number of firms n listed on M increases both f^* and f^e, and decreases L if and only if $|\lambda'| < 4$. An increase in product differentiation between firms t, increases f^*, decreases f^e, and always increases L.*

The main result from Proposition 9 is that M's tendency of setting an excessive fee decreases with factors that increase the firms' competitiveness on M relative to the direct channel provided H is not too concave. An increase in firms' relative competitiveness on M implies that the markup differential (Δ_m) and the surplus differential (Δ_s) both increase, as M's advantage over the direct channel becomes even stronger. But note from Proposition 7 that Δ_m and Δ_s have opposite effects in determining whether M's unregulated fee is excessive. An increase in Δ_s leads to an increase in the unregulated fee (although not one-for-one given log-concave demand), but has no effect on the efficient fee as consumers already take into account the surplus differential when making their choice of channel. On the other hand, an increase in Δ_m leads to an increase in the efficient fee (one-for-one) but it does not get fully passed through into the unregulated fee (given log-concave demand) so does not increase the unregulated fee as much. This means an increase in Δ_s shifts the tradeoff towards the unregulated fee being too high, while an increase in Δ_m shifts the tradeoff towards the unregulated fee being too low. When H is not too concave, the demand faced by M (i.e., $1 - H$) will be sufficiently concave, and the pass-through of changes in Δ_s and Δ_m to the unregulated fee will be small, implying the result will be dominated by the effect of Δ_m on the efficient fee. Thus, the increase in the markup differential as a result of an increase in the relative competitiveness of firms on M pushes up the efficient fee relative to the unregulated fee, and results in less excessive fees.

4 Extensions

In this section we consider three important extensions of our baseline model.

4.1 Incomplete pass-through and elastic demand

Consider the generalization of our baseline model to allow for elastic demand and incomplete pass-through. In addition to generalizing our efficient fee cap formula, allowing for incomplete pass-through also provides a way to understand why collectively firms may prefer lower rather than higher fees.

Consumers get gross utility $u_j(q)$ from buying q units from a firm on channel j, and consumers will buy $q_j(p)$ units from the firm on channel j facing a price of p, where $j = M$ for the platform channel and $j = D$ for the direct channel. Since it can matter to the results once we relax the assumption of unit demands, we allow firms to face a marginal cost of d on each channel. Firms set a symmetric price of $p_M(f)$ on the platform where incomplete pass-through means $0 < p'_M(f) < 1$, and obtain a corresponding profit per consumer of $\pi_M(f) = (p_M(f) - f - d)\, q_M(f)$, where with some abuse of notation, we define $q_M(f) = q_M(p_M(f))$. And suppose, in the direct channel, they set a symmetric price of p_D and obtain a corresponding profit per consumer of $\pi_D = (p_D - d)\, q_D(p_D)$. Firms therefore will continue to join M since in equilibrium they must set $p_M(f) \geq f + d$, given they would not sell at a loss. Consumers get a corresponding net surplus of $u_M(f)$ and u_D in the respective channels.

Consumers choose M over the direct channel if $u_M(f) + b \geq u_D$, or equivalently, if

$$b \geq \Theta(f) \equiv u_D - u_M(f).$$

Consider total welfare, which is

$$W(f) = \int_{\Theta(f)}^{\bar{b}} (u_M(f) + \pi_M(f) + (f - c)\, q_M(f) + b)\, dH(b) + \int_{\underline{b}}^{\Theta(f)} (u_D + \pi_D)\, dH(b). \tag{11}$$

The derivative of welfare with respect to f is the sum of two terms: a *channel selection effect*

$$\left(\pi_D - \pi_M(f) - (f-c)\, q_M(f)\right) q_M(f)\, p_M'(f)\, h(\Theta(f)), \tag{12}$$

which comes from the derivative of W through $\Theta(f)$ changing with f, and a *price compression effect*

$$\left(p_M(f) - c - d\right) q_M'(f)\left(1 - H(\Theta(f))\right), \tag{13}$$

which comes from the derivative of the surplus on M with respect to f for a given $\Theta(f)$. Note we have used that $u_M'(f) = -q_M(f)\, p_M'(f)$ from the Envelope theorem from the consumers' optimization problem to obtain this result.

The first term (12) captures essentially the same welfare tradeoff from channel selection as in the baseline model. Thus, it represents the generalization of our earlier welfare result (Proposition 4) to handle incomplete pass-through and potentially different levels of demand in each channel. The second term (13), which is entirely new, captures the additional welfare gain under elastic demand from lowering the fee to bring down firms' prices on M towards marginal cost $c+d$.

We will focus first on the planner's choice of fee considering only the channel selection effect. This is relevant if we just want to understand the effects of incomplete pass-through and potentially different demand levels in each channel, but where demand remains inelastic on M (i.e., so $q_M'(f) = 0$). To proceed, we set $q_M'(f) = 0$ and fix demands at q_D and q_M in the respective channels. Then the first-order condition from setting (12) to zero can be written as

$$\left((p_D - d)\, q_D - (p_M(f) - f - d)\, q_M - (f-c)\, q_M\right) q_M p_M'(f)\, h(\Theta(f)) = 0.$$

The planner would set f^e such that

$$f^e = c + (p_D - d)\frac{q_D}{q_M} - (p_M(f^e) - d - f^e) = c + \overline{\Delta}_m(f^e), \tag{14}$$

where

$$\overline{\Delta}_m(f^e) = (p_D - d)\frac{q_D}{q_M} - (p_M(f^e) - d - f^e).$$

For the remainder of this section, we assume $\overline{\Delta}_m(c) > 0$, so the markup differential is non-negative when the fee is set equal to the platform's marginal cost, corresponding to the assumption $\Delta_M > 0$ in our benchmark model. We also assume the generalization of (2) holds, which can be written as

$$\int_{\Theta(c)}^{\bar{b}} \left(u_M(c) - u_D + \pi_M(c) - \pi_D + b\right) dH(b) > 0, \tag{15}$$

so the existence of the platform continues to increases total welfare when its fee is set at marginal cost. These additional assumptions ensure a unique solution to (14) exists and characterizes the welfare maximizing fee.

Proposition 10. (Incomplete pass-through and inelastic demand) *Assume (15) holds. When firms have incomplete pass-through on M and possibly different inelastic demand levels across the platform and direct channels, the welfare-maximizing fee is given by (14). In a repeated setting starting from the platform's unregulated fee (or any other fee at which the platform operates) in the first period, and regulating the fee in period $t \geq 2$ to $f^e(t)$ as defined by (14) based on the previous period's levels of the variables on the right-hand-side of (14), then $f^e(t) \to f^e$ as $t \to \infty$.*

The result shows that a similar formula to our baseline efficient fee formula (5) applies. The only difference is that the markup on M now depends on f^e and the markup in the direct channel is scaled by the relative levels of demand in the two channels. Note that when consumers' demand levels in each channel are the same (e.g. the case with unit demands), then the additional scaling factor drops out.

As established in Proposition 10, the formula in (14) can be implemented in practice by requiring the platform to set its fee according to (14) in each period, where f^e in (14) can be calculated using the previous period's values of prices, demands and the fee f. Starting from the platform's unregulated fee f^* (or any other fee at which the platform operates) and repeating this process, as f is adjusted each period according to this rule, the regulated fee moves closer each period to the efficient fee f^e, and eventually converges to f^e.

Incomplete pass-through can help explain why firms may argue in favor of lower fees. With inelastic demand, the firms' total profit is

$$\int_{\Theta(f)}^{\bar{b}} \left(p_M(f) - f - d\right) q_M dH(b) + \int_{\underline{b}}^{\Theta(f)} \left(p_D - d\right) q_D dH(b).$$

The derivative of this with respect to f is

$$\begin{aligned}&-\left(\left(p_M(f) - f - d\right) q_M - \left(p_D - d\right) q_D\right) p_M'(f)\, q_M h\left(\Theta(f)\right)\\ &-\left(1 - p_M'(f)\right) q_M \left(1 - H\left(\Theta(f)\right)\right).\end{aligned}$$

With incomplete pass-through, the term in the second line above is negative. Moreover, $p_M(f) - f$ is decreasing in f under incomplete pass-through, so provided f is not too high, and the measure of consumers using the platform $1 - H(\Theta(f))$ is high, then even if the first term is positive, the term in the second line will dominate, and the firms' total profit will decrease in f. For f high enough, the term in the first line will be positive and will dominate, so that the firms' total profit will eventually increase in f. However, if the firms collectively argue for regulating a very high fee (e.g. one that would lead to few consumers using the platform because their prices would have to be so high), this may raise anticompetitive concerns and/or be blocked due to the harm to consumers. Given this, firms would be better off lobbying for regulators to lower fees.[13]

Finally, we turn to the more complicated case with elastic demand, so that q_M in (14) is now the downward sloping function $q_M(f)$. If platform fee regulation is not intended to force firms that sell on these platforms to lower their prices down to their costs, which would amount to some kind of price regulation on these competing firms rather than regulating the platform's intermediation role, then the characterization in Proposition 10 still holds other than replacing q_M with $q_M(f^e)$. However, to the extent that the planner puts the same weight on (13) as on (12), and so cares about lowering the prices firms set on the platform towards cost, this implies the regulated fee should be lower, implying our fee cap provides an upper bound on the overall welfare maximizing fee.

Proposition 11. (Incomplete pass-through and elastic demand) *Assume (15) holds. When firms have incomplete pass-through on M and possibly different elastic demand levels across the platform and direct channels, the welfare-maximizing fee lies between c and the level f^e defined in (14) after q_M is replaced by $q_M(f^e)$.*

Elastic demand pushes down the welfare-maximizing fee below the efficient fee level determined by (14) as the planner also now takes into account the benefit of lowering firms prices towards costs via a lower platform fee (the price compression effect). However, the welfare-maximizing fee still cannot be below c since then the platform will shut down, and welfare would be lower compared to setting $f = c$. Thus, the welfare-maximizing fee is bounded between c and f^e in (14).

[13] In the Online Appendix we provide a specific model in which there is incomplete pass-through, and provide a condition so that the firms' total profit increases as f is lowered below M's optimal fee.

4.2 Showrooming and comparing regulatory tools

Consider the extension of the *Sequential search model* in Section 2 where a fraction of consumers ρ can showroom. This means after discovering a firm on M, these consumers can costlessly observe the firm's price on each channel and purchase from it in either channel, or continue searching on M. We assume consumers know whether they can showroom or not when deciding which channel to use. Firms cannot distinguish these consumers though, from those coming directly. For this setting, it matters if we interpret b as a joining cost (or benefit) that consumers get from the platform, so even if they switch to buy directly they still incur b, or one they obtain only when making a transaction on the platform. We will focus on the latter interpretation since it has the realistic property that even if everyone who comes to M can showroom (i.e., $\rho = 1$), only some of them will actually showroom (depending on their specific draw of b and the relative prices across the two channels).[14]

When $0 < \rho < 1$, the pricing of competing firms in this search setting is complicated. Even though each firm cannot influence which consumers go to M in the first place, by manipulating their prices they can influence which channel showrooming consumers will complete their transaction on. This means firms' pricing in the two channels is mutually determined. However, their pricing remains straightforward in the two extremes. In the absence of any showrooming (i.e., $\rho = 0$) we know from the Sequential search model in Section 2, the firms' equilibrium prices are $p_M = f + \mu_M$ and $p_D = \mu_D$, where μ_M and μ_D are defined in that example. With all consumers able to showroom (i.e., $\rho = 1$), all consumers would go to M to find their match, but then switch to buy in the direct channel if their draw of b is sufficiently low. Either way, all search will be conducted on M, meaning firms will compete their markups in the direct channel down to the same level as on M (i.e., $p_M = f + \mu_M$ and $p_D = \mu_M$). Provided in the intermediate case (i.e., $0 < \rho < 1$), the symmetric prices solving the first-order conditions of each firm's maximization problem continue to characterize the equilibrium, then as we show in the proof of Proposition 12, equilibrium prices satisfy $f + \mu_M < p_M < f + p_D$ and $\mu_M < p_D < \mu_D$. This implies the equilibrium price on M is higher when only some consumers can showroom (compared to when no one can showroom, or everyone can showroom), while the equilibrium price in the direct channel is lower than the case without any showrooming, but higher than the case with full showrooming.

[14] In the Online Appendix, we show Proposition 12 still holds in case b is a joining benefit.

To characterize the efficient fee in this setting, we make the assumption that

$$\frac{dp_M}{df} > \frac{dp_D}{df}, \tag{16}$$

which just says an increase in M's fee causes firms to increase their price on M more than on the direct channel. This is consistent with firms adjusting relative prices to shift transactions from M to the direct channel in light of a higher fee. This is clearly true in the extreme cases noted above, for which $\frac{dp_M}{df} = 1$ and $\frac{dp_D}{df} = 0$, and it will also be true in general for ρ sufficiently close to zero and ρ sufficiently close to one.[15] Among other things, this ensures the planner's objective function is quasiconcave. We are now ready to characterize the efficient fee in this setting.

Proposition 12. *The welfare-maximizing fee in the presence of showrooming is lower than that in the case without showrooming, and increasingly so as the fraction of consumers that can showroom ρ increases. Specifically, the efficient fee satisfies $p_M(f^e) - p_D(f^e) = c$, and so can still be written in the form $f^e = c + \tilde{\Delta}_m$, with $\tilde{\Delta}_m = p_D(f^e) - (p_M(f^e) - f^e)$. Here f^e is decreasing in ρ, with $f^e = c + \Delta_m$ when $\rho = 0$ and $f^e = c$ when $\rho = 1$.*

The result is intuitive. Firms take into account that some fraction of consumers can be attracted to switch after searching on M. Since these consumers have low search costs, they search more intensely than the consumers who come directly, leading firms to lower their direct price compared to the case without showrooming. This reduces the markup differential between the two channels, and so reduces the need to set a fee above cost to offset that markup differential. In the limit, with everyone able to showroom, the markup differential is eliminated, and therefore the efficient fee, which is still equal to the platform's marginal cost plus the markup differential, just equals the platform's marginal cost.

Taking into account the two types of consumers, we show in the proof of Proposition 12 that welfare at the efficient fee is

$$\begin{aligned} W(f^e) &= (1-\rho)\left(\int_{c-\Delta_s}^{\overline{b}} (\phi_M + b - c)dH(b) + \int_{\underline{b}}^{c-\Delta_s} \phi_D dH(b)\right) \\ &\quad + \rho\left(\int_{c}^{\overline{b}} (\phi_M + b - c)dH(b) + \int_{\underline{b}}^{c} \phi_M dH(b)\right). \end{aligned} \tag{17}$$

[15] These limit results are established in the proof of Proposition 12. We also characterize the condition for (16) to hold for any $0 < \rho < 1$, and establish a sufficient condition for this is that the density function h is log concave in its argument, and μ_M is sufficiently close to zero.

Note that for consumers who can showroom, they get the same high match value regardless of which channel they end up making their purchase on. Showrooming enhances welfare by allowing consumers to combine the more efficient search technology of the platform with their preferred channel to make transactions on. This intuition also applies under efficient fee regulation. Clearly, $W(f^e)$ is increasing in ρ, so under fee regulation, welfare is higher when more consumers can showroom.

But what about enabling showrooming as an alternative to fee regulation? For example, policymakers could make platforms remove rules which stop suppliers informing their customers on the platform of their own direct channels. Removing these rules makes showrooming easier. An example is Apple's and Google's imposition of anti-steering rules which prevent most app developers from linking consumers from inside their apps on these platforms to their direct channels to make purchases. These rules have recently been challenged. The judge in the 2021 Epic Games vs. Apple case in the U.S. required Apple to remove its anti-steering rules in the U.S.[16] In 2024 the European Commission fined Apple about two billion dollars for preventing Spotify and other developers from informing users of options to purchase outside its App Store.[17] Furthermore, in response to the DMA, in 2024 both Apple and Google have relaxed their rules in Europe.[18] Allowing suppliers to direct their customers to their own channels would put pressure on platforms to lower their fees to some extent, to avoid too much showrooming. Thus, we can compare fee regulation with a policy that ensures all consumers are able to showroom if they want to.

Formally, this involves comparing (17) with the welfare arising when $\rho = 1$ but the platform's fee is left unregulated. In the latter case, with showrooming fully enabled, welfare is

$$W_s = \int_{f^*}^{\overline{b}} (\phi_M + b - c) dH(b) + \int_{\underline{b}}^{f^*} \phi_M dH(b), \tag{18}$$

[16] https://techcrunch.com/2024/01/17/apple-allows-devs-to-promote-subscriptions-on-the-web-with-a-27-cut/

[17] https://www.wsj.com/tech/apple-hit-with-near-2-billion-fine-in-europe-over-music-streaming-apps-74062ff7

[18] According to Apple (https://developer.apple.com/support/apps-using-alternative-payment-providers-in-the-eu/,) developers will be able to "(d)irect users to complete a transaction for digital goods and services on (their) external webpage. The presentation of the link out may communicate information for EU users about promotions, discounts, and other deals." According to Google (https://blog.google/around-the-globe/google-europe/complying-with-the-digital-markets-act/) "Starting on March 6, we are launching a program that allows developers of Play-distributed apps to directly lead users in the EEA outside the app, including to promote offers."

where $f^* = c + \lambda(f^*)$. This reflects that when all consumers can showroom, all will go through M to search. This drives the markup differential and the effective surplus differential to zero, with $p_D = \mu_M$. Then consumers make their transaction on M iff $b > f$. Taking this into account, the platform sets f to maximize its profit $(f - c)(1 - H(f))$, which gives rise to the fee f^* and the welfare expression in (18).

Since $f^* > c > c - \Delta_s$, we have

$$W_s - W(f^e) = (1-\rho)\left(\int_{\underline{b}}^{c-\Delta_s} \Delta_s dH(b) + \int_{c-\Delta_s}^{c} (c-b)\, dH(b)\right) - \int_c^{f^*} (b-c) dH(b).$$

Clearly this is decreasing in ρ and is negative for ρ sufficiently close to one. Thus, it follows that:

Proposition 13. *Regulating the platform's fee at the efficient fee level increases total welfare more than a policy that ensures all consumers are free to showroom provided that enough consumers are free to showroom before the policy is enacted. Combining both policies (efficient fee regulation and enabling showrooming) always leads to higher welfare than either policy alone.*

Without ρ sufficiently close to one, the sign of $W_s - W(f^e)$ will be ambiguous. This reflects a tradeoff between three effects of a policy enabling all consumers to showroom rather than regulating the fee at the efficient level: (i) for the additional $1-\rho$ consumers that can now showroom, the policy creates additional surplus of Δ_s as consumers can utilize the better search technology on M; (ii) for the additional $1-\rho$ consumers that can now showroom, the policy creates additional surplus of $c - b$ on the transactions consumers were previously making on M since these transactions should have occurred directly with the same firm given $b < c$; (iii) for all consumers, the policy causes a loss of $b - c$ on the transactions where $c < b < f^*$ which are inefficiently shifted to the direct channel as firms try to avoid the platform's monopoly fee. Thus, in general, it may not be clear which individual policy is better. On the other hand, by combining both policies to enable showrooming and fee regulation, the planner can remove the distortion in (iii), and only the two positive effects in (i) and (ii) would remain.

4.3 Platform competition

In this section we explore whether platform competition reduces the need for platform fee regulation. To do so, suppose there are two competing platforms M_1

and M_2, as well as a direct channel. A consumer's utility of shopping directly is still $\phi_D - p_D = \phi_D - \mu_D$. Provided firms can price discriminate across channels, firms will join both platforms, which makes them at least weakly better off. A consumer's utility of using M_1 is

$$\phi_M - p_{M_1} + b + t\nu_1$$

and her utility of using M_2 is

$$\phi_M - p_{M_2} + b + t\nu_2,$$

where $\nu_1 \geq 0$ and $\nu_2 \geq 0$ are platform-specific benefits. These benefits are independently drawn for each consumer from a distribution function with mean $E[\nu]$. The difference $\Delta\nu \equiv \nu_1 - \nu_2$ is assumed to be symmetrically distributed around zero and therefore $E[\Delta\nu] = 0$. Let $G_{\Delta\nu}$ be the distribution function for $\Delta\nu$ and assume it has a weakly increasing hazard rate with corresponding density function $g_{\Delta\nu}$. The parameter t is akin to the standard mismatch parameter in the Hotelling model. To simplify the analysis, we assume that b is sunk once consumers have decided to shop using platforms. For tractability, a consumer's decision is determined by a two-stage process in stage 2 of the original game:

Stage 2a: **Consumers choose between either the direct channel or the platform channel**. Consumers do not observe ν_1 and ν_2 at this stage. They draw b, observe f_1 and f_2 (or the modal prices on M_1, M_2, and the direct channel), and decide whether to shop directly or use a platform.

Stage 2b: **Consumers choose between M_1 and M_2**. Consumers who chose the platform channel in stage 1 observe the realizations of ν_1 and ν_2 and decide which platform to use.

Similar to our baseline model, consumers only learn firms' channel and price choices after deciding which channel they will visit. Thus, consumers' channel choices are unaffected by firms' actual choices of which channels to join and what prices to set.

In the Appendix we prove the following result.

Proposition 14. *The equilibrium fee under platform competition f^* satisfies $f^* < \hat{f}$, where*

$$\hat{f} = c + \frac{t}{2g_{\Delta\nu}(0)}, \tag{19}$$

is the standard competitive fee in the pseudo-equilibrium where consumers' stage-2a choice about whether to shop directly or on platforms is ignored by platforms when setting fees.

The proposition shows that in our setting (where consumers choose whether to use the direct channel or not, and if not, choose between the two platforms), platform competition leads to a lower fee than the one arising from platform competition in the absence of a direct channel. This is natural since here the direct channel acts like an additional competitive constraint on the fee that platforms would want to charge to attract consumers to their platform in the first place. The upper-bound fee in (19) is low if the mismatch cost of not using the ideal platform is low (i.e., t is low) or many consumers do not have a strong preference towards one of the platforms over the other (i.e., $g_{\Delta\nu}(0)$ is large).

In this framework, the efficient fee remains the same as in the monopoly case as it only depends on c and Δ_m, which are unchanged by platform competition.[19] Therefore, to determine if putting a cap on fees could improve welfare, we need to compare f^* with $f^e = c + \Delta_m$. Given $f^* < \hat{f}$, a sufficient condition for a fee cap to be unnecessary is that $f^e = c + \Delta_m > c + \frac{t}{2g_{\Delta\nu}(0)} = \hat{f}$, which is equivalent to $\Delta_m > \frac{t}{2g_{\Delta\nu}(0)}$. Clearly, this inequality will be true for intense enough platform competition (so low enough t or high enough $g_{\Delta\nu}(0)$), meaning with sufficient platform competition, a fee cap would be unnecessary. If platform competition is not intense enough, our existing fee cap formula may still be needed to bring down the platforms' fees to the efficient level.

5 Conclusion

This paper proposes a simple yet flexible framework for studying the regulation of the fee a platform charges to suppliers when transactions can be done either through the platform or between suppliers and consumers directly. Taking into account that suppliers have lower markups on the platform than in the direct channel due to intensified competition on the platform, we find the efficient fee exceeds the platform's marginal cost by the difference in markups across the two channels. This can reduce the fee set by a monopoly platform which tries to extract too much of the surplus it creates, while eliminating the otherwise excessive use of the platform by consumers if the fee were instead set at the platform's marginal cost.

[19] In the welfare expressions, however, ϕ_M needs to be replaced by $\phi_M + tE[\max\{\nu_1, \nu_2\}]$.

The simple characterization of the efficient fee we offer is relatively easy to implement and robust to some obvious extensions: for example, it is robust to partial pass-through of fee changes by suppliers, or when some consumers can showroom, switching to complete their transaction on the direct channel after searching on the platform. In other cases, it provides an upper bound on the welfare-maximizing fee, such as when consumer demand is elastic or when the welfare objective puts less weight on the platform's profit and/or more weight on the surplus of consumers. Thus, it still serves as a useful fee cap in such cases.

It is important to be cognizant that capping platform fees could lead the platform to adjust its business model in response. One risk with regulating the platform fee too low in a marketplace context is that the platform may then decide to foreclose third-party suppliers by selling its own versions of products, and either steer consumers to its own versions (self-preferencing) or closing down its marketplace to third-party suppliers altogether. The shift towards being a reseller could create new welfare losses: the platform may incur higher costs, provide inferior products, or not be able to offer as much choice compared to a market of third-party suppliers. The regulator may want to avoid driving fees too low with this in mind. However, realistically, as long as any fee regulation still allows the platform to operate profitably, it is unlikely that the platform would aggressively restrict suppliers' access to its platform as that could open the opportunity for a rival platform to emerge offering a marketplace to cater to such suppliers.

A more likely risk therefore is that the platform will try to increase other fees instead. In a marketplace context, the platform could increase fees suppliers pay to be promoted or discovered on the marketplace, or fees for suppliers to be listed. While these alternative types of fees may not always be passed through to consumers and so do not necessarily raise the same issue of distorting consumer choices between channels as the fees we focused on, they may nonetheless represent inefficient ways for the platform to extract surplus. To avoid unintended consequences from an inefficient change in business model, the fee regulation might therefore need to apply as a global cap on the average fee per unit (or per dollar) of transactions, so any attempt to recover fee revenues in other ways would not benefit the platform.

In contrast to Gomes and Mantovani (2024), we have focused on the case where price parity does not hold, so suppliers are free to set different prices across different channels. If suppliers are actually restricted to set the same prices across different channels, our formula for the efficient fee does not apply. The efficient fee in this case should instead be determined by a framework in which the price parity is taken into account directly, such as in Gomes and Mantovani. It remains, however, to explore the possibility of partial price parity, where some transactions are subject to price parity and others are not, and whether combining their cap with ours would be useful in such a setting.

Appendix.

Proof of Proposition 1. The platform's profit is $\Pi(f) = (f - c)(1 - H(f - \Delta_s - \Delta_m))$. Note that the left-hand-side (LHS) of (3) strictly increases from c to $\bar{b} + \Delta_s + \Delta_m$ when f increases from c to $\bar{b} + \Delta_s + \Delta_m$, where note $\bar{b} \to \infty$ if λ is constant. The right-hand-side (RHS) of (3) weakly decreases in f given λ is weakly decreasing (from the assumed weakly increasing hazard rate). Moreover, the RHS of (3) either decreases from a value greater than c to c when f increases from c to $\bar{b} + \Delta_s + \Delta_m$ or is a positive constant greater than c. As a result there is a unique solution to (3). Denote this unique solution as f^*. The function $\Pi(f)$ is strictly increasing in f for $c \leq f < f^*$ as the LHS of (3) is strictly smaller than the RHS, and is strictly decreasing in f for $\bar{b} + \Delta_s + \Delta_m \geq f > f^*$ as the LHS of (3) is strictly greater than the RHS. This implies that f^* is indeed the global maximum. ■

Proof of Proposition 5. Define W^{tus} as W^w with $w_m = 0$ and $w_c = w_f = w$. Differentiating W^{tus} with respect to f we get

$$\frac{dW^{tus}}{df} = w\Delta_m h(f - (\Delta_s + \Delta_m)) - w(1 - H(f - (\Delta_s + \Delta_m))).$$

Define f^{tus} as the unique value of f solving

$$\lambda(f - (\Delta_s + \Delta_m)) = \Delta_m.$$

We have $\frac{dW^{tus}}{df} = h(f - (\Delta_s + \Delta_m))\, w\,[\Delta_m - \lambda(f - (\Delta_s + \Delta_m))]$. Recall λ is weakly decreasing in f. Assume first that it is not constant. If $f < f^{tus}$, we must have $\lambda(f - (\Delta_s + \Delta_m)) \geq \Delta_m$ and $\frac{dW^{tus}}{df} \leq 0$, and vice-versa when $f > f^{tus}$. So f^{tus} characterizes a minimum. The planner will either want to set f as low as possible subject to M wanting to operate (i.e. $f = c$) or as high as possible so that no consumer would go to M (or equivalently, $f < c$, so M would not operate). In either case, M's profit is zero, so W^{tus} corresponds to W. The condition (2) then implies $f^{tus} = c$. Alternatively, suppose λ is constant. Then W^{tus} is either always increasing or always decreasing, in which case from the logic above W^{tus} is maximized by setting $f^{tus} = c$, or W^{tus} is constant, in which case the planner can't do better than setting $f^{tus} = c$. ■

Proof of Proposition 6. First note that if $w_c = w_f$ and $w_c = w_m$, then $f^w = f^e$ as shown in Proposition 4, so in what follows we assume either $w_c > w_f$ or $w_c > w_m$ or both. Let us first show $f^w \geq c$. Suppose instead W^w is maximized for some $f < c$. In this case M stops operating, so W^w just consists of the resulting weighted combination of consumer surplus and firms' profits without M. Suppose we set the same weight on consumer surplus as firms' profit, so set $w'_c = w_f$. Then W^w corresponds to total user surplus. We know from Propositions 2 and 5 that consumer surplus and total user surplus are both maximized when $f = c$, with consumer surplus strictly decreasing in f for $f > c$. Therefore increasing the weight on consumer surplus from $w'_c = w_f$ up to the actual (weakly) higher w_c will still leave the fee maximizing W^w equal to c. Thus, W^w must be strictly higher at $f = c$ than for any fee $f < c$, contradicting that $f < c$ maximizes W^w.

We next show $f^w \leq f^e$. Suppressing the argument $f - (\Delta_s + \Delta_m)$ in h and H in (4), $\frac{dW^w}{df}$ can be rewritten as

$$\frac{dW^w}{df} = \left(\frac{w_f}{w_m}\Delta_m + c - f - \left(\frac{w_c}{w_m} - 1\right)\lambda\right) hw_m.$$

If H is GPD, λ is linear in f, and the sign of $\frac{dW^w}{df}$ can change at most once. We next show that $\frac{dW^w}{df} \leq 0$ at $f = f^e$ and therefore W^w cannot be strictly increasing. This requires

$$\left(\frac{w_c}{w_m} - 1\right)\lambda(c - \Delta_s) \geq \left(\frac{w_f}{w_m} - 1\right)\Delta_m. \tag{20}$$

If $w_f \leq w_m$, the RHS of (20) is non-positive. Given the LHS of (20) is non-negative, (20) always holds. Suppose instead $w_f > w_m$. From Proposition 8, given our assumptions that (2) holds and H takes the GPD form, we must have $f^* > f^e$. This implies

$$\lambda(c - \Delta_s) > \Delta_m,$$

which together with the assumption that $w_c \geq w_f$, implies (20) holds.

There are three remaining possibilities for the shape of W^w given the sign of $\frac{dW^w}{df}$ changes at most once. If W^w is decreasing for $f \geq c$, then $f^w = c < f^e$. If W^w is hump-shaped, $f^w \leq f^e$ given that $\frac{dW^w}{df} \leq 0$ at $f = f^e$. If W^w is U-shaped for $f \geq c$, then the maximum either occurs at $f = c$ or for the highest possible f at which there are no transactions. But since W^w with no transactions is equivalent to W^w with $f < c$ (since M would then shut down) which we have already ruled out as maximizing W^w, W^w must be maximized with $f^w = c$ in this case. ■

Proof of Proposition 7. Compare f^* in (3) and f^e in (5). Notice that the term $\lambda(f - \Delta_s - \Delta_m)$ in (3) weakly decreases in f unless $H(b)$ is an exponential distribution in which case λ is a constant. We show $f^e \leq f^*$ if and only if $\lambda(c - \Delta_s) \geq \Delta_m$. First suppose $f^e \leq f^*$. Then,

$$\begin{aligned} & c + \lambda(f^e - \Delta_s - \Delta_m) \geq c + \lambda(f^* - \Delta_s - \Delta_m) = f^* \geq f^e = c + \Delta_m \\ \Rightarrow\ & \lambda(c - \Delta_s) = \lambda(f^e - \Delta_s - \Delta_m) \geq \Delta_m. \end{aligned}$$

Next suppose $\lambda(c - \Delta_s) \geq \Delta_m$. We want to show $f^e \leq f^*$ but suppose to the contrary that $f^e > f^*$. Then we have

$$f^e = c + \Delta_m \leq c + \lambda(c - \Delta_s) = c + \lambda(f^e - \Delta_s - \Delta_m) \leq \lambda(f^* - \Delta_s - \Delta_m) = f^*,$$

which leads to a contradiction. Thus, it must be that $f^e \leq f^*$, which completes the proof. ■

Proof of Proposition 8. When H takes the GPD form, the total welfare generated by M when $f = c$ is

$$\int_{c-(\Delta_s+\Delta_m)}^{\overline{b}} (\Delta_s + b - c) dH(b) = \left(\frac{\epsilon}{\sigma}\left(\overline{b} + \Delta_s + \Delta_m - c\right)\right)^{\frac{1}{\epsilon}} \left(\frac{\sigma + \epsilon\left(\underline{b} + \Delta_s - c\right) - \Delta_m}{1 + \epsilon}\right).$$

Using (1), the first term in large brackets is positive. Thus, the whole expression is positive iff $\sigma + \epsilon\left(\underline{b} + \Delta_s - c\right) > \Delta_m$. Since $\lambda\left(c - \Delta_s\right) = \sigma + \epsilon\left(\underline{b} + \Delta_s - c\right)$ (see Example 1), this is equivalent to $\lambda\left(c - \Delta_s\right) > \Delta_m$ which from Proposition 7 is the same condition for $f^* > f^e$. ■

Proof of Proposition 10. Assuming $q_M(f)$ is fixed, the derivative of (12) with respect to f is

$$\begin{aligned} & -\left(p_M'(f)\, q_M\right)^2 h(\Theta(f)) \\ & + \left(\pi_D - \pi_M(f) - (f - c) q_M\right) q_M \frac{\partial}{\partial f}\left(p_M'(f)\, h\left(\Theta(f)\right)\right). \end{aligned} \tag{21}$$

At the FOC defined by setting (12) to zero, the second line in (21) is zero, and (21) is therefore clearly negative at the f solving the FOC. So the solution proposed in (14) represents a local maximum. It is a global maximum if (14) has a unique solution $f \in \left[c, \Theta^{-1}\left(\bar{b}\right)\right]$, where $\Theta^{-1}\left(\bar{b}\right)$ is the fee at which $\Theta\left(f\right) = \bar{b}$ so no more consumers come to M. Note $\overline{\Delta}_m\left(f\right)$ is decreasing in f given $p_M\left(f\right)$ is increasing in f, so any solution to (14) must be unique. Thus, a solution to (14) is indeed the global maximum provided it exists. To see why the solution exists, notice that f^e must be greater than c given $\overline{\Delta}_m(c) > 0$. Also, f^e must be lower than $\Theta^{-1}\left(\bar{b}\right)$ since the welfare when M operates at $f = c$ is strictly higher than when $f = \Theta^{-1}\left(\bar{b}\right)$ and M effectively shuts down. Given the welfare function is continuous on $[c, \Theta^{-1}\left(\bar{b}\right)]$, the solution to the welfare maximization problem must be interior and is defined by (14).

To prove the second part of the proposition, let f_0 denote the initial fee before regulation, which if it is unregulated should equal f^*, but we can allow it to be any other feasible level at which M operates. Then applying the regulation in period 1 would imply $f^1 = c + \overline{\Delta}_m\left(f^0\right)$. Given there is some positive pass-through, $\overline{\Delta}_m\left(f\right)$ is increasing in f. Suppose first that $f^0 > f^e$. This implies $\overline{\Delta}_m\left(f^0\right) > \overline{\Delta}_m\left(f^e\right)$ and so $f^1 = c + \overline{\Delta}_m\left(f^0\right) > c + \overline{\Delta}_m\left(f^e\right) = f^e$. Given $f^e = c + \overline{\Delta}_m\left(f^e\right)$ and using the definition of $\overline{\Delta}_m\left(f\right)$ above, we have $f^1 - f^e = \left(c + \overline{\Delta}_m\left(f^0\right)\right) - \left(c + \overline{\Delta}_m\left(f^e\right)\right) = f^0 - f^e - \left(p_M\left(f^0\right) - p_M\left(f^e\right)\right) < f^0 - f^e$ since $p_M\left(f^0\right) > p_M\left(f^e\right)$, thus implying $f^1 < f^0$. This implies $f^1 \in \left(f^e, f^0\right)$. Alternatively if $f^0 < f^e$, by a parallel argument, using $f^1 = c + \overline{\Delta}_m\left(f^0\right)$ implies $f^1 \in \left(f^0, f^e\right)$. Iterating over time, if $f^{t-1} < f^e$, $f^t \in \left(f^{t-1}, f^e\right)$, while if $f^{t-1} > f^e$, $f^t \in \left(f^e, f^{t-1}\right)$. Since the pass-through rate is less than one, we can apply the standard contraction mapping theorem to show the sequence on $[f^e, f^0]$ must converge to a unique fixed point, which by definition must be f^e, and similarly for the sequence on $[f^0, f^e]$. ■

Proof of Proposition 11. In addition to the term (12), we have the extra term (13) when taking the derivative of W in (11). Note that $q_M'(f) < 0$, and $p_M(f) \geq c + d$ as $p_M(f) \geq f + d$ for any $f \geq c$. Thus, the term (13) is negative. Proposition 10 shows that the fee level in (14) is determined by setting (12) to zero, and that this represents a global maximum. Now, with the extra term in the derivative of W that is always negative, the welfare maximizing level of f must be below the level defined by (14). Moreover, given our assumption that the existence of M increases total welfare when its fee is set at marginal cost, shutting down M by setting $f < c$ lowers total welfare compared to setting $f = c$. ■

Proof of Proposition 12. Let the fraction ρ of consumers who can showroom be called showrooming consumers, and the fraction $1 - \rho$ of consumers who cannot be called regular consumers. Given the benefit b is obtained from making a transaction, all showrooming consumers first visit M and then decide which channel to purchase on.

We start by determining the demand function of an individual firm i that sets p_D^i and p_M^i. For regular consumers who visit M, they choose to buy from i if $v^i - p_M^i \geq x_M - p_M$. For regular consumers who search directly, they buy from firm i if $v^i - p_D^i \geq x_D - p_D$. For showrooming consumers, their value of not buying from i but continuing to search on M is $x_M - \min\{p_M - b, p_D\}$. Thus, they buy from firm i on M if $p_M^i - b \leq p_D^i$ and

$$v^i \geq p_M^i - b + x_M - \min\{p_M - b, p_D\},$$

buy from firm i directly if $p_M^i - b > p_D^i$ and

$$v^i \geq p_D^i + x_M - \min\{p_M - b, p_D\},$$

and continue to search on M otherwise. Firm i's demand has four parts: (i) demand from regular consumers who buy on M $D_{rM} \equiv (1-\rho)(1 - H(f - \Delta_s - \widetilde{\Delta}_m))\frac{1-G(x_M - p_M + p_M^i)}{1-G(x_M)}$, where $\widetilde{\Delta}_m \equiv p_D - (p_M - f)$; (ii) demand from regular consumers who buy directly $D_{rD} \equiv (1-\rho)H(f - \Delta_s - \widetilde{\Delta}_m)\frac{1-G(x_D - p_D + p_D^i)}{1-G(x_D)}$; (iii) demand from showrooming consumers who buy from i on M

$$D_{sM} \equiv \rho \int_{p_M^i - p_D^i}^{\bar{b}} \left(\frac{1 - G(p_M^i - b + x_M - \min\{p_M - b, p_D\})}{1 - G(x_M)} \right) dH(b);$$

and (iv) demand from showrooming consumers who switch and buy from i

$$D_{sD} \equiv \rho \int_{\underline{b}}^{p_M^i - p_D^i} \left(\frac{1 - G(p_D^i + x_M - \min\{p_M - b, p_D\})}{1 - G(x_M)} \right) dH(b).$$

Firm i solves

$$\max_{p_D^i, p_M^i} \; p_D^i(D_{rD} + D_{sD}) + (p_M^i - f)(D_{rM} + D_{sM})$$

Focusing on a solution with symmetric prices, the FOC with respect to p_M^i can be written

$$
\begin{aligned}
& ((1-\rho)(1-H(p_M-p_D-\Delta_s))+\rho\,(1-H(p_M-p_D)))\left(\frac{p_M-f}{\mu_M}-1\right) \quad (22)\\
= \; & -\rho(p_M-f-p_D)h(p_M-p_D),
\end{aligned}
$$

and the FOC with respect to p_D^i can be written

$$
\begin{aligned}
& (1-\rho)H(p_M-p_D-\Delta_s)\left(1-\frac{p_D}{\mu_D}\right)+\rho H(p_M-p_D)\left(1-\frac{p_D}{\mu_M}\right) \quad (23)\\
= \; & -\rho(p_M-f-p_D)h(p_M-p_D).
\end{aligned}
$$

As noted in Section 4.2, we assume the solutions to these FOCs characterize the equilibrium (symmetric) prices p_M and p_D.

We first derive some inequalities on each of the equilibrium prices. Suppose $p_M \geq f+p_D$ so the RHS of (22) is non-positive. For the LHS of (22) to also be non-positive requires $p_M \leq f+\mu_M$. Since the RHS of (23) is also non-positive, for the LHS of (22) to also be non-positive requires $p_D > \mu_M$. But $p_M \geq f+p_D$ and $p_D > \mu_M$ imply $p_M > f+\mu_M$, contradicting that $p_M \leq f+\mu_M$. Thus, we must have (i) $p_M < f+p_D$. Then (22) implies we must have (ii) $p_M > f+\mu_M$, and (23) implies we must have (iii) $p_D < \mu_D$. Finally, if $p_D < \mu_M$, then since we have $p_M < f+p_D$ from (i), this implies $p_M < f+\mu_M$ which contradicts (ii). Thus, we must have (iv) $p_D > \mu_M$. Thus, (i)-(iv) establishes the properties noted in the text of Section 4.2.

Now consider welfare. Taking into account the two types of consumers, welfare is

$$
\begin{aligned}
W \; = \; & (1-\rho)\left(\int_{p_M-p_D-\Delta_s}^{\bar{b}}(\phi_M+b-c)dH(b)+\int_{\underline{b}}^{p_M-p_D-\Delta_s}\phi_D dH(b)\right) \quad (24)\\
& +\rho\left(\int_{p_M-p_D}^{\bar{b}}(\phi_M+b-c)dH(b)+\int_{\underline{b}}^{p_M-p_D}\phi_M dH(b)\right).
\end{aligned}
$$

Differentiating with respect to f and noting that both p_M and p_D can vary with f, we get the FOC for the efficient fee f^e:

$$
(p_M-p_D-c)\left((1-\rho)\,h(p_M-p_D-\Delta_s)+\rho h(p_M-p_D)\right)\left(\frac{dp_D}{df}-\frac{dp_M}{df}\right)=0.
$$

Given our assumption that $\frac{dp_M}{df} > \frac{dp_D}{df}$, there is a unique solution to the FOC given by $p_M(f^e) = c + p_D(f^e)$, or equivalently, $f^e = c + \tilde{\Delta}_m$, with $\tilde{\Delta}_m = p_D(f^e) - (p_M(f^e) - f^e)$. Note the condition $\frac{dp_M}{df} > \frac{dp_D}{df}$ implies for any lower f, $\frac{dW}{df} > 0$, and for any higher f, $\frac{dW}{df} < 0$, so f^e is indeed the welfare maximizing fee.

Next we show conditions under which $\frac{dp_M}{df} > \frac{dp_D}{df}$ holds. To do so, define $x \equiv p_M - p_D$. After substituting in x for $p_M - p_D$ in (22) using $p_M = x + p_D$, we can solve (22) for p_D. Substituting this p_D into (23) and rearranging implies x is determined by

$$(1-\rho)\left(\frac{\mu_D - \mu_M}{\mu_D}\right) + (x - f)\, y(x) = 0, \tag{25}$$

where

$$\begin{aligned} y(x) \quad = \quad & \frac{\rho h(x)}{H(x-\Delta_s)} + \left((1-\rho)\frac{\mu_M}{\mu_D} + \rho\frac{H(x)}{H(x-\Delta_s)}\right) \times \\ & \left(\frac{\rho h(x)}{(1-\rho)(1-H(x-\Delta_s)) + \rho\,(1-H(x))} + \frac{1}{\mu_M}\right). \end{aligned}$$

Totally differentiating (25), we get

$$\frac{dx}{df} = \frac{y(x)^2}{y(x)^2 - (1-\rho)\left(\frac{\mu_D - \mu_M}{\mu_D}\right) y'(x)}.$$

Note as $\rho \to 0$, we get $y'(x) \to 0$ and so $\frac{dx}{df} \to 1$, and as $\rho \to 1$, $\frac{dx}{df} \to 1$, which show the limit results noted in the text of Section 4.2 hold. More generally, given that $x - f < 0$, it follows that $\frac{dx}{df} > 0$ for $0 < \rho < 1$ provided $y'(x)$ is not too positive. This requirement is satisfied if μ_M is sufficiently close to zero since then the sign of $y'(x)$ is determined by the sign of $\left[\frac{H(x)}{H(x-\Delta_s)}\right]'$, which is negative given h is log concave, thus establishing the result in footnote 14.

To show a higher ρ decreases f^e, note given $H(x) > H(x - \Delta_s)$ and $\mu_D > \mu_M$, $y(x)$ will be increasing in ρ. Since $y(x)$ is multiplied by $x - f$ in (25), which is negative, this together with the fact that the first term in (25) is decreasing in ρ implies that the LHS of (25) would become negative if $x - f$ didn't change in response to an increase in ρ. Thus, if f^e didn't change in response to a higher ρ, we would need an increase in x for (25) to hold. Since $x = c$ at the initial f^e, this implies $x > c$ after the increase in ρ. Thus, to decrease x so that $x = c$, which is required for welfare maximization, f must decrease given $\frac{dx}{df} > 0$. This proves that f^e is decreasing in ρ.

Finally, substituting $p_M(f^e) - p_D(f^e) = c$ into (24) implies (17) as required. ■

Proof of Proposition 14.

In stage 2a, a consumer chooses the platform channel over the direct channel iff

$$\phi_M + b + \mathbb{E}[\max\{tv_1 - p_M(f_1), tv_2 - p_M(f_2)\}] \geq \phi_D - \mu_D,$$

where $p_M(f) = f + \lambda(x_M)$. Then, we have

$$\Delta_s + \Delta_m + b + \mathbb{E}[\max\{tv_1 - f_1, tv_2 - f_2\}] \geq 0. \tag{26}$$

In stage 2b, ν_1 and ν_2 are realized. Note that from (26), for a consumer who chooses to compare platforms rather than shopping directly, the ex-ante expected utility of shopping on a platform must be positive since $\phi_D - \mu_D \geq 0$. However, the realized ν_1 and ν_2 might be lower than $\mathbb{E}[\nu]$. But the consumer will still choose a platform to make purchases as b is assumed to be sunk. In stage 2b, the consumer's equilibrium utility $\phi_M - p_M^* + t\nu_i$ will be positive if $\phi_M - p_M^* \geq 0$. This will be true if ϕ_M is sufficiently large. In stage 2b, a consumer prefers M_1 to M_2 iff

$$\phi_M - p_{M1} + t\nu_1 \geq \phi_M - p_{M2} + t\nu_2 \quad \Leftrightarrow \quad \Delta\nu \geq \frac{f_1 - f_2}{t}.$$

Let us characterize the symmetric equilibrium fee $f^* = f_1^* = f_2^*$. First, consider the pseudo-equilibrium where consumers' stage-2a choice about shopping channels are ignored. Suppose the other platform sets the symmetric pseudo equilibrium fee $\hat{f}$. Then M_i solves

$$\max_{f_i} \left\{ (f_i - c)\left(1 - G_{\Delta\nu}\left(\frac{f_i - \hat{f}}{t}\right)\right)\right\}.$$

Taking the logarithm of the objective function and applying the FOC, we obtain

$$\frac{1}{f_i - c} = \frac{\frac{1}{t} g_{\Delta\nu}(\frac{f_i - \hat{f}}{t})}{1 - G_{\Delta\nu}(\frac{f_i - \hat{f}}{t})}. \tag{27}$$

Setting $f_i = \hat{f}$ implies (19) in Proposition 14. Here, we used the fact $G_{\Delta\nu}(0) = 1/2$. Notice that the LHS is strictly decreasing in f_i from infinity to some finite value as f_i increases from c, while since $G_{\Delta\nu}(\cdot)$ has an increasing hazard rate, the RHS is increasing in f_i from some finite value to infinity as f_i increases to the point where $G_{\Delta\nu} = 1$. Thus, (27) uniquely defines the pseudo-equilibrium fee.

We now check whether the pseudo equilibrium fee continues to be the equilibrium fee in the full model. Note that if M_j sets $f_j = \hat{f}$ defined in (19), M_i has no incentive to set $f_i > \hat{f}$ since its loss in profit from setting $f_i > \hat{f}$ is even greater than in the analysis when stage 1 consumer choices are ignored (as a higher f_i will decrease the measure of consumers who come to either platform). However, when M_i sets $f_i \leq \hat{f}$, its actual profit is

$$(f_i - c)\left(1 - G_{\Delta\nu}\left(\frac{f_i - \hat{f}}{t}\right)\right)\left(1 - H(-\mathbb{E}[\max\{t\nu_1 - f_1, t\nu_2 - \hat{f}\}] - \Delta_s - \Delta_m)\right).$$

Let $y(f_1) \equiv -\mathbb{E}[\max\{t\nu_1 - f_1, t\nu_2 - \hat{f}\}] - \Delta_s - \Delta_m$. Note that $y'(f_1) > 0$. An increase in f_1 leads to a lower expected value of $\max\{t\nu_1 - f_1, t\nu_2 - \hat{f}\}$ as the resulting distribution will be first-order stochastically dominated by the initial distribution. Take the logarithm of M_i's profit function above and apply the FOC

$$\frac{1}{f_i - c} - \frac{\frac{1}{t} g_{\Delta\nu}(\frac{f_i - \hat{f}}{t})}{1 - G_{\Delta\nu}(\frac{f_i - \hat{f}}{t})} = \frac{y'(f_i)}{\lambda(y(f_i))}, \tag{28}$$

where recall λ is the inverse hazard rate which is assumed weakly decreasing in its argument. Given that both $1 - G_{\Delta\nu}(\cdot)$ and $1 - H(\cdot)$ are log-concave, the product of the two remains log-concave and the FOC in (28) identifies the maximizer. Compare (28) with (27). The LHS of (28) is zero at $f_i = \hat{f}$ but the RHS is positive for any f_i including $f_1 = \hat{f}$. Moreover, the LHS is decreasing in f_i, since the hazard rate of $G_{\Delta v}$ is assumed weakly increasing. Thus, the only way for the two sides to be equal is if $f_i < \hat{f}$ so the LHS can also be positive. Note the LHS decreases from ∞ to 0 when f_1 increases from c to $\hat{f}$, while the RHS is always positive. This ensures that there exist a $f_i < \hat{f}$ such that the LHS is equal to the RHS. Since f_1 and f_2 are strategic complements, we must have the symmetric equilibrium fee $f^* < \hat{f}$. ∎

References

Anderson, S. P. and Renault, R. (1999) "Pricing, product diversity, and search costs: A Bertrand-Chamberlin-Diamond model," *RAND Journal of Economics*, Vol. 30(4), pp. 719-735.

Baker, J. and Scott Morton, F. (2018) "Antitrust Enforcement Against Platform MFNs," *Yale Law Journal*, Vol. 127(7), pp. 2176-2202.

Baye, M. R. and Morgan, J. (2001). "Information Gatekeepers on the Internet and the Competitiveness of Homogeneous Product Markets," *American Economic Review*, Vol. 91, pp. 454-474.

Bergemann, D. and Bonatti, A. (2023). "Data, Competition, and Digital Platforms," Working Paper.

Bisceglia, M. and Tirole, J. (2023). "Fair Gatekeeping in Digital Ecosystems," *Working Paper.*

Boik, A. and Corts, K. (2016) "The Effects of Platform Most-Favored-Nation Clauses on Competition and Entry," *Journal of Law and Economics*, Vol. 59(1), pp. 105-134.

Edelman, B. and Wright, J. (2015) "Price Coherence and Excessive Intermediation," *Quarterly Journal of Economics*, Vol. 130(3), pp. 1283-1328.

Gomes, R. and Mantovani, A. (2024) "Regulating Platform Fees under Price Parity," *Journal of the European Economic Association*, forthcoming.

Galeotti, A. and Moraga-González, J. L. (2009) "Platform Intermediation in a Market for Differentiated Products," *European Economic Review*, 53(4), pp. 417-428.

Haan, M. and Moraga-González, J. L. (2011) "Advertising for Attention in a Consumer Search Model," *Economic Journal*, Vol. 121 (552), pp. 552-579.

Hagiu, A. and Wright, J. (2024) "Marketplace leakage," Management Science, forthcoming.

Johnson, J. P. (2017). "The agency model and MFN clauses," Review of Economic Studies, 84(3), 1151-1185.

Lu, Y. (2020) "Can Shopify Compete With Amazon Without Becoming Amazon?" New York Times Magazine, November 24, 2020.

Mitchell, S. (2021) "Amazon's Toll Road," Institute for Local Self-Reliance Report.

Perloff, J.M. and Salop, S.C. (1985) "Equilibrium with Product Differentiation," *Review of Economic Studies*, Vol. 52, pp. 107-120.

Rochet, J.-C. and Tirole, J. (2002) "Cooperation Among Competitors: Some Economics of Payment Card Associations," *RAND Journal of Economics*, Vol. 33, pp. 549-570.

Rochet, J.-C. and Tirole, J. (2011) "Must-Take Cards: Merchant Discounts and Avoided Costs," *Journal of the European Economic Association*, Vol. 9, pp. 462-495.

Ronayne, D. (2021). "Price comparison websites," *International Economic Review*, Vol. 62, pp. 1081-1110.

Ronayne, D., and Taylor, G. (2020). "Competing sales channels," *Department of Economics working paper series*, University of Oxford, no. 843.

Ronayne, D., and Taylor, G. (2022). "Competing sales channels with captive consumers," *Economic Journal*, Vol. 132, pp. 741-766.

Salop, S. C. (1979) "Monopolistic competition with outside goods," *Bell Journal of Economics*, Vol. 10, pp. 141-156.

Sullivan, M. (2022) "Price controls in a multi-sided market," *Working paper*.

Wang, C. and Wright, J. (2020) "Search Platforms: Showrooming and Price Parity Clauses," *RAND Journal of Economics*, Vol. 51(1), pp. 32-58.

Wolinsky, A. (1986) "True monopolistic competition as a result of imperfect information." *Quarterly Journal of Economics*, Vol. 101(3), pp. 493-511.

Wright, J. (2004) "Determinants of Optimal Interchange Fees," *Journal of Industrial Economics*, Vol. 52, pp. 1-26.

Wright, J. (2012) "Why Payment Card Fees are Biased Against Retailers," *RAND Journal of Economics*, Vol. 43, pp. 761-780.

Online Appendix: Regulating platform fees

Chengsi Wang[1] and Julian Wright[2]

We present some extensions referred to in the main paper.

A Percentage fees

Suppose firms face a marginal cost d, and M sets a percentage fee r, where $0 < r < 1$. In the case with transaction fees, a firm maximizes $(p_M^i - f - d)D_i^M(p_M^i, p_M)$ and the symmetric equilibrium price is $p_M = f + d - \frac{D_M^i(p_M,p_M)}{D_M^{i\prime}(p_M,p_M)}$. In the case of percentage fees, firm i maximizes $((1-r)p_M^i - d)D_M^i(p_M^i, p_M)$ and the symmetric equilibrium price is $p_M = \frac{d}{1-r} - \frac{D_M^i(p_M,p_M)}{D_M^{i\prime}(p_M,p_M)}$. Therefore, as long as $\mu_M = -\frac{D_M^i(p_M,p_M)}{D_M^{i\prime}(p_M,p_M)}$ is a constant that is independent of p_M, we can write $p_m = d + \frac{r}{1-r}d + \mu_M$ in the case of percentage fee. Note $p_D = d + \mu_D$ as before.

As in our baseline setting, consumers will choose M iff

$$\phi_M - p_M + b \geq \phi_D - p_D$$

or equivalently

$$b \geq f - \Delta_s - \Delta_m,$$

where we have redefined $f \equiv \frac{r}{1-r}d$ given that f is one-to-one increasing in r. Note that with this definition, we have $p_M = d + f + \mu_M$, so that it takes the same form as our baseline model.

The welfare maximizing choice of f is then determined by the same consideration as before, so

$$f^e = c + \Delta_m.$$

This implies the efficient percentage fee is

$$r^e = \frac{c + \Delta_m}{c + d + \Delta_m}.$$

In contrast, M chooses r to maximize

$$(rp_M - c)\left(1 - H\left(\frac{r}{1-r}d - \Delta_s - \Delta_m\right)\right)$$

[1]Monash University
[2]Department of Economics, National University of Singapore

or using the definition of f above, sets f to maximize

$$\Pi_M = \left(f - c + \frac{f\mu_M}{d+f} \right) (1 - H(f - \Delta_s - \Delta_m)). \tag{29}$$

Note if $\mu_M = 0$, which could for instance arise in certain microfoundations if $n \to \infty$, this is the same expression for profit maximization as in the baseline. In this case, percentage fees lead to equivalent results as the case with fixed per-unit fees.

The FOC corresponding to maximizing (29) is

$$\frac{f + \frac{f\mu_M}{d+f} - c}{1 + \frac{d\mu_M}{(d+f)^2}} = \frac{1 - H(f - \Delta_s - \Delta_m)}{h(f - \Delta_s - \Delta_m)}.$$

The LHS is strictly increasing in f. By the weakly increasing hazard rate, the RHS is weakly decreasing in f. Consider an increase of f from the level at which M's profit margin equals zero to $\bar{b} + \Delta_s + \Delta_m$. The LHS increases from 0 to a strictly positive value, while the RHS weakly decreases from a value greater than zero to zero, or is a positive constant if λ is constant (in this case $\bar{b} = \infty$ and the LHS approaches ∞ when f goes to $\bar{b} + \Delta_s + \Delta_m$). So there is a unique f^* solving the FOC. Assumption (1) ensures this unique f^* is such that some consumers go to M. Then, we can conclude that if and only if the RHS is greater than the LHS at $f^e = c + \Delta_m$, then $f^* \geq f^e$, i.e.,

$$f^* \geq f^e \quad \Leftrightarrow \quad \frac{1 - H(c - \Delta_s)}{h(c - \Delta_s)} \geq \frac{\Delta_m + \frac{(c+\Delta_m)\mu_M}{d+c+\Delta_m}}{1 + \frac{d\mu_M}{(d+c+\Delta_m)^2}}. \tag{30}$$

We have $f^e < f^*$ otherwise. Note provided $\mu_M > 0$, the term on the RHS of the inequality is higher than in the fixed per-unit fee case, where it was just Δ_m, implying that the condition for the profit-maximizing fee to exceed the efficient fee is now harder to satisfy. Thus, even though the efficient fee is equivalent under percentage and fixed per-unit fees, the percentage fee that maximizes M's profit is lower than the corresponding one under fixed per-unit fees.

The comparative statics of the comparison on the RHS in (30) with respect to Δ_s are similar to before; a greater Δ_s will make it easier for $f^* \geq f^e$. After substituting in $\Delta_m = \mu_D - \mu_M$ into the above expression, it can be confirmed that the RHS is increasing in μ_D and the RHS is decreasing in μ_M given that $\mu_D \geq \mu_M$, so a higher μ_D shifts the tradeoff towards M's fee being too high and a higher μ_M shifts the tradeoff towards M's fee being too low. Thus, the direction of the effect on Δ_s and μ_M and μ_D are preserved from the case with a fixed per-unit fee.

B Total user surplus created by the platform

The total user surplus created by M is

$$\int_{f-(\Delta_s+\Delta_m)}^{\overline{b}} (\Delta_s + b - f)dH(b),$$

is U-shaped in f (or in case b is exponentially distributed, decreasing in f given (2)) and equal to zero at $f = \Delta_m + \Delta_s + \overline{b}$ (since there would be no transactions on the platform).

In order that the existence of M weakly increases total user surplus, f must satisfy

$$f \leq \Delta_s + \widehat{b}\,(f)\,, \tag{31}$$

where $\widehat{b}\,(f) = \mathbb{E}[b|b \geq f - (\Delta_s + \Delta_m)]$ is the expected value of b for consumers using M given the fee f. If H is strictly log-concave, $\widehat{b}\,(f)$ is increasing in f but at a rate less than one. Thus, for strictly log-concave H, (31) implies there will be a unique cap for f, denoted f^c, below which total user surplus is higher whenever M operates. That is, $f^c = \Delta_s + \widehat{b}\,(f^c)$, and the existence of M increases total user surplus for all $f < f^c$. If we also assume (2) holds, then $c < f^c < \Delta_m + \Delta_s + \overline{b}$.

Gomes and Mantovani (2024) show that at their welfare-maximizing fee, the existence of the platform never reduces total user surplus. This is not always the case in our setting. To see this, assume b is strictly log-concave GPD distributed on $\left[\underline{b}, \overline{b}\right]$, which implies $\widehat{b}\,(f) = \frac{f-(\Delta_s+\Delta_m)+\sigma+\epsilon\underline{b}}{1+\epsilon}$, and so (31) implies the relevant cap is

$$f^c = \underline{b} + \Delta_s + \frac{\sigma - \Delta_m}{\epsilon}.$$

Using this as an extra constraint on the regulated fee implies

$$f^{reg} = \min\left(f^e, \underline{b} + \Delta_s + \frac{\sigma - \Delta_m}{\epsilon}\right).$$

In the limit case of the exponential distribution ($\epsilon \to 0$), the extra constraint is not binding whenever (2) holds, which implies no additional cap is needed beyond f^e to ensure total user surplus is enhanced by the existence of M. More generally, for high enough $\epsilon > 0$, the constraint can become binding, in which case the fee should be set below f^e if we want to ensure M's existence increases total user surplus.

C Sequential search specification

We consider how the sequential search framework of Wolinsky (1986) and Anderson and Renault (1999) fits our setup. We assume a continuum of consumers and firms, of measure one in each case. Each firm produces a horizontally differentiated product. The firms' production cost is normalized to zero. There are search frictions in both the direct market and on M, and consumers need to conduct sequential search in order to find out information about match values and prices. The search cost consumers incur each time they sample a firm on the platform s_M is assumed to be lower than its counterpart in the direct channel s_D.

The match value ξ between a consumer and a firm is drawn i.i.d. from the common distribution function G over $[0, \overline{\xi}]$. We assume G is twice continuously differentiable with a strictly increasing hazard rate and a strictly positive density function g over $[0, \overline{\xi}]$. Under these assumptions, the inverse hazard rate of ξ, $\lambda_\xi(z) = \frac{1-G(z)}{g(z)}$, strictly decreases in z.

Expecting that firms charge the symmetric equilibrium price on the direct channel, a consumer's expected value from searching directly, or the reservation value for using the direct channel, is x_D, implicitly determined by

$$\int_{x_D}^{\overline{\xi}} (\xi - x_D) dG(\xi) = s_D.$$

It is well known from literature, when there are infinitely many firms, a consumer will stop and buy the product at firm i if $\xi^i \geq x_D - p_D + p_D^i$, and will continue to search otherwise. We assume s_D is sufficiently small such that a unique value of x_D exists satisfying $0 < \lambda_\xi(x_D) < x_D$. Specifically, we assume $s_D < \underline{s}$, where $\underline{s} = \int_{\underline{x}}^{\overline{\xi}} (\xi - \underline{x}) dG(\xi)$ and $\underline{x}$ is uniquely defined by $\underline{x} = \lambda_\xi(\underline{x})$. We can similarly define consumers' reservation value of using M as x_M. Consumers who search on M will stop and buy at firm i if $\xi^i \geq x_M - p_M + p_M^i$, and will continue to search otherwise. Since $s_D > s_M$ and $\lambda_\xi(\cdot)$ is a decreasing function, we must have $\lambda_\xi(x_M) < \lambda_\xi(x_D) < x_D < x_M$. Note since consumers know f and their draw of b before making their channel choice, if they choose M and search on M, they will always buy. This reflects that there is a continuum of firms.

Given consumers' optimal stopping rule, firm i chooses p_D^i to maximize

$$p_D^i(1 - G(x_D - p_D^i + p_D))$$

As shown in Wang and Wright (2020), the FOC and symmetry imply the symmetric equilibrium price on M is given by

$$p_D = \lambda_\xi(x_D).$$

On M, firm i chooses p_M^i to maximize

$$(p_M^i - f)(1 - G(x_M - p_M^i + p_M))$$

The FOC and symmetry imply the symmetric equilibrium price on M is given by

$$p_M = f + \lambda_\xi(x_M).$$

Thus, mapping this to our general setting, we have $\phi_M = x_M$, $\phi_D = x_D$, $\mu_M = \lambda_\xi(x_M)$ and $\mu_D = \lambda_\xi(x_D)$, with $\phi_M > \phi_D$, $\mu_D > \mu_M$, $\phi_D \geq \mu_D$ and $\phi_M > \mu_M$.

D Perloff-Salop specification

We consider how the discrete-choice framework in Perloff and Salop (1985) fits our setup. There are $n \geq 3$ ex-ante symmetric firms and a unit mass of consumers. Each consumer can only encounter some fixed number $n_d = 2, ..., n-1$ firms in the direct market. After, but not before, a consumer visits the direct market, she can see prices and match values of the n_d firms. So a unilateral change in direct price does not change consumers' visiting decisions. The identity of the n_d firms are randomly distributed across consumers. The utility a consumer can get from buying at firm i is

$$u^i = v - p^i + \beta\xi^i,$$

where ξ^i is distributed according to a CDF G on $[\underline{\xi}, \overline{\xi}]$, which is iid across firms and consumers. The parameter β measures consumers' taste for product differentiation. The consumer's outside option is assumed to be 0. We assume v is great enough compared to $\max\{f^* + \mu_M, \mu_D\} - \beta\underline{\xi}$ such that the market is always fully covered.

In the direct market, the demand of a firm i when it charges p_D^i while all other firms charge p_D is

$$D^i(p_D^i, p_D) = \Pr\left[u^i \geq \max_{j\neq i} u^j\right] = \int_{\underline{\xi}}^{\overline{\xi}} \left(1 - G\left(\xi + \frac{p_D^i - p_D}{\beta}\right)\right) dG(\xi)^{n_D-1}.$$

Firm i chooses p_D^i to maximize $p_D^i D^i(p_D^i, p_D)$. The FOC and symmetry imply

$$\int_{\underline{\xi}}^{\overline{\xi}} (1 - G(\xi))\, dG(\xi)^{n_D-1} - \frac{p_D}{\beta} \int_{\underline{\xi}}^{\overline{\xi}} g(\xi) dG(\xi)^{n_D-1} = 0.$$

In a symmetric equilibrium, the first term is equal to $1/n_D$.[3] The symmetric equilibrium price then is

$$p_D = \mu_D = \frac{\beta}{n_D \int_{\underline{\xi}}^{\overline{\xi}} g(\xi) dG(\xi)^{n_D-1}}.$$

A consumer's expected gross surplus of visiting the direct market is

$$\phi_D = v + \beta \int_{\underline{\xi}}^{\overline{\xi}} \xi dG(\xi)^{n_D}.$$

Similarly, the symmetric equilibrium price on M is

$$p_M = f + \mu_M = f + \frac{\beta}{n \int_{\underline{\xi}}^{\overline{\xi}} g(\xi) dG(\xi)^{n-1}}$$

and a consumer's expected gross surplus of visiting M is

$$\phi_M = v + \beta \int_{\underline{\xi}}^{\overline{\xi}} \xi dG(\xi)^{n}.$$

Clearly, $\phi_M > \phi_D$ as $G(\xi)^n$ first-order stochastically dominates $G(\xi)^{n_D}$. Zhou (2017) shows that when $g(\xi)$ is log-concave, $\mu_M < \mu_D$ and the markup converges to zero when the number of firms goes to infinity.[4] The expressions in (7)-(8) follow directly from the above results.

Our specification in the direct market corresponds to the case of symmetry consideration sets and zero latent demand in Gomes and Mantovani (2024). If the

[3] If there exist some consumers who do not have access to the direct market, then we will have a latent demand $D_0 \in (0, 1)$. We then replace $1/n$ by $(1 - D_0)/n$.

[4] See Zhou, J. (2017) "Competitive bundling," *Econometrica*, Vol. 85(1), pp. 145-172.

latent demand is not zero, consumers who attend the direct market will either know n_D firms or no firms.

E Salop circular-city specification

Consider the application of the baseline model to the Salop circular-city model of competition. Assume there are a finite number $n \geq 2$ of firms which are equally located around a circle according to the standard Salop circular city model, with measure one of consumers uniformly located around the same circle. Consumers are willing to pay v for one unit of the good, but face a mismatch cost parameter of t. We will assume consumers' value for the product v is large enough and the mismatch cost t is low enough so that the competitive pricing equilibrium always arises given M's optimal choice of f. As shown below, sufficient conditions are $v > c+\lambda\left(\frac{t}{8}\right)+\frac{3t}{4}$ and $t < c+\lambda\left(\frac{t}{4}\right)$, which we assume hold in this Salop circular-city setting. Note since $c+\lambda\left(\frac{t}{8}\right) \geq c+\lambda\left(\frac{t}{4}\right) > t$ the two conditions imply $v > t$.

We assume each consumer can only visit one channel, either one of the firms directly or M. After selecting a firm to visit directly, the consumer observes this firm's price and location (so match value), and then decides whether to buy. Alternatively, if the consumer chooses to visit M, after observing price and location information on all listed firms, she decides which one of them to buy from.

Consider the direct market. Expecting that locations are random and that prices are symmetric, consumers randomly choose a firm. This means a consumer's visiting choice does not signal any information about her location on the Salop circle relative to the firm, and therefore her location can be viewed as being uniformly distributed on the Salop circle from the firm's perspective. After choosing a firm, the consumer will eventually buy from the firm if and only if $v - p - tx \geq 0$, or equivalently $x \leq \frac{v-p}{t}$. In addition, given firms and measure one of consumers are located along a circle of circumference one, the shortest distance between a consumer and a firm cannot exceed $\frac{1}{2}$. So the firm behaves as if it is a monopolist in the direct market and sets price to solve the following maximization problem

$$\max_{p} \left\{2p\min\left\{\frac{v-p}{t}, \frac{1}{2}\right\}\right\}.$$

The symmetric equilibrium price and so markup is $\mu_D = v - \frac{t}{2}$ and $\phi_D = v - \frac{t}{4}$, both of which will be positive given that $v > t$.

On M, depending on the magnitude of platform fee, there are three possible pricing equilibrium between firms as labelled in Salop (1979):

- Competitive equilibrium: If $f < v - \frac{3t}{2n}$, the market is covered and each firm charges $p_M = f + \frac{t}{n}$, so $\mu_M = \frac{t}{n}$, and the gross surplus of a consumer will be $\phi_M = v - \frac{t}{4n}$.

- Kinked equilibrium. If $v - \frac{3t}{2n} \leq f < v - \frac{t}{n}$, each firm charges $p_m = v - \frac{t}{2n}$ and all consumers visiting M end up buying. In particular, the consumers who are located exactly halfway between any two firms are indifferent about buying.

- Monopoly equilibrium. If $f \geq v - \frac{t}{n}$, each firm charges $p_m = \frac{v+f}{2}$ and consumers with $x < \frac{v-f}{2t}$ buy after visiting M. The total demand by consumers conditional on visiting M is $\frac{n(v-f)}{t}$.

In the competitive equilibrium price range, we have $\Delta_s = \frac{t}{4} - \frac{t}{4n} > 0$ and $\Delta_m = v - \frac{t}{2} - \frac{t}{n} > 0$ given $v > t$. Anticipating the competitive equilibrium, consumers choose M iff $b \geq f - \Delta_s - \Delta_m = f + \frac{5t}{4n} + \frac{t}{4} - v$. The fraction of them going through M will be $1 - H(f - \Delta_s - \Delta_m)$. The requirement for the competitive equilibrium to emerge $f \leq v - \frac{3t}{2n}$ also ensures that all consumers will buy after visiting M. We then need to determine M's optimal fee to see if it satisfies this constraint.

Lemma 1. *The platform sets $f^* < v - \frac{3t}{2n}$ in equilibrium.*

Proof of Lemma 1. We show M cannot do better setting a fee that induces a kinked equilibrium or a monopoly equilibrium. In the kinked equilibrium range, consumers choose M if and only if $b + v - \frac{t}{4n} - \left(v - \frac{t}{2n}\right) \geq \frac{t}{4}$, or equivalently, $b \geq \frac{t}{4} - \frac{t}{4n} = \frac{(n-1)t}{4n}$. For all kinked-equilibria, the equilibrium price on M is fixed and so is consumers' participation. Thus, the highest profit M can obtain in the kinked-equilibrium regime is by setting the highest possible fee that induces such a kinked equilibrium, i.e., $f = v - \frac{t}{n}$. Denote the corresponding platform profit by

$$\Pi^k(v) = \left(v - \frac{t}{n} - c\right)\left(1 - H\left(\frac{(n-1)t}{4n}\right)\right).$$

Moreover, $\frac{d\Pi^k(v)}{dv} = 1 - H\left(\frac{(n-1)t}{4n}\right)$.

The optimal fee among all fees giving rise to a competitive equilibrium is derived by solving

$$\max_f \ (f - c)\left(1 - H\left(\frac{t}{4} + \frac{5t}{4n} - v + f\right)\right), \text{ subject to } f \leq v - \frac{3t}{2n}.$$

Under our assumption $v > c + \lambda\left(\frac{(n-1)t}{4n}\right) + \frac{3t}{2n}$, we will next show that the constraint $f \leq v - \frac{3t}{2n}$ is never binding. Denote the unconstrained maximum profit that M can obtain by inducing a competitive equilibrium by

$$\Pi^c(v) = \max_f \ (f - c)\left(1 - H\left(\frac{t}{4} + \frac{5t}{4n} - v + f\right)\right).$$

By the implicit function theorem, $\frac{d\Pi^c(v)}{dv} = (f^* - c)h\left(\frac{t}{4} + \frac{5t}{4n} - v + f^*\right)$ where f^* is pinned down by the FOC with respect to f

$$1 - H\left(\frac{t}{4} + \frac{5t}{4n} - v + f^*\right) = (f^* - c)h\left(\frac{t}{4} + \frac{5t}{4n} - v + f^*\right). \tag{32}$$

(32) can be rewritten as

$$f^* - c = \lambda\left(\frac{t}{4} + \frac{5t}{4n} - v + f^*\right). \tag{33}$$

The LHS of (33) strictly increases from 0 to $v - \frac{3t}{2n} - c$ as f^* goes from c to $v - \frac{3t}{2n}$ (note our assumption on v implies $v - \frac{3t}{2n} > c$). The RHS weakly decreases from $\lambda\left(\frac{t}{4} + \frac{5t}{4n} - v + c\right)$ to $\lambda\left(\frac{(n-1)t}{4n}\right)$ as f^* goes from c to $v - \frac{3t}{2n}$. So we just need $v - \frac{3t}{2n} - c > \lambda\left(\frac{(n-1)t}{4n}\right)$ to have an intersection between the LHS and RHS. Moreover, $v - \frac{3t}{2n} - c$ is increasing in n and $\lambda\left(\frac{(n-1)t}{4n}\right)$ is decreasing in n. Thus, if the required inequality holds at $n = 2$, i.e., $v - \frac{3t}{4} - c > \lambda\left(\frac{t}{8}\right)$, it holds for all $n > 2$. This corresponds to our assumption on v, so we have shown there is a unique solution to (33) that lies in the interior of $\left[c, v - \frac{3t}{2n}\right]$. Given the quasiconcavity of the platform's objective function, this implies the constraints are not binding.

Since $f^* < v - \frac{3t}{2n}$, we have $1 - H\left(\frac{t}{4} + \frac{5t}{4n} - v + f^*\right) > 1 - H\left(\frac{(n-1)t}{4n}\right)$, and (32) implies

$$\frac{d\Pi^c(v)}{dv} > \frac{d\Pi^k(v)}{dv}.$$

If $\Pi^c(t) \geq \Pi^k(t)$, we can conclude $\Pi^c(v) \geq \Pi^k(v)$ for all $v \geq t$.

Note

$$\Pi^c(v) = \lambda\left(\frac{t}{4} + \frac{5t}{4n} - v + f^*\right)\left(1 - H\left(\frac{t}{4} + \frac{5t}{4n} - v + f^*\right)\right)$$

and since $1 - H\left(\frac{t}{4} + \frac{5t}{4n} - v + f^*\right) > 1 - H\left(\frac{(n-1)t}{4n}\right)$, as noted above, a sufficient condition for $\Pi^c(t) \geq \Pi^k(t)$, i.e., $\lambda\left(\frac{t}{4}+\frac{5t}{4n}-t+f^*\right) \geq t-\frac{t}{n}-c$ is

$$\lambda\left(\frac{t}{4}+\frac{5t}{4n}-t+t-\frac{3t}{2n}\right) \geq t-\frac{t}{n}-c$$

$$\Rightarrow \; c+\lambda\left(\frac{(n-1)t}{4n}\right) \geq \frac{(n-1)t}{n} \tag{34}$$

This is because $f^* < v-\frac{3t}{2n}$ and $\lambda(\cdot)$ is a non-increasing function. The LHS of (34) strictly decreases in n, while the RHS of (34) strictly increases in n. Taking $n \longrightarrow \infty$, (34) holds for all $n \geq 2$ provided $t \leq c+\lambda\left(\frac{t}{4}\right)$.

Finally, in the monopoly equilibrium range, if M treats consumer participation as constant in f, it would maximize its profit by solving

$$\max_f \; \left\{(f-c)\left(\frac{n(v-f)}{t}\right)\right\} \text{ subject to } \; f \geq v-\frac{t}{n}.$$

The solution is $f=\frac{v}{2}$ if $\frac{t}{n} \geq \frac{v}{2}$ and $f=v-\frac{t}{n}$ otherwise. Since $v>t$, the solution must be $f=v-\frac{t}{n}$ which coincides with the one inducing the kinked equilibrium. Taking into account that consumer participation is decreasing in f (reflecting that the firms' prices are increasing in f in this range), just reinforces that the constraint $f \geq v-\frac{t}{n}$ must be binding. Thus, in this range M will want to set $f=v-\frac{t}{n}$, which corresponds to the solution with the kinked equilibrium with $f=v-\frac{t}{n}$. Since we already showed this involves lower profit than in the competitive equilibrium range, the monopoly equilibrium solution must be worse for M. ■

F Comparative statics

Applying (10) to the three competition applications in the paper we get the following results:

Proposition 15. (Comparative statics) *The level of f^*, f^e and the tendency for the platform to set its fee too high ($L \equiv f^*-f^e$) change with the primitives of the respective competition models in the following ways:*

- *Sequential search model: Let λ_G be the inverse hazard rate of the distribution G on the match value ξ, which is decreasing. A decrease in search costs s_M on the platform or an increase in search costs s_D in the direct channel increases both f^* and f^e, and decreases L if and only if $|\lambda'| \leq |\lambda'_G|$.*

- *Random-utility model: An increase in the importance of match value (β) increases both f^* and f^e, with*

$$\frac{\partial L}{\partial \beta} < 0 \quad \Leftrightarrow \quad |\lambda'| < \frac{\frac{1}{\int_{\underline{\xi}}^{\overline{\xi}} g(\xi)dG(\xi)^{n_D-1}} - \frac{1}{\int_{\underline{\xi}}^{\overline{\xi}} g(\xi)dG(\xi)^{n-1}}}{\int_{\underline{\xi}}^{\overline{\xi}} \xi dG(\xi)^n - \int_{\underline{\xi}}^{\overline{\xi}} \xi dG(\xi)^{n_D}}. \tag{35}$$

An increase in the number of firms that consumers can evaluate on the platform increases both f^* and f^e, with

$$\frac{\partial L}{\partial n} < 0 \textit{ iff } \quad |\lambda'| \leq \frac{\frac{\int_{\underline{\xi}}^{\overline{\xi}} (g(\xi))^2 \left(G(\xi)^{n_D-2} + (n_D-1)G(\xi)^{n_D-2} \ln G(\xi)\right) d\xi}{\left(\int_{\underline{\xi}}^{\overline{\xi}} (g(\xi))^2 (n_D-1) G(\xi)^{n_D-2} d\xi\right)^2}}{\int_{\underline{\xi}}^{\overline{\xi}} \xi g(\xi) \left(G(\xi)^{n_D-1} + n_D G(\xi)^{n_D-1} \ln G(\xi)\right) d\xi} \equiv \epsilon(n), \tag{36}$$

and

$$\frac{\partial L}{\partial n_D} > 0 \textit{ iff } \quad |\lambda'| \leq \epsilon(n_D). \tag{37}$$

- *Circular-city model: An increase in the number of firms n listed on M increases both f^* and f^e, with*

$$\frac{\partial L}{\partial n} < 0 \quad \Leftrightarrow \quad |\lambda'| < 4. \tag{38}$$

An increase in product differentiation between firms t increases f^*, decreases f^e, and always increases L.

Proof of Proposition 15.

☐ Sequential search specification. In the sequential search model, the primitives of interest are s_D and s_M. Since $\phi_M = x_M$ is determined by $\int_{x_M}^{\overline{\xi}} (\xi - x_M) dG(\xi) = s_M$ and $\phi_D = x_D$ is determined by $\int_{x_D}^{\overline{\xi}} (\xi - x_D) dG(\xi) = s_D$, we know that ϕ_M decreases in s_M and ϕ_D decreases in s_D.

Holding s_M fixed, consider an increase in s_D. An increase in s_D decreases ϕ_D and therefore increases Δ_s. Moreover, a decrease in ϕ_D increases $\mu_D = \lambda_G(\phi_D)$ and thus Δ_m. Given that an increase in s_D increases both Δ_s and Δ_m, we can conclude that an increase in s_D increases both f^* and f^e. Holding s_D fixed, consider an increase in s_M. The analysis is exactly the opposite of an increase in s_D, so an increase in s_M decreases both f^* and f^e.

Next, consider how L changes with s_D and s_M. Totally differentiating $\int_{x_D}^{\overline{\xi}} (\xi - x_D) dG(\xi) = s_D$, we get $\frac{dx_D}{ds_D} = \frac{-1}{1-G(x_D)}$ and so $\frac{d\Delta_s}{ds_D} = \frac{1}{1-G(x_D)}$. Since $\Delta_m = \lambda_G(x_D) - \lambda_G(x_M)$, we have $\frac{d\Delta_m}{ds_D} = \frac{-\lambda'_G}{1-G(x_D)}$. Substituting these results into (10) and noting

λ' and λ'_G are both negative implies a decrease in s_M decreases L if and only if $|\lambda'| \leq |\lambda'_G|$. A similar argument applies to s_D.

□ Perloff-Salop specification. We have $\phi_D = v + \beta \int_{\underline{\xi}}^{\overline{\xi}} \xi dG(\xi)^{n_D}$ and $\phi_M = v + \beta \int_{\underline{\xi}}^{\overline{\xi}} \xi dG(\xi)^n$. Moreover, $\mu_D = \frac{\beta}{n_D \int_{\underline{\xi}}^{\overline{\xi}} g(\xi) dG(\xi)^{n_D - 1}}$ and $\mu_M = \frac{\beta}{n \int_{\underline{\xi}}^{\overline{\xi}} g(\xi) dG(\xi)^{n-1}}$. Holding n fixed, consider an increase in n_D. An increases in n_D increases ϕ_D and thus decreases Δ_s. Furthermore, an increase in n_D decreases μ_D and thus decreases Δ_m. Both effects work in the same direction, which is to decrease f^* and f^e. Holding n_D fixed, consider an increase in n. The analysis is exactly the opposite of an increase in n_D. So an increases in n increases both f^* and f^e. Furthermore, note that both Δ_s and Δ_m increase in β given that $G(\xi)^n$ FOSDs $G(\xi)^{n_D}$ and $G(\xi)^{n-1}$ FOSDs $G(\xi)^{n_D - 1}$, which implies an increase in β increases both f^* and f^e.

Finally, differentiating (6) and (7) with respect to β and using (10) and that $\Delta_s > 0$ and λ' is negative we get the result in (35). Differentiating (6) and (7) with respect to n_D and using (10) and that λ' is negative we get the result in (37). A parallel argument gives the result in (36).

□ Salop circular-city specification. In this setup, $\mu_D = v - \frac{t}{2}$ and $\mu_M = \frac{t}{n}$. Moreover, $\phi_D = v - \frac{t}{4}$ and $\phi_M = v - \frac{t}{4n}$. That is, $\Delta_s = \frac{t}{4} - \frac{t}{4n}$ and $\Delta_m = v - \frac{t}{2} - \frac{t}{n}$. An increase in either n increases both Δ_m and Δ_s, and thus increases f^* and f^e. An increase in t increases Δ_s but decreases Δ_m, and thus increases f^* but decreases f^e given $n \geq 2$.

Since we have $\frac{d\Delta_s}{dt} = \frac{1}{4} - \frac{1}{4n}$ and $\frac{d\Delta_m}{dt} = -\frac{1}{2} - \frac{1}{n}$, (10) implies

$$\frac{dL}{dt} \geq 0 \quad \Leftrightarrow \quad \lambda' \leq \frac{\frac{1}{2} + \frac{1}{n}}{\frac{1}{4} - \frac{1}{4n}}$$

which is always true since the right-hand-side is positive and $\lambda' < 0$. On the other hand, $\frac{d\Delta_s}{dn} = \frac{t}{4n^2}$ and $\frac{d\Delta_m}{dn} = \frac{t}{n^2}$, and using (10) and that λ' is negative gives the result in (38). ■

G Model with incomplete pass-through

We provide a tractable microfounded model with incomplete pass-through and unit demands. Consumers get value v from buying from a firm regardless of the channel. In either channel, firms are either monopolists or they compete head-to-head with one other firm according to homogenous Bertrand competition. When they decide whether to participate on M they do not know which situation they will be in – monopoly or competition. This is determined randomly. Assume on M, with probability $0 \leq \theta_M < 1$ they will be a monopolist and set a price of v to extract all the surplus they offer consumers, and with probability $1 - \theta_M$ they will be competing head-to-head and so price at their perceived marginal cost, earning nothing. Assume in the direct channel, the corresponding probabilities are θ_D and $1-\theta_D$, where $\theta_M \leq \theta_D \leq 1$. If $\theta_M = 0$ and $\theta_D = 1$ then the model captures the case of homogenous Bertrand-type price competition on M and monopoly pricing in the direct channel.

As in Section 4.1, we assume the number of transactions consumers make on each channel is independent of the fee (or price), provided that consumers obtain a weakly positive surplus. Firms get expected profit $\theta_M (v - f - d_M) q_M$ from each consumer they reach on M, so will join M provided $f \leq v - d_M$. Assuming v is high enough, this will not be a constraint on the analysis of fees. Firms get expected profit $\theta_D (v - d_D) q_D$ from each consumer they reach on the direct channel. We assume $d_D \geq d_M$. The firms' expected markup (per transaction) will be higher in the direct channel provided $f > d_D - d_M$ given that $\theta_D \geq \theta_M$.

Consumers get $(1 - \theta_M)(v - f - d_M) q_M$ from transactions on M since with probability θ_M they will have all their surplus extracted and with probability $1-\theta_M$ the firm will compete head-to-head with another firm on M, splitting the market equally and passing through their costs to consumers. Likewise, consumers get $(1 - \theta_D)(v - d_D) q_D$ from transactions on the direct channel.

Assume b is the benefit (or cost) of using M. So consumers choose M over the direct channel if

$$(1 - \theta_M)(v - f - d_M) q_M + b \geq (1 - \theta_D)(v - d_D) q_D$$

so

$$b \geq q_M (f - \Delta_s - \Delta_m(f)),$$

where $\Delta_s = (v - d_M) - (v - d_D)\frac{q_D}{q_M}$ and $\Delta_m(f) = \theta_D (v - d_D) \frac{q_D}{q_M} - \theta_M (v - f - d_M)$. The expected price on M will be $p_M(f) = \theta_M v + (1 - \theta_M)(f + d_M)$ and the expected price on the direct channel will be $p_D = \theta_D v + (1 - \theta_D) d_D$. The pass-through rate on M is $1 - \theta_M < 1$. Note $f - \Delta_s - \Delta_m(f) = p_M(f) - v - (p_D - v)\frac{q_D}{q_M}$.

Let's consider the various objectives of interest in this setting. First, welfare is

$$W = \int_{q_M\left(p_M(f)-v-(p_D-v)\frac{q_D}{q_M}\right)}^{\bar{b}} \left(q_M\left(v-c-d_M\right)+b\right) dH\left(b\right) + \int_{\underline{b}}^{q_M\left(p_M(f)-v-(p_D-v)\frac{q_D}{q_M}\right)} q_D\left(v-d_D\right) dH\left(b\right),$$

which is maximized when

$$p_M\left(f\right) = c + p_D\frac{q_D}{q_M} - \left(d_D\frac{q_D}{q_M} - d_M\right)$$

or equivalently

$$f^w = \frac{c + \left(\theta_D\frac{q_D}{q_M} - \theta_M\right)v - \left(\theta_D d_D\frac{q_D}{q_M} - \theta_M d_M\right)}{1-\theta_M}.$$

Note we can rewrite this as

$$f^w = c + \Delta_m\left(f^w\right),$$

consistent with our general result in Section 4.1 of the main paper.

Next consider M's choice of profit maximizing fee, which solves

$$f^* = c + \lambda\left(q_M\left(1-\theta_M\right)\left(f^*+d_M\right) - \left(1-\theta_D\right)d_D q_D - \left(\theta_D q_D - \theta_M q_M\right)v - (q_M-q_D)v\right).$$

The firms' total profit is

$$\Pi^{firms} = \int_{q_M\left(p_M(f)-p_D-v+v\frac{q_D}{q_M}\right)}^{\bar{b}} q_M\left(p_M\left(f\right)-f-d_M\right) dH\left(b\right) + \int_{\underline{b}}^{q_M\left(p_M(f)-p_D-v+v\frac{q_D}{q_M}\right)} q_D\left(p_D-d_D\right) dH\left(b\right).$$

The FOC is

$$\begin{aligned}\frac{d\Pi^{firms}}{df} &= -\left(q_M(p_M\left(f\right)-f-d_M\right) - q_D\left(p_D-d_D\right))\,p_M'\left(f\right)h - q_M\left(1-p_M'\left(f\right)\right)\left(1-H\right) \\ &= \left[q_D\left(\theta_D\left(v-d_D\right) - q_M\theta_M\left(v-f-d_M\right)\right)\left(1-\theta_M\right) - \theta_M\lambda\right]h,\end{aligned}$$

where h, H, and λ are all functions of $p_M\left(f\right)-p_D-v+v\frac{q_D}{q_M}$. Note that this derivative increases in f as λ decreases in f. Thus, if there exists $\tilde{f}$ such that $\frac{d\Pi^{firms}}{df} = 0$ at $f = \tilde{f}$, then $\frac{d\Pi^{firms}}{df} < 0$ for $f < \tilde{f}$ and $\frac{d\Pi^{firms}}{df} > 0$ for $f > \tilde{f}$. Therefore, $f = \tilde{f}$

achieves the minimum of firm profits.

We next want to see if $\tilde{f} \geq f^*$, in which case regulating a lower fee than f^* will increase firms' total profit, explaining why firms would prefer to lower f from M's unregulated level. To show this possibility, we focus on the special case of $q_M = q_D$ and the GPD distribution of b. The profit-maximizing platform fee becomes

$$f^* = \frac{c + \sigma + \epsilon\left(\left(1-\theta_D\right)d_D - \left(1-\theta_M\right)d_M + \left(\theta_D - \theta_M\right)v + \underline{b}\right)}{1 + \epsilon\left(1-\theta_M\right)}.$$

The fee that is worst for firms, i.e., $\tilde{f}$, is

$$\tilde{f} = \frac{\left(1-\theta_M\right)\left(\theta_D d_D - \theta_M d_M - \left(\theta_D - \theta_M\right)v\right) + \theta_M\left(\sigma + \epsilon\left(\left(\theta_D - \theta_M\right)v + \left(1-\theta_D\right)d_D - \left(1-\theta_M\right)d_M + \underline{b}\right)\right)}{\left(1+\epsilon\right)\theta_M\left(1-\theta_M\right)}.$$

After rearranging, we have $\tilde{f} \geq f^*$ if and only if

$$\frac{c - \left(d_D - d_M\right) + \sigma + \epsilon\left(\left(\theta_D - \theta_M\right)\left(v - d_D\right) + \underline{b}\right)}{1 + \epsilon\left(1-\theta_M\right)} \leq \frac{\sigma + \epsilon\left(\left(\theta_D - \theta_M\right)\left(v - d_D\right) + \underline{b}\right)}{\left(1+\epsilon\right)\left(1-\theta_M\right)} - \frac{\left(\theta_D - \theta_M\right)\left(v - d_D\right)}{\left(1+\epsilon\right)\theta_M}.$$

Then if $c \leq d_D - d_M$, and given $1 + \epsilon\left(1-\theta_M\right) > \left(1+\epsilon\right)\left(1-\theta_M\right)$, this inequality is true provided θ_D is not too much above θ_M. And in particular, this inequality will be true in case $\theta_M = \theta_D = \theta$, so the differential markup is purely endogenous to the setting of the fee. Thus, we have shown the possibility that regulating a lower fee than f^* will increase firms' total profit.

H Showrooming with joining benefits

Suppose b is a joining benefit that consumers get from M, so even if they switch to buy directly they still incur b. In addition, we assume $\underline{b} > -\Delta_s$, which just says no consumers face a cost of visiting M that is greater than the surplus differential. Assume for now all showrooming consumers who visit M will eventually choose to switch and buy directly. We will verify this later by showing that $p_D < p_M$ in equilibrium. We first characterize the demand facing an individual firm.

The demand from regular consumers who buy directly: These consumers search and buy products only in the direct market. They choose the direct channel because $\phi_M - p_M + b < \phi_D - p_D$. They buy from firm i if and only if $v^i \geq \phi_D - p_D + p_D^i$. So their demand is

$$D_{rD}^i \equiv (1-\rho)H(f - \Delta_s - \tilde{\Delta}_m)\frac{1 - G(\phi_D - p_D + p_D^i)}{1 - G(\phi_D)},$$

where $\tilde{\Delta}_m = p_D - (p_M - f)$.

The demand from regular consumers who buy on M: These consumers search and buy on M because they have $\phi_M - p_M + b \geq \phi_D - p_D$ and they cannot switch channel to purchase. They buy from firm i if and only if $v^i \geq \phi_M - p_M + p^i_D$. Their demand is

$$D^i_{rM} \equiv (1-\rho)(1 - H(f - \Delta_s - \tilde{\Delta}_m))\frac{1 - G(\phi_M - p_M + p^i_M)}{1 - G(\phi_M)}.$$

The demand from showrooming consumers who purchase directly: Since $b \geq -\Delta_s$, all showrooming consumers visit M. They will switch to buy from firm i directly if $p^i_M > p^i_D$ and

$$v^i - p^i_D \geq x_M - p_D.$$

Their demand is

$$D^i_{sD} \equiv \rho \mathbb{1}_{p^i_M > p^i_D} \frac{1 - G(\phi_M - p_D + p^i_D)}{1 - G(\phi_M)}.$$

The demand from showrooming consumers who purchase on M: Showrooming consumers will purchase from firm i on M if $p^i_M \leq p^i_D$ and

$$v^i - p^i_M \geq \phi_M + p_D.$$

In this case, their demand is

$$D^i_{sM} \equiv \rho(1 - \mathbb{1}_{p^i_M > p^i_D})\frac{1 - G(\phi_M - p_D + p^i_M)}{1 - G(\phi_M)}.$$

Note that D^i_{sD} and D^i_{sM} cannot co-exist.

We now explain why firm i will set $p^i_M > p^i_D$. Note that except for the fee f, $p^i_D D^i_{sD}$ and $(p^i_M - f)D^i_{sM}$ are identical if we interchange p^i_D and p^i_M. Thus, if we ignore all other terms in i's profit function, $p^i_M > p^i_D$ follows given $f \geq c$. Now consider the maximization of the remaining terms in firm i's profit function:

$$p^i_D D^i_{rD} + (p^i_M - f)D^i_{rM}.$$

Note that $p^i_D D^i_{rD}$ only depends on p^i_D while $(p^i_M - f)D^i_{rM}$ only depends on p^i_M. So the maximization problem reduces to $\max_{p^i_D} p^i_D D^i_{rD}$ and $\max_{p^i_M} (p^i_M - f)D^i_{rM}$. The solution of p^i_D to the former maximization must be smaller than the solution of p^i_M to the latter maximization given $f \geq c$ and $p_D < p_M$. Thus, from i's perspective, setting $p_M^i > p_D^i$ is optimal for each component of its profit, and so it must also be true for the sum of its profit, implying $\mathbb{1}_{p_M^i > p_D^i} = 1$ and $D_{sM}^i = 0$. This implies we can separately characterize p_D and p_M in equilibrium. Since p_M^i is only used to maximize $\max_{p_M^i} (p_M i - f) D_{rM}^i$, the FOC along with symmetry implies $p_M = f + \mu_M$.

Each firm i chooses p_D^i to maximize its profit in the direct market

$$\begin{aligned} & p_D^i (D_{rD}^i + D_{sD}^i) \\ =& p_D^i \left(\rho \frac{1 - G(\phi_M - p_D + p_D^i)}{1 - G(\phi_M)} + (1-\rho) H(f - \Delta_s - \widetilde{\Delta}_m) \frac{1 - G(\phi_D - p_D + p_D^i)}{1 - G(\phi_D)} \right). \end{aligned}$$

Suppose the symmetric equilibrium price p_D is characterized by the FOC such that p_D solves

$$p_D = \frac{(\rho + (1-\rho) H(f - \Delta_s - \widetilde{\Delta}_m)) \mu_D \mu_M}{\rho \mu_D + (1-\rho) H(f - \Delta_s - \widetilde{\Delta}_m) \mu_M}.$$

It is straightforward to show $\mu_M < p_D < \mu_D$ for $0 < \rho \leq 1$, with $p_D = \mu_D$ if $\rho = 0$ and $p_D = \mu_M$ if $\rho = 1$.

We next show that $p_D < p_M$ whenever $\rho > 0$ and $f > 0$. Suppose instead $p_D \geq p_M$. Using the expression of p_D and that $p_M = f + \mu_M$, this implies

$$(1-\rho) H(f - \Delta_s - \widetilde{\Delta}_m) (\mu_D - \mu_M) \mu_M \geq f \left(\rho \mu_D + (1-\rho) H(f - \Delta_s - \widetilde{\Delta}_m) \mu_M \right). \tag{39}$$

Since $f - \Delta_s - \widetilde{\Delta}_m = f - \Delta_s - (p_D - (p_M - f)) = p_M - p_D - \Delta_s$, the supposition $p_D \geq p_M$ implies $f - \Delta_s - \tilde{\Delta}_m \leq \Delta_s < \underline{b}$, implying $H(f - \Delta_s - \tilde{\Delta}_m) = 0$. Thus, the LHS of (39) is equal to zero, while the RHS is strictly positive for any $f > 0$ and $\rho > 0$, leading to a contradiction. This shows that indeed $p_D < p_M$ and all showrooming consumers who go to M will indeed want to switch and buy directly.

The social planner chooses f to maximize

$$\rho \int_{\underline{b}}^{\bar{b}} (\phi_M + b) \, dH(b) + (1-\rho) \left(\int_{f - \Delta_s - \widetilde{\Delta}_m}^{\bar{b}} (\phi_M + b - c) dH(b) + \int_{\underline{b}}^{f - \Delta_s - \widetilde{\Delta}_m} \phi_D dH(b) \right),$$

which implies the efficient fee has a familiar form

$$f^e = c + \widetilde{\Delta}_m,$$

where $\tilde{\Delta}_m = p_D(f^e) - \mu_M$.

Finally, note that the expression of p_D can be written as

$$p_D = \mu_D - \frac{\mu_D(\mu_D - \mu_M)}{\mu_D + \left(\frac{1}{\rho} - 1\right) H(f - \Delta_s - \tilde{\Delta}_m)\mu_M}.$$

Using the property that $p_D(f^e) = f^e + \mu_M - c$ and $f^e - \Delta_s - \widetilde{\Delta}_m = c - \Delta_s$, this can be rewritten as

$$f^e + \mu_M - c = \mu_D - \frac{\mu_D(\mu_D - \mu_M)}{\mu_D + \left(\frac{1}{\rho} - 1\right) H(c - \Delta_s)\mu_M},$$

which clearly shows f^e is decreasing in ρ. This completes the proof of Proposition 12 for the case where b arises from joining M.