B Accompanying appendix

B.1 Search based microfoundation

The paper assumes that consumers and firms interact in a probabilistically manner. One can consider a search process as the microfoundation for this interaction. Suppose consumers conduct sequential search with recall. The first search is free, and all subsequent search costs θ. When consumers search, they discover firms following the search process λ as defined in the main paper.

As is standard in the literature, when consumers face such as search process they stop searching whenever their expected gains from search is less than or equal to the cost of search. Define an R which represents the utility required to induce such a stopping rule.

Now turn our attention to a firm maximizing profits. As before, the mass of consumers entering the market is independent of a firm's pricing strategy because, in expectation, a firm is unable to unilaterally deviate in prices to affect the mass of consumers entering the market. Additionally, to ensure demand, the firm has to provide at least as much surplus as R. Therefore, the firm maximizes its profits $\pi(\alpha_j) = p_j \lambda_j n$ subject to the constraints $\alpha_j - p_j \geq R$ and $p_j > 0$. The optimal pricing strategy is $p_j^* = \max\{\min\{\frac{\alpha_j-\sigma}{2}, \alpha_j - R\}, 0\}$.

$$p_j^* = \begin{cases} \frac{\alpha_j-\sigma}{2} & \text{if } R < \sigma \\ \frac{\alpha_j-\sigma}{2} & \text{if } R > \sigma \text{ and } \alpha_j \geq 2R - \sigma \\ \alpha_j - R & \text{if } R > \sigma \text{ and } R < \alpha_j < 2R - \sigma \\ 0 & \text{otherwise.} \end{cases}$$

Since $p_j > 0$, firms are only active on the market if $\alpha_j > \max\{\sigma, R\}$. Then if $R > \sigma$, those with $R < \alpha_j < 2R - \sigma$ set the price $\alpha_j - R$, and those with $\alpha_j \geq 2R - \sigma$ set the price $\frac{\alpha_j-\sigma}{2}$. And if $\sigma > R$, all firms with $\alpha_j > \sigma$ are active on the platform, selling at price $p_j = \frac{\alpha_j-\sigma}{2}$.

For tractability, focus on the situation where $\sigma > R$. This means that consumers' expected utility is given by $\int_\sigma^1 \frac{\alpha_j-p_j-\sigma}{\int_\sigma^1 \alpha_h-p_h-\sigma \, d\alpha_h}(\alpha_j - p_j)\, d\alpha_j = \frac{1+2\sigma}{3}$. Then the reservation value for consumers can be computed following $E[\max\{\alpha_j - p_j^*, R\} - R] \geq \theta$. When $\sigma > R$, we know $E[\alpha_j - p_j] > R$, which means the stopping rule resolves to $R = \frac{1+2\sigma}{3} - \theta$. As in the main model, the platform's profit function is $\frac{1+2\sigma}{3} r \frac{1-\sigma}{3}$ and, in equilibrium, $\sigma = \frac{1}{4}$. Then ensuring $\sigma > R$, we require that $\theta > \frac{1}{4}$ for this equilibrium to hold.

This exercise highlights how recasting this probabilistic function as the density of a search outcome can lead to very familiar results from the search literature. Namely, in equilibrium, consumer stop searching after inspecting the first product.

Therefore, the proposed contest fashion of framing the recommender function can be microfounded in a model of directed consumer search.

B.2 No recommender system

When the platform does not implement a recommender system, I search for the consumer surplus maximizing equilibrium where firms follow a symmetric pricing strategy. Doing so allows me to find the smallest possible change in consumer surplus following the implementation of the recommender system. Moreover, this seems to be the most relevant benchmark as many regulators seem to focus on consumer outcomes.

Following the backward induction logic, consumers only purchase from firms with quality $\alpha > p(\alpha)$ and the mass of consumers active on the platform is $n = \int_{j\in\mathbf{N}} \frac{\alpha_j - p(\alpha_j)}{\int_{h\in\mathbf{N}} 1\ d\alpha_h}\ d\alpha_j$. The firm j makes profit $n(1-r)\frac{p(\alpha_j)}{\int_{h\in\mathbf{N}} 1\ d\alpha_h}$ and the profits of the platform is given by $nr\int_{j\in\mathbf{N}} \frac{p(\alpha_j)}{\int_{h\in\mathbf{N}} 1\ d\alpha_h}\ d\alpha_j$.

Total consumer surplus is given by $\int_0^n E[u] - c\ dc$. Let $\overline{p} = \int_{j\in\mathbf{N}} \frac{p(\alpha_j)}{\int_{h\in\mathbf{N}} 1\ d\alpha_h}\ d\alpha_j$, then consumer surplus can be rewritten as $(\int_{j\in\mathbf{N}} \frac{\alpha_j}{\int_{h\in\mathbf{N}} 1\ d\alpha_h}\ d\alpha_j - \overline{p})^2/2$. Note that $1 > \alpha > p(\alpha)$, it must be that consumer surplus is decreasing in $\overline{p}$. To maximize consumer surplus, $p(\alpha) = 0$ such that $\overline{p}$ is minimized. This leads to a consumer surplus of 1/8.

Therefore, the consumer surplus maximizing price is $p(\alpha) = 0$ such that all firms are active in the market. Firms and the platform make zero profits, and total welfare is driven by consumer surplus, 1/8.

B.3 Transparency and public education

The main analysis assumes consumers fully understand the equilibrium implications of how platforms develop recommender systems. However, given the complexity of recommender systems, there are concerns that consumers may not fully understand its implications. I model such naive consumers as consumers who may observe they are receiving useful recommendations from the platform but do not anticipate how firms adjust prices in response to the platform's recommender system. In other words, while a platform may select some $\sigma \neq 0$ and firms may price their products in response to this specification, naive consumers fail to recognize how firms account for σ in their pricing strategy, and over anticipate firm exit.

Recall that consumers join the platform based on their expected consumption utility, such that they are unable to unilaterally affect the mass of consumers joining the platform. Hence a firm's decision to be active and its price is independent of consumers decision to join the platform. Therefore, firms adopt the same pricing strategy regardless of consumers' naïveté, $p_j^* = \frac{\alpha_j - \sigma}{2}$ and there is a unique cutoff $\bar{\alpha} = \sigma$ such that only firms with quality above the cutoff are active on the platform (Corollary 1).

Naifs failing to anticipate how σ affects prices, means they believe each firm sets $p_j^N = \frac{\alpha_j}{2}$, and expect consumption surplus of $\frac{\alpha_j}{2}$ from each transaction. They anticipate to be matched to firms following

$$\lambda(\alpha_j, p_j^N, \mathbf{p}_{\text{-}j}^N, \sigma) = \begin{cases} \frac{\alpha_j - p_j^N - \sigma}{\int_{h \in \mathbf{N}^N} \alpha_h - p_h^N - \sigma \; d\alpha_h} & \text{if } \alpha_j - p_j^N - \sigma \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, they believe firms are active only if $\alpha_j > 2\sigma$. And naifs expected consumption utility from the platform is $\int_{2\sigma}^{1} \frac{\alpha_j - 2\sigma}{\int_{2\sigma}^{1} \alpha_h - 2\sigma \; d\alpha_h} \frac{\alpha_j}{2} \; d\alpha_j$, the mass of naifs joining the platform is $\frac{1+\sigma}{3}$.

Corollary 8. *Naifs' decision to join a platform differ from non-naifs in two ways: (i) they expect the consumption utility of $\alpha_j - p_j^N = \frac{\alpha_j}{2}$ from firm α_j; (ii) they overestimate the amount of screening, and believe only firms with $\alpha_j > 2\sigma$ are active.*

Proof of Corollary 8. Naifs' expected consumption utility is $E[u^N] = \int_{h \in \mathbf{N}^N} \lambda^N(\alpha_h)(\alpha_h - p^N(\alpha_h)) \; d\alpha_h$. Effect (i) is immediate as they believe they receive $\alpha_j - \frac{\alpha_j}{2} = \frac{\alpha_j}{2}$ from a transaction with firms of quality α_j. To see effect (ii) consider the condition for firm exit. Firms exit if (a) they have to set negative prices, or (b) they receive no demand. Since naifs believe $p^N = \frac{\alpha_j}{2} > 0$, they do not anticipate firm exit due to negative prices. Firms receive demand based on $\lambda(\alpha_j)$, this means naifs believe firms become inactive if $\frac{\alpha_j}{2} - \sigma < 0 \Leftrightarrow \alpha_j < 2\sigma$. This also means the mass of naive consumers joining the platform is $n^N = E[u^N] = \frac{1+\sigma}{3}$. □

The first effect in Corollary 8 shows how naifs, failing to account for how recommender systems that are more precise of value-for-money can drive price competition, anticipate that firms set higher prices. Thus, conditional on interacting with a particular firm, consumers expect less value. The second effect in Corollary 8 discusses how naive consumers overestimate the degree of screening which occurs. Because firms are only recommended to consumers if $\alpha_j - p_j^N - \sigma > 0 \Leftrightarrow \alpha_j > 2\sigma$, consumers overestimate the quality of firms on the platform (recall from Corollary 1 firms join if $\alpha_j > \sigma$). This second effect dominates and, in expectation, naifs are more willing to join the platform.

Proposition 8. *There exists a unique subgame perfect Nash equilibrium, and in equilibrium a monopolist platform facing naive consumers sets $\sigma^N = 0$.*

Proof of Proposition 8. To see this, observe that the mass of consumers joining the platform is $\frac{1+\sigma}{3}$. The platforms and firms correctly anticipate firm's strategy and all firms with quality $\alpha_j > \sigma$ join the platform. Thus the platform's profit function is $\Pi = r\frac{1+\sigma}{3} \int_{\sigma}^{1} \frac{\alpha_j - \sigma}{\int_{\sigma}^{1} \alpha_h - \sigma \; d\alpha_h} \frac{\alpha_j - \sigma}{2} \; d\alpha_j = r\frac{1+\sigma}{3}\frac{1-\sigma}{3}$. Then taking the first derivative with respect to σ, $\frac{-2\sigma}{9} < 0$ indicating that the platform prefers $\sigma = 0$. □

For some intuition, first notice that firms' optimal strategy is independent of consumer naïveté and the platform's per-transaction revenue is decreasing in σ. Then observe that naifs are inherently more willing to join the platform, the mass of naifs joining the platform is less elastic than non-naifs. In other words, a more precise recommender system improves demand from naive consumers by a smaller amount. This translates to a smaller benefit to platforms. Therefore, it is immediate to see that a platform prefers less precise recommender systems when it faces naive consumers.

Observe that platform's profit when consumers are naive is $\frac{r}{9}$, which is lower than when consumers are not naive, $\frac{r}{8}$. Further, recall from Corollary 2, consumer surplus improves when recommender systems are more precise, and the recommender system has to be sufficiently precise for consumers to be as well off as no recommender system. Hence, when consumers are naive, consumer surplus is $1/18$. Their overconfidence in the recommender system can lead to worse consumer surplus outcomes than no recommender systems.

Corollary 9. *When consumers are naive, both consumer surplus and total welfare are lower than the fully rational benchmark.*

Share of naive users It is possible to consider the case where only some β fraction of users are naive. This way, only a share of users behave as in Corollary 8, and all others behave as in the main model. In this case, the platform's profit function becomes $\Pi = (\beta\frac{1+\sigma}{3} + (1-\beta)\frac{1+2\sigma}{3})r\frac{1-\sigma}{3}$, and the profit maximizing level of precision is $\sigma = \frac{1-\beta}{2(2-\beta)}$ which is decreasing in β. This shows that when users are partially naive, the optimal level of precision is between the naive level and the monopoly level, $\sigma \in (\sigma^N, \sigma^m)$, and the level of precision is decreasing in the share of naive users.

B.4 No free returns

I show the effects of consumers not having free returns by relaxing the assumption that firms must offer a positive consumption utility to get the attention of consumers. Instead, consumers have a probability of engaging with all firms, this probability is increasing in the value-for-money the firms offer and even those with negative value-for-money have a positive probability engaging with consumers. To capture this, let

$$\lambda(\alpha_j, p_j, \mathbf{p}_{-j}, \sigma) = \frac{\alpha_j - p_j - \sigma + \overline{u}}{\int_{h\in\mathbf{N}} \alpha_h - p_h - \sigma + \overline{u} \; d\alpha_h}$$

where the additional $\overline{u}$ is the highest consumption utility consumers can get from a particular firm in equilibrium. This allows all firms to have at least some positive demand. Recall then that some mass of consumers choose to join the platform and firms are unable to unilaterally influence this mass of consumers. The firms maximize their profits

$$(1-r)n\frac{\alpha_j - p_j - \sigma + \overline{u}}{\int_{h\in\mathbf{N}} \alpha_h - p_h - \sigma + \overline{u} \; d\alpha_h} p_j$$

by setting the price $\frac{\alpha-\sigma+\overline{u}}{2}$, rewriting firms profit function,

$$(1-r)n\frac{(\alpha_j-\sigma+\overline{u})^2}{2\int_{h\in\mathbf{N}}\alpha_h-\sigma+\overline{u}\;d\alpha_h}.$$

Observe that the firms profits are increasing in α_j and note that a firm is only active if $\alpha_j-\sigma+\overline{u}>0$. Further, consumption utility from α_j is $\alpha_j-p_j=\frac{\alpha_j+\sigma-\overline{u}}{2}$ which is increasing in α. Therefore, $\overline{u}=\frac{1+\sigma}{3}$, the equilibrium consumption utility received from the highest quality firm. I now check which firms are active at a given σ. Firms which are active require $\alpha_j-\sigma+\overline{u}>0\Leftrightarrow\alpha_j>\frac{2\sigma-1}{3}$. Therefore all firms are active unless $\sigma>\frac{1}{2}$.

Suppose first that $\sigma\leq\frac{1}{2}$ such that all firms are active. Then consumption utility and the mass of consumers joining the platform evaluates to $n=\frac{2+4\sigma-4\sigma^2}{15-12\sigma}$. The firm's profit function becomes $\pi(\alpha_j)=(1-r)n\frac{(1+3\alpha_j-2\sigma)^2}{3(5-4\sigma)}$. The platform's profit function is $r\frac{2+4\sigma-4\sigma^2}{15-12\sigma}\frac{7-10\sigma+4\sigma^2}{15-12\sigma}$, which is strictly increasing over $\sigma\in[0,\frac{1}{2}]$. Therefore, it can never be optimal for the platform to set such a σ.

Now suppose otherwise, that $\sigma>\frac{1}{2}$. The mass of consumers joining the platform and the profits of a firm with quality α_j are

$$\begin{aligned} n&=\frac{1+4\sigma}{9}\\ \pi(\alpha_j)&=(1-r)\frac{1+4\sigma}{9}\frac{(1+3\alpha_j-2\sigma)^2}{4(2-\sigma)^2}. \end{aligned}$$

This means the platform's profit function is

$$r\frac{1+4\sigma}{9}\frac{2(2-\sigma)}{9}$$

and the profit maximizing level of σ is $\frac{7}{8}$. As before, I recover the standard two-sided platform arguments that the platform is balancing elasticities across both sides of the market. Importantly, in this scenario the platform also prefers only positive σ, recommender systems that are more precise of value-for-money.

B.5 Uninformative recommendations and distributional assumptions

The main model restricts $\sigma\in\mathbb{R}_+$ such that recommender systems are at least as precise as value-for-money. Relaxing this assumption, allow $\sigma\in\mathbb{R}$. Further, to understand how consumer costs affects the platform's decision, relax the assumption that $c_i\sim U[0,1]$ and allow this to follow some generic distribution, B with support $[0,1]$. Lemma 9 addresses the question of when a monopolist platform may prefer recommendations less precise than value-for-money.

Proposition 9. *A monopolist platform prefers recommendations which are more precise than value-for-money unless the distribution of consumer cost is too 'bottom-heavy'. In the special case where B is uniform, a monopolist always prefers recommendations which are more precise than value-for-money.*

The intuition is simple, suppose consumer cost is bottom-heavy, that is most consumers have low cost of joining the platform. Then a platform improving the precision of its recommender system is only able to attract marginally more consumers. However, recall more precise recommender systems lead firms to compete more aggressively in prices, which lowers the platform's per-transaction revenue. Hence, if the platform is only able to attract marginally more consumers, it may instead prefer uninformative recommender systems.

Proof of Proposition 9. I first consider the special case where B is uniform $[0, 1]$. Observe from (4) that if $\sigma < 0$ consumers only purchase from a firm if $\alpha_j + \sigma > 0$. Otherwise, consumers may purchase and return the product, which is equivalent to zero demand. To ensure there is at least some firm where consumers purchase, it must be $\sigma > -1$. This also implies that only firms with sufficiently high quality $\alpha_j > -\sigma$ are sell on the platform. Next observe that firms are only recommended to consumers with positive probability if $\alpha_j - \sigma > 0$. To ensure there is at least some firm which is recommended to consumers, it must be $\sigma < 1$. Turning to the firm problem, firms join the platform as long as they make (weakly) positive profit. There is no direct sales channel or outside option. Firms are unable to individually influence platform demand, n. To maximize their profits, firms set $p^* = \frac{\alpha_j - \sigma}{2}$. This means firms want to sell whenever $\alpha_j > \sigma$. As a result, if $\sigma < 0$, all firms want to sell in the market, but only those with $\alpha_j > -\sigma$ make sales. Now consider a platform strategy. Note that if $\sigma \geq 0$, I recover the result in the main model. Suppose the platform sets $\sigma < 0$. Then $E[u] = \frac{1-\sigma-2\sigma^2}{3(1-3\sigma)}$. The platform's profit function is $\Pi = \frac{1-\sigma-2\sigma^2}{3(1-3\sigma)} r \frac{1-4\sigma+7\sigma^2}{3(1-3\sigma)}$. This is strictly increasing in σ, $\frac{\partial \Pi}{\partial \sigma} = \frac{1+3\sigma+3\sigma^2-59\sigma^3+84\sigma^4}{9(1-3\sigma)^3} > 0 \; \forall \sigma \in (-1, 0)$. Therefore, $\sigma < 0$ cannot be an equilibrium.

Now consider the case where consumer cost follows a generic distribution B and let b represent its PDF.

Solving the game by backward induction, recall that consumers form expected consumption utility following (1) and firms set prices $\frac{\alpha_j - \sigma}{2}$, this is true for all $\sigma \in [-1, 1]$. Then suppose $\sigma < 0$. Consumers expected consumption utility is $\frac{1-\sigma-2\sigma^2}{3(1-3\sigma)}$. The mass of consumers joining the platform is $B(\frac{1-\sigma-2\sigma^2}{3(1-3\sigma)})$, and since $b > 0$ and $\frac{1-\sigma-2\sigma^2}{3(1-3\sigma)}$ is increasing in σ, it must be that the mass of consumers joining the platform is increasing in σ. The platform's profit function is now $\Pi = rB(\frac{1-\sigma-2\sigma^2}{3(1-3\sigma)})\frac{1-4\sigma+7\sigma^2}{3(1-3\sigma)}$. Then $\frac{\partial \Pi}{\partial \sigma} = rb(\cdot)\frac{2(1-2\sigma+3\sigma^2)}{3(1-3\sigma)^2}\frac{1-4\sigma+7\sigma^2}{3(1-3\sigma)} - \frac{rB(\cdot)(1-14\sigma+21\sigma^2)}{3(1-3\sigma)^2}$. This is positive as long as b is sufficiently large, $\frac{b(\frac{1-\sigma-2\sigma^2}{3(1-3\sigma)})}{B(\frac{1-\sigma-2\sigma^2}{3(1-3\sigma)})} > \frac{3(1-3\sigma)(1-14\sigma+21\sigma^2)}{2(1-2\sigma+3\sigma^2)(1-4\sigma+7\sigma^2)}$, it is never optimal for the platform to select $\sigma < 0$. In other words, the distribution of consumers cannot be too 'bottom-heavy' such that most consumers are easily attracted to join the platform.

Then considering the case where $\sigma > 0$, we know the platform's profit function is $\Pi = rB(\frac{1+2\sigma}{3})\frac{1-\sigma}{3}$, and $\frac{\partial \Pi}{\partial \sigma} = rb(\cdot)\frac{2}{3}\frac{1-\sigma}{3} - \frac{rB(\cdot)}{3}$. Evaluating this at $\sigma = 0$, we have $rb(\frac{1}{3})\frac{2}{9} - \frac{rB(\frac{1}{3})}{3}$, which is positive if b is sufficiently large. In other words, that the firm can attract sufficiently many consumers following an increase in σ, it is willing to make its recommender system at least as precise as value-for-money. The distribution of consumers cannot be too 'bottom-heavy'.

Therefore, as long as the distribution of consumers is sufficiently not 'bottom-heavy', consumers are not so easily attracted to join the platform, the platform prefers more precise recommendations.

$\square$

B.6 General recommender function

To relax the contest success function that represents the recommender system, observe that an increase in σ is identical to a right shift in the distribution of consumption utility. This is akin to a first order stochastic dominance. Hence, allow a general recommender technology which depends on consumption utility. Then, without loss, consider a recommender technology $Q(\alpha_j, p_j)$ which first order stochastic dominates another recommender technology $L(\alpha_j, p_j)$ in consumption utility.

Proposition 10. *A monopolist platform prefers Q to L if and only if its per-transaction revenue using Q is sufficiently high such that (7) holds. This per-transaction revenue can be lower than that of L. When consumers are naive, the monopolist platform is less likely (compared to before) to prefer Q to L if and only if $E_Q[\alpha_j - p_j^*] > E_Q[\alpha_j - p_j^L]$.*

If the recommender system is sufficiently precise of consumption utility, then per-transaction revenues can be made up for by higher volume of sales. This is why the per-transaction revenue that leads the platform to prefer Q to L does not need to be as high as that of L. The results on naive consumers are in-line with those in the main text, showing that naive consumers lead the platform to become less likely to adopt the more precise recommender system.

Proof of Proposition 10. Consider the following recommender systems $Q(\alpha_j, p_j)$ and $L(\alpha_j, p_j)$ and without loss of generality, suppose Q first order stochastic dominates L such that $E_Q[\alpha_j - p_j] \geq E_L[\alpha_j - p_j]$. Let their PDFs be $q(\alpha_j, p_j)$ and $l(\alpha_j, p_j)$ respectively, which are both weakly increasing in consumption utility, $\alpha_j - p_j$, such that the likelihood a consumer interacts with a firm depends on the relative value-for-money firms provide. In other words, they are weakly increasing in the first argument, quality, and weakly decreasing in the second, price.

Recall consumers have a uniformly distributed cost of joining the platform. Then the profit function of any firm on the platform adopting the recommender system $\Gamma \in \{Q, L\}$ is $\int_0^1 \gamma(\alpha_h, p_h)u(\alpha_h, p_h)\ d\alpha_h\, \gamma(\alpha_j, p_j)p_j(1-r)$. Note that because the recommender functions can possibly assign a firm with quality α_j to receive a recommendation with probability 0, this is equivalent to saying that the integral is over all firm qualities, rather than the set of firms $\mathbf{N}$. Following the logic that an individual firm is unable to unilaterally influence consumer's decision to join the platform, this means a firm's profit maximizing price is $p_j^{*\Gamma} = -\frac{\gamma(\alpha_j, p_j^*)}{\gamma_2'(\alpha_j, p_j^*)}$. Then a platform's profit function is $\Pi_\Gamma = -r\int_0^1 \gamma(\alpha_h, p_h^{*\Gamma})(\alpha_h + \frac{\gamma(\alpha_h, p_h^{*\Gamma})}{\gamma_2'(\alpha_h, p_h^{*\Gamma})})\ d\alpha_h \int_0^1 \frac{(\gamma(\alpha_j, p_j^{*\Gamma}))^2}{\gamma_2'(\alpha_j, p_j^{*\Gamma})}\ d\alpha_j$. Note that $\int_0^1 \gamma(\alpha_h, p_h^{*\Gamma})(\alpha_h + \frac{\gamma(\alpha_h, p_h^{*\Gamma})}{\gamma_2'(\alpha_h, p_h^{*\Gamma})})\ d\alpha_h = E_\Gamma[\alpha_j - p_j^{*\Gamma}]$.

Then the recommender system Q is preferred to L if

$$\begin{aligned} -rE_Q[\alpha_j - p_j^{*Q}]\int_0^1 \frac{(q(\alpha_h, p_h^{*Q}))^2}{q_2'(\alpha_h, p_h^{*Q})}\ d\alpha_h &> -rE_L[\alpha_j - p_j^{*L}]\int_0^1 \frac{(l(\alpha_h, p_h^{*L}))^2}{l_2'(\alpha_h, p_h^{*L})}\ d\alpha_h \\ \Leftrightarrow -\int_0^1 \frac{(q(\alpha_h, p_h^{*Q}))^2}{q_2'(\alpha_h, p_h^{*Q})}\ d\alpha_h &> -\int_0^1 \frac{(l(\alpha_h, p_h^{*L}))^2}{l_2'(\alpha_h, p_h^{*L})}\ d\alpha_h \frac{E_L[\alpha_j - p_j^{*L}]}{E_Q[\alpha_j - p_j^{*Q}]}, \end{aligned} \tag{7}$$

where $\frac{E_L[\alpha_j - p_j^{*L}]}{E_Q[\alpha_j - p_j^{*Q}]} \in (0, 1)$ and $-\int_0^1 \frac{(\gamma(\alpha_j, p_j^{*\Gamma}))^2}{\gamma_2'(\alpha_j, p_j^{*\Gamma})}\ d\alpha_j$ is the average per-transaction revenue. Hence, if the average per-transaction revenue following Q is sufficiently large (and need not be larger than L), the platform prefers Q. Note it is possible for the per-transaction revenue of Q to be less than L and the platform would still prefer Q because more consumers join the platform if the recommender function sufficiently favors products which provide higher value-for-money.

Now, allow consumers to be naive such that they do not correctly anticipate the effect that the change in recommender technology has on firms pricing strategy.

Suppose, without loss of generality, that consumers believe firms pricing strategy always follows the recommender system L. Then the profit function of firms when the recommender system is L are $\int_0^1 l(\alpha_h, p_h)u(\alpha_h, p_h)\ d\alpha_h\, l(\alpha_j, p_j)p_j(1-r)$ and their corresponding pricing strategy are $p_j^L = -\frac{l(\alpha_j, p_j^L)}{l_2'(\alpha_j, p_j^L)}$. The platform's profit function is $\Pi^L = -r\int_0^1 l(\alpha_h, p_h^L)(\alpha_h + \frac{l(\alpha_h, p_h^L)}{l_2'(\alpha_h, p_h^L)})\ d\alpha_h \int_0^1 \frac{(l(\alpha_j, p_j^L))^2}{l_2'(\alpha_j, p_j^L)}\ d\alpha_j$.

When the recommender system is Q, because consumers are naive, $n^Q = \int_0^1 q(\alpha_h, p_h^L)u(\alpha_h, p_h^L)\ d\alpha_h$. Hence, the profit function of firms are $\int_0^1 q(\alpha_h, p_h^L)u(\alpha_h, p_h^L)\ d\alpha_h\, q(\alpha_j, p_j)p_j(1-r)$. However, because firms cannot unilaterally influence consumer's decision to join the platform, their profit maximizing price is $p_j^Q = -\frac{q(\alpha_j, p_j^Q)}{q_2'(\alpha_j, p_j^Q)}$. The platform's profit function is now $\Pi^Q = -r\int_0^1 q(\alpha_h, p_h^L)(\alpha_h + \frac{l(\alpha_h, p_h^L)}{l_2'(\alpha_h, p_h^L)})\ d\alpha_h \int_0^1 \frac{(q(\alpha_j, p_j^Q))^2}{q_2'(\alpha_j, p_j^Q)}\ d\alpha_j$.

Then the recommender system Q is preferred to L if

$$-rE_Q[\alpha_j - p_j^L]\int_0^1 \frac{(q(\alpha_h, p_h^Q))^2}{q_2'(\alpha_h, p_h^Q)}\, d\alpha_h > -rE_L[\alpha_j - p_j^L]\int_0^1 \frac{(l(\alpha_h, p_h^L))^2}{l_2'(\alpha_h, p_h^L)}\, d\alpha_h$$
$$\Leftrightarrow -\int_0^1 \frac{(q(\alpha_h, p_h^Q))^2}{q_2'(\alpha_h, p_h^Q)}\, d\alpha_h > -\int_0^1 \frac{(l(\alpha_h, p_h^L))^2}{l_2'(\alpha_h, p_h^L)}\, d\alpha_h \frac{E_L[\alpha_j - p_j^L]}{E_Q[\alpha_j - p_j^L]}.$$

In other words, if $E_Q[\alpha_j - p_j^Q] > E_Q[\alpha_j - p_j^L]$ then the platform is less likely to prefer recommender system Q to L when compared to non-naive consumers.

□

B.6.1 Specific functional forms

In this section, I consider two specific functional forms which are common in the literature. While these functional forms yield qualitatively similar trade-offs for a monopolist platform, their complexity can already be seen in the monopolist setting. Hence, for tractability, I do not consider these functional forms in a setting with competing platforms.

Tullock contest with utility exponent. In this section, I define $\lambda(\alpha_j, p_j, \mathbf{p}_{\text{-}j}, \sigma) \equiv \frac{(\alpha_j - p_j)^\sigma}{\int_{h\in\mathbf{N}}(\alpha_h - p_h)^\sigma \, d\alpha_h}$, $\sigma > 0$. This formulation takes the familiar form of the Tullock contest success function with exponent parameter. A larger σ corresponds to recommender systems more precise of value-for-money, and $\sigma = 1$ represents a recommender system which is exactly based on value-for-money.

Consider firms pricing strategy, and recall that the mass of consumers joining the platform is $n = E[u]$ given by (1) such that individual firms are unable to influence the mass of consumers joining the platform. The firms profit function is $n(1-r)\frac{(\alpha - p)^\sigma}{\int_{h\in\mathbf{N}}(\alpha_h - p_h)^\sigma \, d\alpha_h} p$, and the profit maximizing pricing strategy is $p^* = \frac{\alpha}{1+\sigma}$. This means that all firms are always active on the platform, and because prices are decreasing in σ and higher σ means better match quality, the mass of consumers active on the platform increases in σ.

The mass of consumers joining the platform is therefore given by $n = \frac{\sigma}{2+\sigma}$ and consumer surplus is given by $\frac{\sigma^2}{2(2+\sigma)^2}$. Consumer surplus is increasing in σ.

The platform's profit function is therefore $\frac{\sigma}{2+\sigma} r \frac{1}{2+\sigma}$, which is maximized at $\sigma = 2$.

Therefore, like the main model, higher levels of σ causes firms to set lower prices and the platform prefers recommender systems which are more precise of value-for-money, which provides more consumer surplus than recommender systems that are purely based on value-for-money.

Logistic contest success function. In this section, I follow Casner and Teh (Forthcoming) and define $\lambda(\alpha_j, p_j, \mathbf{p}_{\text{-}j}, \sigma) \equiv \frac{\exp(\frac{\alpha_j - p_j}{\sigma})}{\int_{h\in\mathbf{N}} \exp(\frac{\alpha_h - p_h}{\sigma})\, d\alpha_h}$, allowing the platform to set any $\sigma \in \mathbf{R}_+$, where $\sigma = 1$ corresponds to matching based purely on value-for-money, and $\sigma \to \infty$ corresponds to completely random matching, and a lower level of σ reflects more precise recommender systems.

Consider firms pricing strategy, and recall that the mass of consumers joining the platform is $n = E[u]$, given by (1), such that firms are unable to influence the mass of consumers joining the platform. Therefore, firms optimal prices are $p^* = \sigma$. Since firms want to make positive profits, they choose to sell only if $p^* > 0$. Further, because consumers have free returns, only firms with quality $\alpha > \sigma$ sell on the platform and all other firms exit. Therefore, the platform's strategy is constrained to $\sigma \in (0, 1]$. Then the mass of consumers choosing to enter the market evaluates to $n = \frac{\exp(1/\sigma)(1-\sigma)}{\exp(1/\sigma)-\exp(1)} - \sigma$. And the platform's profit function is $rn \int_\sigma^1 \frac{\exp(\alpha_h/\sigma)}{(\exp(1/\sigma)-\exp(1))\sigma} \sigma \, d\alpha_h = rn\sigma$. The platform's profit function is represented in Figure 5, and the optimal recommender system specification is $\sigma = 0.413$. This means, as in the main model, the platform prefers its recommender system to be more precise than value-for-money alone.

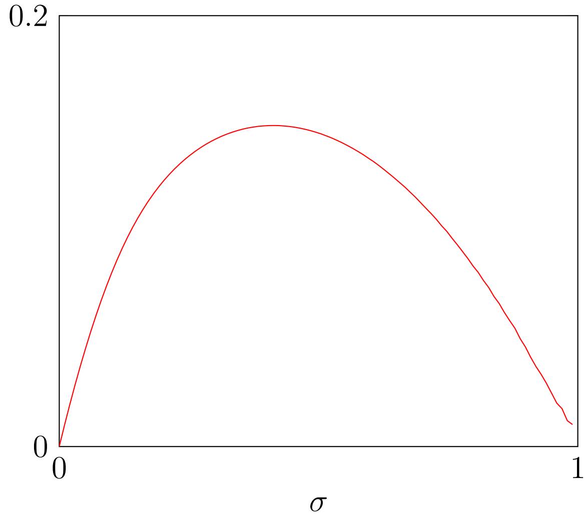

Figure 5: Platform profit function.

Qualitatively, the platform's trade-offs are similar to the main model. More precise recommender systems leads to lower per-transaction revenue, but can attract more consumers to the platform and increase the volume of transactions. However, there exists a screening effect that works in contrary to the main model. Whereas more precise recommendations lead to more screening in the main model, here, less precise recommender systems drive screening. This is because screening in the main model is driven by firms needing to set negative prices to obtain demand. And here, screening is driven by prices being too high when the recommender system is uninformative, such that consumers return all low quality products.

As in the main model, the presence of naive consumers means the platform is less able to use precise recommendations to influence consumers' decision to join the platform, and hence prefers to improve its profit by developing less precise recommendations to raise prices instead.

B.7 Entrant: Asymmetric consumer cost

From the base model, allow for the following modification: Suppose consumer cost for each platform is drawn from different distributions. For tractability, let the cost of the incumbent platform $c_{i,I}$ be drawn from a uniform distribution R with support $[0,1]$ and the cost of the entrant $c_{i,E}$ be drawn from a triangular T distribution with support $[0,1]$ with peak 1. Proposition 11 summarizes the results of this modification.

Proposition 11. *There exists a unique equilibrium, and in equilibrium* $\sigma_E = 0.547 > \sigma_I = 0.335$.

This shows that the platform facing higher consumer cost prefers to develop a more precise recommender system. Intuitively, this is because the entrant has to create more incentives for consumers to join the platform. Hence, this platform is more willing to give up more per-transaction revenue than the incumbent to attract consumers.

Proof of Proposition 11. The expected consumption utility from joining either platform is $\frac{1+2\sigma_k}{3}\ \forall k \in \{I, C\}$. The mass of consumers joining either platform is given by

$$\begin{aligned}
n_I &= \int_0^{E[u_E]} R(E[u_I] - E[u_E] + c_{i,E})t(c_{i,E})\ dc_{i,E} + R(E[u_I])(1 - T(E[u_E])) \\
&= \frac{26 + 54\sigma_I - 6\sigma_E - 12\sigma_E^2 - 8\sigma_E^3}{81} \\
n_E &= \int_0^{E[u_I]} T(E[u_E] - E[u_I] + c_{i,I})r(c_{i,I})\ dc_{i,I} + T(E[u_E])(1 - R(E[u_I])) \\
&= \frac{7 + 36\sigma_E^2 - 6\sigma_I + 8\sigma_I^3 + 6\sigma_E(5 - 4\sigma_I - 4\sigma_I^2)}{81}
\end{aligned}$$

Since the firms' pricing strategy is independent of the mass of consumers joining the platform, $p^*_{j,k} = \frac{\alpha_j - \sigma_k}{2}\ \forall k \in \{I, C\}$. And the per-transaction revenue for a platform k is $r\frac{1-\sigma_k}{3}$. Thus, on the incumbent

On the incumbent,

$$\Pi_I = r\frac{26 + 54\sigma_I - 6\sigma_E - 12\sigma_E^2 - 8\sigma_E^3}{81}\frac{1-\sigma_I}{3}$$

and we can solve that $\sigma_I = \frac{14+3\sigma_E+6\sigma_E^2+4\sigma_E^3}{54}$ maximizes profits.

On the entrant,

$$\Pi_E = r\frac{7 + 36\sigma_E^2 - 6\sigma_I + 8\sigma_I^3 + 6\sigma_E(5 - 4\sigma_I - 4\sigma_I^2)}{81}\frac{1-\sigma_E}{3}$$

and $\sigma_E = \frac{1+4\sigma_I+4\sigma_I^2 \pm \sqrt{70-46\sigma_I-48\sigma_I^2+8\sigma_I^3+16\sigma_I^4}}{18}$ maximizes its profit. Solving for σ_I and σ_E, we can show that the negative solution for σ_E does not satisfy the constraints of $\sigma_E \in [0, 1)$. And it is possible to obtain the numerical solutions $\sigma_I = 0.335$ and $\sigma_E = 0.547$.

□

B.8 Entrant: General recommender function

I relax the contest success function in the case of competition by adapting the setting provided in Section B.6. As before, without loss, consider a recommender technology $Q(\alpha_j, p_j)$ which first order stochastic dominates another recommender technology $L(\alpha_j, p_j)$ in consumption utility. Further, without loss of generality, impose that L is the monopolist optimal recommender system. I focus on the symmetric equilibrium where both platforms selects recommender systems of the same precision.

Proposition 12. *In the symmetric equilibrium, competing platforms prefer the recommender system Q to the monopolist profit-maximizing recommender system L if and only if their per-transaction revenue following Q is sufficiently large such that* (8) *holds. This per-transaction revenue can be lower than that of L.*

This provides the conditions for which my main results on how market power drives platform degradation continues to hold.

Proof of Proposition 12. Consider the following recommender systems $Q(\alpha_j, p_j)$ and $L(\alpha_j, p_j)$ and without loss of generality, suppose Q first order stochastic dominates L such that $E_Q[\alpha_j - p_j] \geq E_L[\alpha_j - p_j]$. Let their PDFs be $q(\alpha_j, p_j)$ and $l(\alpha_j, p_j)$ respectively, which are both weakly increasing in consumption utility, $\alpha_j - p_j$. In other words, they are weakly increasing in the first argument, quality, and weakly decreasing in the second, price.

Further, without loss of generality, impose that L is the monopolist optimal. In other words, $\Pi_L^m = -r\int_0^1 l(\alpha_h, p_h^*)(\alpha_h + \frac{l(\alpha_h,p_h^*)}{l_2'(\alpha_h,p_h^*)})\, d\alpha_h \int_0^1 \frac{(l(\alpha_j,p_j^*))^2}{l_2'(\alpha_j,p_j^*)}\, d\alpha_j$ is the highest possible profit a monopolist platform can achieve. For notational simplicity, recall that $E_L[\alpha_j - p_j] = \int_0^1 l(\alpha_h, p_h^*)(\alpha_h + \frac{l(\alpha_h,p_h^*)}{l_2'(\alpha_h,p_h^*)})\, d\alpha_h$.

Following the logic that an individual firm is unable to unilaterally influence consumer's decision to join a platform, this means a firm's profit maximizing price is $p_j^* = -\frac{\gamma(\alpha_j,p_j^*)}{\gamma_2'(\alpha_j,p_j^*)}$ as in Section B.6.

In the symmetric equilibrium, the mass of consumers joining each platform is given by (5). Therefore, each platforms' profit function is given by $-rE_L[\alpha_j - p_j^*]\left[1 - \frac{E_L[\alpha_j-p_j^*]}{2}\right]\int_0^1 \frac{(l(\alpha_j,p_j^*))^2}{l_2'(\alpha_j,p_j^*)}\, d\alpha_j = \Pi_L^m - \frac{(E_L[\alpha_j-p_j^*])^2}{2}(-r\int_0^1 \frac{(l(\alpha_j,p_j^*))^2}{l_2'(\alpha_j,p_j^*)}\, d\alpha_j) < \Pi_L^m$. This means that when a platform faces competition, any recommender function Q does not need to offer as much profit as L does when the platform is a monopolist.

Then any Q is preferred to L if and only if

$$-E_Q[\alpha_j - p_j^*]\left[1 - \frac{E_Q[\alpha_j - p_j^*]}{2}\right]\int_0^1 \frac{(q(\alpha_j, p_j^*))^2}{q_2'(\alpha_j, p_j^*)}\,d\alpha_j >$$
$$-E_L[\alpha_j - p_j^*]\left[1 - \frac{E_L[\alpha_j - p_j^*]}{2}\right]\int_0^1 \frac{(l(\alpha_j, p_j^*))^2}{l_2'(\alpha_j, p_j^*)}\,d\alpha_j.$$

By construction, $E_Q[\alpha_j - p_j] \geq E_L[\alpha_j - p_j]$. This means that $E_Q[\alpha_j - p_j^*]\left[1 - \frac{E_Q[\alpha_j - p_j^*]}{2}\right] \geq E_L[\alpha_j - p_j^*]\left[1 - \frac{E_L[\alpha_j - p_j^*]}{2}\right]$, and Q attracts more consumers to each platform than L. We can then rewrite the condition as

$$\frac{E_Q[\alpha_j - p_j^*]\left[1 - \frac{E_Q[\alpha_j - p_j^*]}{2}\right]}{E_L[\alpha_j - p_j^*]\left[1 - \frac{E_L[\alpha_j - p_j^*]}{2}\right]} > \frac{\int_0^1 \frac{(l(\alpha_j, p_j^*))^2}{l_2'(\alpha_j, p_j^*)}\,d\alpha_j}{\int_0^1 \frac{(q(\alpha_j, p_j^*))^2}{q_2'(\alpha_j, p_j^*)}\,d\alpha_j}. \tag{8}$$

As before, we recover that Q is preferred to L even if it offers less per-transaction revenue, and this is because it is able to attract more consumers to the platforms. Therefore as long as this condition holds, competing platforms prefer the more precise recommender system Q to the monopolist maximizing recommender system L. $\square$

C Empirical exercise

Using data from Prosper Marketplace prior to 2008, I conduct a simple empirical exercise and find patterns consistent with a platform's key trade-off described by the model.[23]

Prosper Marketplace

Prosper Marketplace was an auction site for peer-to-peer loans. It was one of the first, and the largest, such marketplaces in the USA. On the website, borrowers could request loans between 1,000 and 25,000 USD for a fixed term of three years. Borrowers submitted a listing comprising the loan amount and the maximum interest rate they were willing to pay. These listings were then published publicly, along with some information about the borrower's financial health (see Table 3 for the list of variables). Lenders could bid on these listings by proposing the amount they were willing to fund and setting a private minimum interest rate they were willing to charge. The loans were then allocated through a reverse auction.[24] On funded loans, Prosper charged between 1.5% and 2.5% of the loan amount in fees to the borrower. There are no listing fees.

On 12 February 2007, Prosper increased the amount of information it revealed about borrowers' financial health on each listing.[25] This change was unannounced and not anticipated. Of note, LendingClub, a marketplace offering similar services, was founded in October 2006 and began operations as part of Facebook Platform's launch applications in May 2007. Indicating the increase in information happened around the time a rival was planning to enter the market.

Mapping to model

The model proposed in this paper predicts that when platforms direct consumers to products with higher value-for-money, this causes prices to fall, and consequently improve consumer participation on the platform. Additionally, I show that one potential reason for platforms to improve how precisely they direct consumers to higher value-for-money products is competition.

It is difficult to obtain data about how a platform directs consumers to higher value-for-money products. For example, when the recommender system on Amazon changes there may be anecdotal evidence that consumers are more (or less) satisfied with the new system. However, this is only anecdotal, and there is no way to precisely measure if the system is improving the precision about value-for-money.[26] Hence, even if one has access

[23] My gratitude to Lauri Puro for providing the publicly available database used in Puro et al. (2011).

[24] The auction starts at the borrower's maximum interest rate and lenders bid for lower and lower interest rates until the loan is funded. If the loan is unfunded after 14 days, the listing expires. Partial funding was possible.

[25] This expanded information about credit lines and credit score.

[26] Other factors that may influence consumer satisfaction could be the independent of the product listings, for example the color scheme of the website, how long it takes to load, etc....

to data from Amazon, without knowing exactly the change made to the recommender systems it is not possible to identify if consumers understanding about value-for-money has increased.

On Prosper, lenders are provided objective information about borrowers financial health. More information allows lenders to better assess the financial health of borrowers, and hence evaluate the risk associated with financing the loan. Borrowers with a higher risk profile can be seen as lower value-for-money products as they are more likely to default on their loans. Since an increase in information objectively allows lenders to create a better picture of borrowers risk profile, this dataset provides a clean way to analyze the effects of improving the precision of product value-for-money.

On 12 February 2007, Prosper increased the number of variables it disclosed about borrowers credit, an exogenous shock to how precisely lenders can classify the risk associated with a listing. In the model, σ represents the precision about transaction value-for-money, and on Prosper this corresponds to more detailed borrower credit information. Therefore, this dataset allows for an objective measure in the change in σ. Hereafter, I refer to the 12 February 2007 as the treatment.

Studying Prosper is also appealing because of their fee structure. Prosper's fee structure is a percentage of the loan amount. The actual amount borrowers receive is the requested loan amount less Prosper's fee. This means when interest rates increase, borrowers have to request a higher amount to receive the desired payout, raising Prosper's per-transaction revenue. Since the total interest fees paid by the borrower are also proportional to the loan amount, this means Prosper's fees are a proportion of the surplus that borrowers receive from the loan. This corresponds to the proportional commission fee on prices that firms pay to the platform in my model.

The auction setting on Prosper has slightly different features from the model, two of which are relevant to this analysis. First, in the model, prices are determined by the firm. On prosper, prices correspond to the interest rates paid by borrowers, and these are determined following an auction. This difference can be mitigated by analyzing the maximum interest rate that borrowers set when submitting a listing. Therefore, this is analogous to a firm (borrower) submitting announcing a price (maximum interest rate) that he is willing to sell the product (returns on the loan). Note that an increase in maximum interest rate means that lenders (consumers) obtain more surplus from the loan, corresponding to lower prices in the model.

Second, the model assumes that firms provide a single product and consumers demand a single unit of product. On Prosper, borrowers can take multiple loans, and lenders can fund multiple loans. This means participation measures in the Prosper dataset cannot be mapped one-to-one with the model. Thus, in the analysis I look at multiple measures of lender activity to approximate consumer participation levels.

It is also interesting to observe that Prosper had been active since the end of 2005 and had not changed the amount of information displayed in listings until February 2007. Coincidentally, this occurred between the incorporation of LendingClub, a main rival in the market, and its official launch in May 2007 as part of a select group of applications for Facebook Platform at Facebook's first third-party developer conference, F8.[27] While not causal, the release of more information occurred during a period where competition was being introduced to the market, which aligns with the model's prediction that competition leads to recommendations more precise about value-for-money.

Data

Listing information	
Borrower max interest rate (%)	16.90 (7.03)
Loan amount (USD)	8225.67 (6831.49)
Credit grade	
AA = 7	0.035
A = 6	0.034
B = 5	0.049
C = 4	0.087
D = 3	0.125
E = 2	0.177
High risk = 1	0.484
No credit score = 0	0.009
Delinquencies (#, last 7 years)	0.694 (4.67)
(#, last 10 years)	0.043 (0.323)
Credit lines (#, last 7 years)	2.32 (8.31)
Debt to income ratio	0.760 (1.97)
Home ownership (Homeowner = 1)	0.297 (0.457)
Listing completed (Completed = 1)	0.057 (0.233)
Observations	52414

Borrower information	
Revolving payment (USD, monthly)	269.49 (444.41)
Income (USD, monthly)	5143.07 (6518.43)
Observations	2944

Bidding information	
Lender min interest rate (%)	16.64 (5.09)
Observations	461411

Marketplace information	
Loans completed (#, daily)	34.06 (23.21)
New lenders (#, daily)	143.97 (81.71)
Observations	151

Table 3: Summary statistics. Mean values with standard deviations in parentheses. Data from 1 December 2006 to 30 April 2007.

I use data from the two month window around the treatment (1 December 2006 to 30 April 2007). The choice of a two-month window was made to avoid any spillovers that may arise from the launch of LendingClub. Table 3 provides the summary statistics.

[27] The first employees of LendingClub also joined in February.

Within the sample period, there were 52414 listings observed, from this 3009 listings were funded. All listing information were publicly available. The variable borrower max interest rate is the maximum interest rate the borrower is willing to pay for the loan; the loan amount is the amount requested by the borrower; credit grade is assigned by prosper and reflects the borrower's credit score; delinquencies is the number of loans with late payments or defaulted; public records is the number of credit records from court databases; credit lines is the number of credit lines taken; home ownership reflects it the borrower is a homeowner; listing completed reflects it the loan was funded.

Additional information about borrowers were available for completed listings, and these were not known to lenders. These were revolving payments which is the monthly payment on revolving loans such as credit cards and income levels. However, a small sample of 65 reported income levels were unverified and I drop these from the sample, leaving 2944 observations.

Within the sample period, I observed 461411 bids, and use this information to compute the average number of bids per loan. Note that lender min interest rate is the minimum interest rates lenders were willing to be paid to fund the loan, and given only for lenders who did not win the auction.[28] I also observe 151 days (5 months) of aggregated marketplace information, which provides details about the number of completed loans and new lenders.

Findings

I first examine how more information about borrowers on a listing affects the borrower's choice of maximum interest rate then explore whether this has any impact on lender participation.

Borrower max interest rates Figure 3 shows the mean borrower's maximum interest rate and also when the borrowers are split by credit grades.[29] The mean maximum interest rate does not seem to increase following more information. However, this could be a result of other factors such as the prevailing interest rates on credit in the market, which behave as an outside option for borrowers, placing an exogenous cap on interest rates on Prosper. When splitting the sample by credit grades, borrowers with no or low credit grades set higher maximum interest rates, this is relatively flat and does not change following the treatment, possibly bounded by market interest rate. Borrowers with high credit grades choose higher interest rates following the treatment, suggesting it is possible that borrowers have a tendency to choose higher interest rates following the treatment.

Table 4 reports regression results about borrower's maximum interest rate following the treatment (12 February 2007), controlling for publicly available borrower features,

[28]The reserve price of winning bids was not made available.

[29]These classification of high and low credit grades were chosen arbitrarily and results are robust to other definitions, e.g. high grades being AA, A, B and low grades being E, high risk.

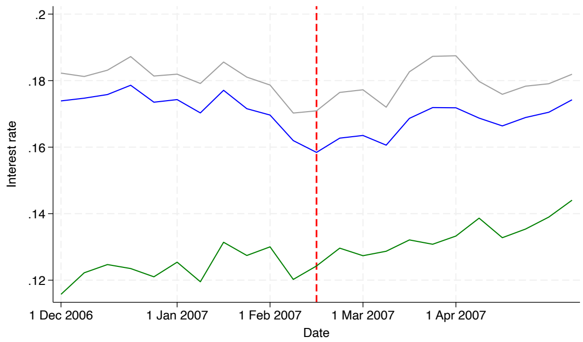

Figure 3: Mean weekly borrower max interest rates. Middle blue line: all listings. Bottom Green line: all high credit grades (AA, A, B, C) listings. Top gray line: all low credit grades (D, E, high risk, no credit score) listings. Red dashed line: 12 February 2007.

	Borrower max interest rate (%)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treatment	0.198	0.640***	0.618***	0.061	0.120	0.091	0.161
	(0.146)	(0.155)	(0.157)	(0.165)	(0.277)	(0.279)	(0.151)
High credit				-1.23***	-2.18***	-2.23***	
				(0.211)	(0.331)	(0.318)	
Treatment×High credit				0.663***	0.938***	0.958***	
				(0.179)	(0.318)	(0.314)	
Listing controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Borrower controls	No	No	Yes	No	No	Yes	No
Constant	19.47***	26.24***	26.25***	19.31***	26.04***	26.03***	19.35***
	(0.145)	(0.261)	(0.261)	(0.159)	(0.288)	(0.295)	(0.147)
Observations	52414	3009	2944	52414	3009	2944	49405

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Standard errors are two-way clustered at the borrower and listing date levels.

Table 4: Columns 1–3 report OLS estimates for all listings, completed listings, and completed listings with borrower controls, respectively. Columns 4–6 report Difference-in-Differences estimates, comparing high credit and low credit borrowers before and after the treatment for all listings, completed listings, and completed listings with borrower controls, respectively. Column 7 reports OLS estimates for uncompleted listings. Treatment is an indicator equal to 1 for listings posted on or after 12 February 2007, and high credit is an indicator equal to 1 for borrowers with credit grades AA, A, B, and C.

loan amount and private information about the borrower. Columns 1 to 3 report OLS estimates, and Columns 4 to 6 report Difference-in-Difference estimates, comparing high credit and low credit borrowers before and after the treatment. Column 1 and 4 provide estimates for all listings, Columns 2 and 5 provide estimates only for completed listings, and Columns 3 and 6 provide estimates only for completed listings and include private borrower information. Column 7 reports OLS estimates for uncompleted listings.

Column 1 shows the treatment has no effect on borrowers max interest rates. Columns 2 and 3 show that, on average, the treatment has a positive effect on borrowers max interest rates if listings were completed. This is congruent with observations from Figure 3 which suggests the treatment likely has heterogeneous effect depending on the borrowers financial profile. Since it is likely borrowers with better profiles are likely to get loans, the completed listings subsample possibly skews in the direction of borrowers with better financial health.

Columns 4 to 6 repeats the analysis, treating borrowers with a low credit score as a control group. This control group is motivated by stylized observation in Figure 3, and that borrowers face the option of market interest rates. Further, Column 7 of Table 4 shows that the treatment had no effect on the way low credit score borrowers set their maximum interest rate. It is possible that borrowers face exogenous interest rate limits, such as a large centralized financial market and regulatory controls.[30] This results in an upper bound on maximum interest rates, or equivalently, a maximum surplus lenders can achieve from a transaction. Therefore, it is natural to focus on the group of borrowers that have more leeway to change their maximum interest rates.

On average, Columns 4 to 6 show that high credit borrowers set lower interest rates than low credit borrowers, corresponding to higher quality firms setting higher prices. Importantly, high credit borrowers do increase their maximum interest rate following the treatment, with the average effect ranging from between 0.663% to 0.958%. This approximately reflects a 3.5% increase in the interest rates across all specifications.

	Loan amount (USD)
Treatment	824.03***
	(107.21)
Completed	-3271.49***
	(197.05)
Treatment×Completed	-581.27***
	(220.18)
Listing controls	Yes
Constant	4780.11
	(102.65)
Observations	52414

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Standard errors are two-way clustered at the borrower and listing date levels.

Table 5: The table provides OLS estimates for all listings. Treatment is an indicator equal to 1 for listings posted on or after 12 February 2007, and Completed is an indicator equal to 1 for loans that were funded.

Platform fees Because Propser's fees are proportional to the loan amount, I use loan amount as a proxy for the platform's per-transaction revenue. Table 5 shows that on average the loan amount for completed loans is significantly smaller than those that were not completed. Following the treatment, the average requested loan amounts increased, with weak evidence that loan amounts for completed loans decreased. Since the platform only takes commission from completed loans, this provides some evidence that Proper's per-transaction revenue fell falling the treatment.

[30] California had a cap of 19.2% for loans of \$1000 to \$2550 and 36% for larger loans. States with similar regulations include Arizona, Kentucky, Maine, Massachusetts, Minnesota and New Hampshire (Rigbi, 2013). This is consistent with the plot for low credit grades in Figure 3 which hover just below 20%.

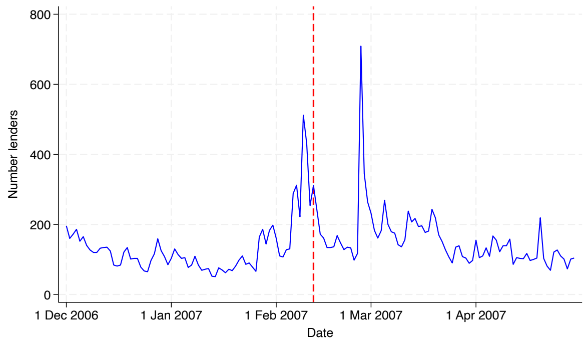

Figure 6: Blue line: new number of registered lenders, daily. Red dashed line: 12 February 2007.

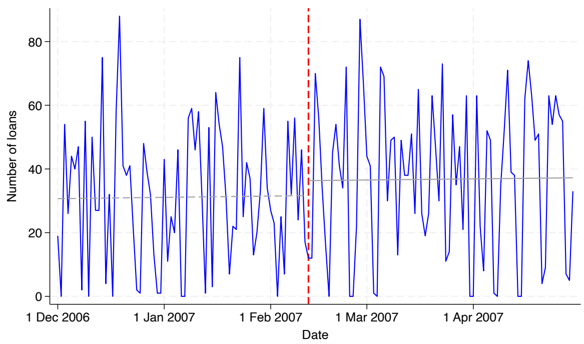

Figure 7: Blue line: number of loans completed, daily. Red dashed line: 12 February 2007. Gray lines: dashed = pre-trends, solid = post-trends.

Lender participation Turning to explore how more information affects lenders' participation in the marketplace. Figure 6 shows the daily change in registered lenders on Prosper. The number of daily new lenders is seemingly flat, with two spikes, one just before the release of more listing information and another towards the end of February. This seems to have no correlation with the release of data. Since the treatment was not announced, the spike before 12 February is likely unrelated to the change in available information. Moreover, in both instances, the growth of new lenders was not sustained, which suggests that they are unlikely a result of the treatment and more likely an outcome of other external forces such as advertisement or news coverage.

An alternative method of measuring market participation is to consider the number of loans completed each day. Figure 7 shows that there is an increase in the number of loans completed daily following the release of more listing information. Figure 8 shows a similar outcome by highlighting the share of completed listings by listing date. However, these may be driven by the composition of listings submitted. To isolate listing activity, Figure 4 shows the number of bids made per active listing. This shows that there is a noticeable increase in the number of active bids per listing following the treatment. Combined with the increase in the number of loans completed, this suggests that lenders become more active in the market.

In summary, the analysis provides suggestive evidence that more information about borrowers increases the maximum interest rates they are willing to offer. It also indicates that lender activity rises following an increase in borrower credit information. This provides some suggestive evidence for the on-platform mechanism in the model where more precise recommender systems lead to lower prices and more consumers. Additionally, I show that the platform's per-transaction revenues fall following more information, providing suggestive evidence consistent with the model's prediction that platforms face a trade-off between more trade volume and per-transaction revenue when deciding on the precision of their recommender system.

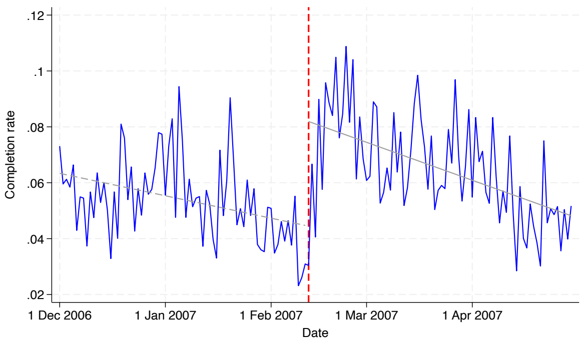

Figure 8: Blue line: completed as share of open listings, daily. Red dashed line: 12 February 2007. Gray lines: dashed = pre-trends, solid = post-trends.

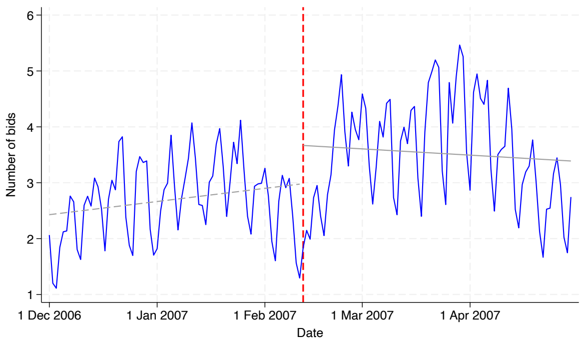

Figure 4: Blue line: mean number of bids per active listing, daily. Red dashed line: 12 February 2007. Gray lines: dashed = pre-trends, solid = post-trends.