Search Steering in Two-Sided Platforms

Pedro Guinsburg[*], Gustavo Saraiva[†]

Latest draft: click here. This draft: March 10, 2023.
First draft: April 10th, 2022.

Abstract

We study the incentives from a two-sided platform to segment the market by providing personalized search results. In our environment, a monopolistic platform matches sellers with buyers. Upon being matched, each pair of buyer and seller negotiate prices. If they choose to transact, the platform receives a commission fee proportional to the value of the transaction plus a flat fee per transaction. The platform is assumed to have full information over customers' and sellers' outside options. We derive sufficient conditions under which the platform's optimal matching is suboptimal from the perspective of buyers and sellers.

Keywords: Market Segmentation, Information Design, Two-sided markets

1 Introduction

With the proliferation of online marketplaces, such as Amazon and eBay, many expected that information frictions in these markets would disappear in the long run, thus improving customers' welfare. But theoretical and empirical research have emerged suggesting that platforms may have incentives to obfuscate information or segment the market in a way that could reduce customers' welfare. In this article, we conduct a theoretical analysis of a type of market distortion that has been overlooked by the literature, which occurs when the platform steers customers towards certain products based on their willingness to pay. We show that, in an environment in which the platform can discriminate search results, but cannot choose the prices of the goods sold by third-party sellers through its platform, the platform may have incentives to match those with higher willingness to pay to more expensive products in a way that will tend to generate more transactions than the socially optimum. This occurs if part of the commission fee charged by the platform is proportional to the product's price, as it is common in many online marketplaces, such as in Amazon.com and eBay.com.[1] We also present some conditions under which the platform is more likely to implement inefficient matches.

This research is motivated by growing concerns that e-commerce websites may be discriminating search results based on users' data. As an example, a study in 2012 found that the travel website Orbitz offered consistently higher prices for hotels to Mac users, as opposed to Windows users (Mattioli (2012)). Other studies, such as Hannak et al. (2014), have also found evidence that many e-commerce websites practice

[*]C6 Bank, Av. Nove de Julho, 3186 - Jardim Paulista, São Paulo-SP. E-mail address: pedro.guinsburg@gmail.com

[†]Business School, Pontificia Universidad Católica de Chile, Vicuña Mackenna 4860, Macul, Santiago, Chile. E-mail address: gsaraiva@uc.cl

[1]For Amazon's fee policy, see https://sellercentral.amazon.com/gp/help/external/200336920. For eBay, see https://pages.ebay.com/seller-center/seller-updates/2022-winter/fees-update.html.

personalized search results as a function of customers' cookies and browsing history. In 2000, it became public that Amazon.com was conducting A/B testing experiments on the prices of DVDs. Such policy was then quickly abandoned after many customers complained about the practice (Weiss (2000)).

Motivated by this empirical evidence, we build a theoretical framework in which a platform that intermediates transactions dictates one-to-one matches between buyers and sellers in exchange for a commission fee per sale, part of which is proportional to the final price chosen by the seller. We show that the platform has incentives to maximize the number of transactions, which will tend to be greater than the number of transactions that maximize buyers' and sellers' total surplus. The way the platform would create this excessive number of transactions is by allocating sellers with high production costs to customers with high willingness to pay, while allocating sellers with low production costs to customers with low willingness to pay, so as to increase market prices and, therefore, the commissions paid to the platform, and maximize the total number of transactions. For such a policy to be feasible, the platform must have information on customers' willingness to pay, as well as sellers' cost structure or willingness to sell.

The intuition as to why a platform would want to implement such a policy can be best described with the help of a simple example. Consider an economy with 5 sellers and 5 buyers. Each seller is willing to sell only one unit of their product, and each buyer has unitary demand. Each seller $s_j \in \{s_1, s_2, \cdots, s_5\}$ sells a product with quality $q_j \geq 0$, with a production cost of $c_j \geq 0$, where c_j could also be interpreted as the seller's outside option, i.e., the price at which they can sell their product outside the platform. Each buyer $b_i \in \{b_1, b_2, \cdots, b_5\}$ has an outside option $u_i \geq 0$. In figure 1 we order sellers' net value, $q_j - c_j$, in descending order, and buyers' outside option, u_i, in ascending order.

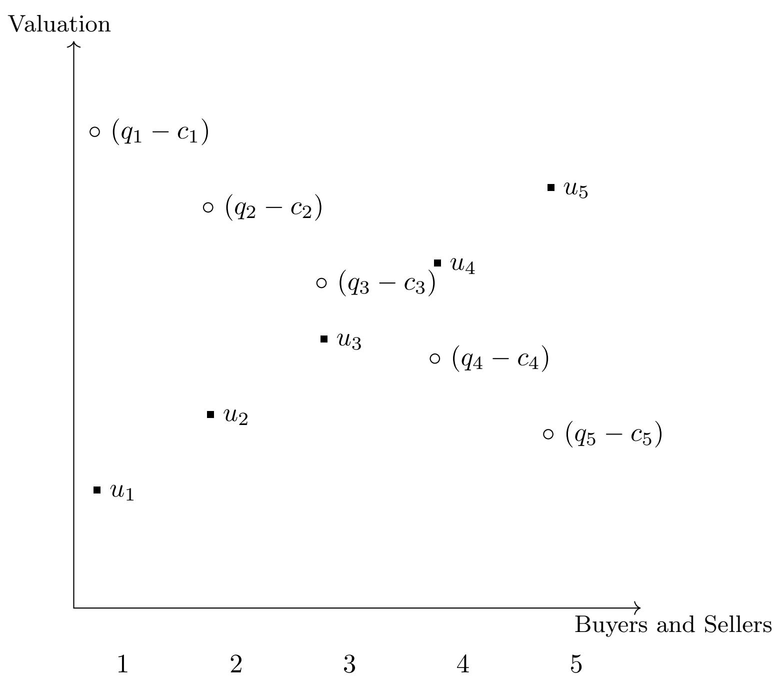


Figure 1: A plot depicting the net value by each seller, and the outside option from each buyer.

From figure 2, we can see that the total surplus from this economy is maximized when buyers b_1, b_2, b_3 (i.e., those with highest willingness to pay) transact with sellers s_1, s_2, s_3 (those who generate most net value), and the remaining buyers and sellers do not transact, so that only 3 transactions take place in the economy. In this case, the total surplus could be approximated by the blue region from the graph.

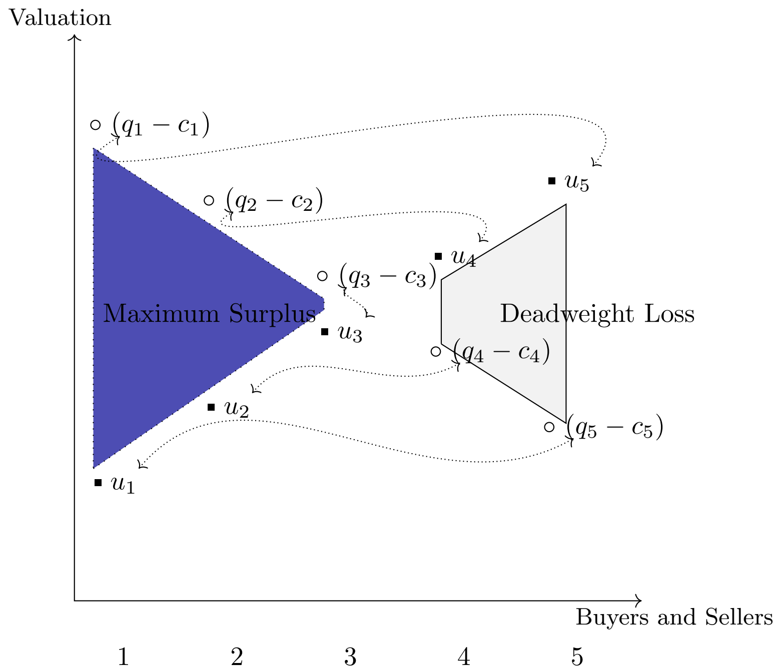


Figure 2: The plot depicts an approximation of the total surplus obtained when agents are matched so as to maximize the total number of transactions.

But the platform could, in principle, induce 5 transactions by matching each buyer $b_i \in \{b_1, b_2, \cdots, b_5\}$ with seller s_{6-i}. This matching is displayed in figure 2 by the curved arrows. Because in each match the net value created by the seller surpasses the customer's outside option, each match should yield a transaction, thus generating a total surplus given by the blue area minus the deadweight loss given by the grey area in the figure.

Assume now that the platform only collects a fixed fee per transaction that does not depend on the final price negotiated between buyer and seller, and assume that all products have the same quality q_j. In this case, the platform would have incentives to maximize the total number of transactions. Zhao et al. (2011) have shown that the number of transactions can be maximized by iteratively matching sellers with low costs with buyers with *low* willingness to pay, until a new match generates no positive surplus. In other words, they propose a matching criterion that prioritizes finding potential buyers to sellers with low production costs (i.e., they prioritize finding matches to more efficient sellers). So after implementing their algorithm, we have that if a seller with cost c_j manages to sell its product, a seller with cost $c_j' < c_j$ will also be able to sell its product. It can be shown that the final allocation obtained after implementing this algorithm is not necessarily Pareto efficient. So this result already illustrates how the platform may have incentives to sacrifice efficiency in exchange for higher revenues.

But we show that, to maximize the total number of transactions, the platform could implement an alternative and even more inefficient algorithm: it could prioritize securing feasible matches to sellers with *high* costs, by iteratively matching them with buyers with *high* willingness to pay, until a new match yields no transaction. This could lead to situations in which a seller with a high cost manages to sell its product, but a seller with low cost (or low outside option) does not. Though the platform has no strict incentives to adopt this more inefficient strategy in this scenario where commissions do not depend on the value of the transaction, things change when the platform also charges a commission fee that is proportional to the final price paid by customers, as it is common in many e-commerce websites, such as Amazon.com or eBay.com. We show that, in this case, the platform will have strict incentives to secure good matches to high cost sellers so as to inflate market prices, which in turn increases the commissions received by the platform per transaction. So in our example, if the commission fees are sufficiently small, it may be beneficial for the platform to match each buyer $b_i \in \{b_1, b_2, \cdots, b_5\}$ with seller s_{6-i}, so as to inflate market prices, and thus, the commissions paid by sellers.

We also provide a definition of stability and derive conditions under which there is a matching that maximizes buyers' and sellers' surplus and is more likely to be stable than any other matching (including profit-maximizing ones). Stability is a desirable property in this market, as it partially captures the platform's concerns about its reputation. Indeed, if customers were systematically matched with "bad sellers" in the sense that they are able to form better incentive-compatible matches on their own, they would probably be less inclined to revisit the platform in the future.

Notice that in our environment price negotiations between buyers and sellers happen after the match is determined by the platform. Though this would be more consistent with market structures in which some transactions are made through auctions (e.g., eBay.com) or bargaining (e.g., Upwork.com or Zillow.com), our results also apply to situations in which sellers first commit to a price, and then the platform determines the matching, and sellers are not sophisticated enough to predict the effect of their prices on the match that they get from the platform; or to situations in which, though sellers can predict the effect of prices on the matching assignment, they are bound to charge a price (net of commission fees) less than or equal the price they practice outside of the platform lest they suffer public criticism for practicing price discrimination, or violate some price parity clause imposed by the platform. In the latter case, seller s_i would charge a price equal to his outside option c_i, i.e., equal to the price it can sell its product outside of the platform, plus the commissions charged by the platform.

Also notice that the source of inefficiency from the model comes from the fact that the platform does not internalize the costs of manufacturing more expensive products. Indeed, if the platform was the owner and manufacturer of the products sold, prioritizing the sale of more expensive products would not be ideal, as the platform would have to pay the cost of manufacturing those more expensive products. So we expect our results to be more applicable to situations in which the platform is intermediating transactions from third-party sellers, instead of manufacturing the products themselves.

Though the results derived in this article have been framed in terms of a platform that intermediates transactions, it also applies to other two-sided markets, such as the one intermediated by real estate agencies.

2 Related Literature

The theoretical literature on incentives of e-commerce platforms is filled with examples that show how platforms may be willing to create information frictions that can potentially hurt customers, such as the seminal work of Diamond (1971) and Anderson and Renault (1999). In their theoretical model, customers sequentially search for a product through undirected search. So in each period, customers are randomly matched with a product and decide whether to buy it at the advertised price or keep searching for a better match. Like in our environment, they assume that prices of the final good are set by the third-party sellers, so that the platform cannot directly control the final prices of the products in the market. They show that, if there are infinitely many sellers operating in the platform selling the same homogeneous product, then, for any given search cost, no matter how small, there is an equilibrium in which all sellers choose the monopoly price. This result, known as the Diamond Paradox, is arguably very puzzling, as one would expect that competition in a market with many providers selling the same homogeneous product would drive prices down to the perfectly competitive level.

An implication of this result is that, if part of the platform's commissions are proportional to the price of the product sold (as it is common in many two-sided platforms, such as Amazon, eBay and Airbnb), then the platform may have incentives to foster search frictions to increase the final price charged by sellers. This observation has led others to present specific mechanisms through which platforms can increase search costs, such as the work of Eliaz and Spiegler (2011) and Casner (2020), who show that a platform may have incentives to allow some low quality sellers to enter the market in order to obfuscate search, or the work of Ellison (2005) and Ellison and Ellison (2009) who show that in a competitive environment firms may be able to achieve higher profits when add-on prices (e.g., shipping fees, resort fees, etc.) are hidden. These results illustrate how two-side platforms may sometimes be willing to add inefficiencies into their market to regain some control over the prices chosen by third-party sellers. These studies assume, however, that the platform is not in control over who gets matched with whom in the search process. In contrast, in our environment, the platform is the one determining the matches, and it uses this ability as a tool to regain some control over market prices.

In a different approach, Hagiu and Jullien (2011) studies a search model in which, instead of the search results being random, the platform is the one dictating the order of the search results from each customer. In their specification, customers are willing to buy more than one product, and products exhibit no substitutability whatsoever. So contrary to Diamond (1971) and Anderson and Renault (1999), in their specification, high search costs unambiguously hurt the platform, as it implies customers inspect fewer products and therefore make fewer purchases on the intensive and extensive margins. In our environment, on the other hand, products are perfect substitutes, so customers have unitary demand, in which case the platform may have incentives to limit customers' search ability to avoid competition between sellers, as competition causes equilibrium prices to go down, which reduces the commissions paid to the platform. So in our environment what motivates the platform to steer customers towards certain products is its desire to increase purchases on the extensive margin (but not on the intensive margin) and also to affect equilibrium prices. Additionally, our environment gives a fuller characterization as to how steering customers towards certain products affects efficiency, whereas Hagiu and Jullien (2011) is more focused on studying the conditions under which the platform may have incentives to steer customers towards less preferred products.

This article is also related to the literature on double auctions (i.e., where an auctioneer intermediates transactions between buyers and sellers), such as the work of Deshmukh et al. (2002), Ausubel et al. (2017) and Satterthwaite et al. (2022). But because the auctioneer is usually interested in implementing efficient and/or incentive compatible mechanisms, this literature has usually proposed and studied auction formats in which the sellers with lowest opportunity costs are the ones who end up transacting (assuming that the products sold are homogeneous) as it is the case in the uniform and pay-as-bid auctions. Indeed, both the uniform and pay-as-bid auctions are Pareto efficient when participants submit their true valuations, and the auctioneer does not collect distortionary commission fees. Though neither mechanism is strategy-proof (e.g., see Ausubel et al. (2014) or Binmore and Swierzbinski (2000)), the uniform price auction is known to be "*almost*" strategy-proof when the number of participants in the market is large, and each participant has a small market share on their side of the market (Azevedo and Budish (2019)).

More in line with our approach, Zhao et al. (2011) assumes that the auctioneer already knows the valuations from buyers and sellers, so that it can abstract from incentive compatibility issues. They propose the most efficient algorithm to match buyers with sellers among the ones that maximize the total number of transactions in the economy. This algorithm is attractive if the auctioneer collects a fixed revenue per transaction that does not depend on the final price at which the product is sold. This algorithm can generate more transactions than the socially optimum. In our specification, on the other hand, we assume that the platform (i.e., the auctioneer) also collects a fee per transaction that is proportional to the final price at which the product is sold. So we predict that the matching mechanism used by the platform will be even more inefficient than the one proposed by Zhao et al. (2011), as the platform will not only try to maximize the total number of transactions, but it will also prioritize finding eligible buyers to sellers with a high outside option (i.e., with high cost), so as to inflate the final prices paid by customers, and therefore, the fees received by the platform.

Our environment is very similar to the one in Boerner and Quint (2022). Though they also provide sufficient conditions under which the optimal policy maximizes surplus, those conditions only hold asymptotically and in expectation, and they require the market to be unbalanced. Meanwhile, we prove results for small markets that are valid for any realization of preferences, and the conditions that we derive that guarantee that the platform's optimal policy does not maximize buyers' and sellers' surplus is quite mild: we only require that there are either more buyers than sellers (i.e., the market is unbalanced) or that the "demand crosses supply" (see section 1 for details). Moreover, we show how to generalize the simulations conducted in Boerner and Quint (2022) to allow for horizontal differentiation, by making use of the Hungarian method.

Zhou and Zou (2022) developed an environment with two sellers and a platform in charge of recommending them to a continuum set of buyers. Different from our environment and the ones from Hagiu and Jullien (2011) and Boerner and Quint (2022), they allow the platform's recommendation to influence seller's pricing decisions. They show that, in some instances, the platform could be made better off if it was able to commit to not use prices as a recommendation criterion. They only derive results, however, for a case in which the market is comprised of only two firms, whereas our results apply to situations in which the market has multiple firms. Moreover, while they rely on the assumption that the market has a continuum set of customers, we follow the mechanism design approach by assuming that there is a discrete set of customers and sellers, which not only is more realistic, but allows variants of the algorithms we present to be more easily implemented in practical applications.

This article is also related to the literature on market segmentation, such as the work of Bergemann et al. (2015), Haghpanah and Siegel (2019) and Yang (2022). This literature, however, usually assumes that the seller segmenting the market is in control over the prices of each one of its products. In such an environment, market segmentation usually has the effect of improving total welfare. As an example, for the extreme case in which the seller practices perfect price discrimination by charging each customer their willingness to pay, total surplus is maximized, though in this case customers earn zero surplus. In our environment, on the other hand, the platform intermediating transactions between buyers and sellers cannot dictate the prices charged by sellers. As a result, attempts by the platform to segment the market in its favor will generate inefficiencies not predicted by the previous literature.

3 Model

A market is comprised of a set of n buyers, $B \equiv \{b_1, b_2, \cdots, b_n\}$, and a set of m sellers, $S \equiv \{s_1, s_2, \cdots, s_m\}$. Buyers have a unitary demand for the products sold by the sellers, and each seller can only sell one unit of their product to a single customer. For each $i \leq n$ and each $j \leq m$ we denote $q_{ij} \in \mathbb{R}_+$ as how much buyer $b_i \in B$ values the product from seller $s_j \in S$. We denote $c_j \in \mathbb{R}_+$ as the cost or the outside option from seller s_j of selling its product.

Without loss of generality, we order buyers and sellers in ascending order of their outside options and costs, respectively. That is, unless specified otherwise, we assume that

$$u_1 \leq u_2 \leq \cdots \leq u_n \tag{1}$$

and

$$c_1 \leq c_2 \leq \cdots \leq c_m. \tag{2}$$

A monopolistic platform is responsible for intermediating transactions between buyers and sellers. More precisely, the platform will match customers to sellers through a matching function μ. In this environment we will only allow for one-to-one matches, i.e., each seller can be matched with at most one customer and vice versa, though, as we explain later, the intuition of our main results may be extended to many-to-one matching environments. The matching function μ will map agents on one side of the market to agents on the other side of the market, with the possible exception of an agent being matched with himself to account for the possibility that the agent remains unmatched.

Definition 1. *A matching μ is a function mapping $B \cup S$ into itself, such that:*

i) $\forall b \in B$, $\mu(b) \in S \cup \{b\}$.

ii) $\forall s \in S$, $\mu(s) \in B \cup \{s\}$.

iii) $\forall b \in B$ *and* $s \in S$, $s \in \mu(b)$ *if and only if* $\mu(s) = b$.

To simplify the exposition, we sometimes also use the notation $s_j : b_i$ to indicate that seller s_j is matched to buyer b_i. In other words, $s_j : b_i$ is equivalent to writing $\mu(s_j) = b_i$ or $\mu(b_i) = s_j$.

Once the matching is determined, matched agents can transact. Suppose that the platform receives a fixed share $\psi \in [0, 1]$ of the price paid by the customer for each realized purchase, plus a flat commission fee $\tau \geq 0$ per transaction. Prices are determined ex-post, after the matches are formed, and after agents learn about the commission fees ψ and τ charged by the platform. In practice, prices could be determined after the matches take place if the products are being sold through an auction (e.g., through eBay), or if buyers and sellers are allowed to bargain after they are matched with one another, as in Upwork.com or Zillow.com. Though in principle one would expect an auction platform to match many potential buyers to a single seller in order to maximize the auctioneer's expected revenue, under the assumption of private independent values, the final price paid by the winning bidder is usually only affected by a few bids. Under the second price auction mechanism, for instance, agents have a weakly dominant strategy of bidding their true valuation, and the price paid only depends on the second highest bid. In this case, if the platform knew buyers' valuations, it would be wasteful to match more than two bidders to a given seller, as the bids below the second highest price would not affect the final price of the product, and those bidders could have actually helped increase the final price from other auctions. So our one-to-one matching environment can still give us some insight as to what could happen in platforms that intermediate some of the transactions through auction mechanisms, such as eBay.com.

We assume that prices are determined through Nash Bargaining, where each buyer has bargaining power $\alpha \in [0, 1]$ and each seller has bargaining power $1 - \alpha$. More precisely, if $\psi > 0$ and seller s_j is matched with a buyer b_i such that $(1-\psi)(q_{ij}-u_i) \geq c_j + \tau$, then a transaction will occur, and the seller will set its price at

$$p_j^*(b_i, \psi, \tau) = (1-\alpha)(q_{ij}-u_i) + \alpha \frac{c_j+\tau}{1-\psi}, \tag{3}$$

generating a revenue of

$$\psi p_j^*(\psi, \tau, b_i) + \tau = \psi \left[(1-\alpha)(q_{ij}-u_i) + \alpha \frac{c_j+\tau}{1-\psi} \right] + \tau$$

to the platform. If $(1-\psi)(q_{ij}-u_i) < c_j + \tau$, then no transaction transpires from the match between seller s_j and buyer b_i, so that the platform earns zero revenues from this match. If $\psi = 1$ and $c_j = \tau = 0$, the seller is indifferent between not selling and selling its product at any price. To guarantee that the profit function is upper semicontinuous, we will assume that in this case the seller will charge $p_j = \max\{q_{ij} - u_i, 0\}$.[2]

Notice that, when $\alpha = 1$, i.e., when buyers have all bargaining power, seller s_j elicits the price $(c_j + \tau)/(1-\psi)$ regardless of his match. Because in this case the matching algorithm chosen by the platform does not affect sellers' pricing strategy, we could invert the timing of the model by assuming that sellers first elicit a price, and then, based on those prices, the platform decides who gets matched with whom, which would be more consistent with prominent e-commerce websites such as Amazon.com. In practice, this could happen if, for instance, the platform imposes a price parity clause that prohibits each seller s_j from charging a price, net of commission fees, higher than the price c_j at which he sells its products outside of the platform (i.e., it prohibits each seller s_j from charging a price above $(c_j + \tau)/(1-\psi)$).

Taking the commission rates (ψ, τ) as given, we define, for each matching function μ,

$$FM_\mu \equiv \{(s_j, b_i) \in S \times B; \mu(s_j) = b_i \text{ and } (1-\psi)(q_{ij}-u_i) \geq c_j + \tau\},$$

i.e., FM_μ represents all the matches associated with μ that generate transactions.

Once a matching is formed it generates a total surplus to buyers and sellers equal to

$$TS(\mu) \equiv \sum_{(s_j,b_i) \in FM_\mu} (q_{ij} - \psi p_j^*(\psi, \tau, b_i) - c_j - u_i - \tau).$$

For a given pair of commissions (ψ, τ), we say a matching is constrained Pareto efficient (CPE) if it maximizes $TS(\mu)$. As in our economy agents have quasilinear utility, this definition of efficiency is equivalent to the notion that no seller or buyer can be made better off without making another buyer or seller worse off. Notice that this definition does not include the surplus extracted by the platform, as our main interest is in evaluating the platform's optimal policy on consumers' and sellers' total surplus. Also notice that this definition takes the fees as exogenous, so it also does not take into account the distortionary effects caused by the fees (ψ, τ). The reason we use this second-best notion of Pareto efficiency is because, if $\psi > 0$ or $\tau > 0$, the final allocation would almost always be inefficient for any matching μ due to the distortionary effects caused by the fees. So the only way to guarantee efficiency would be to completely eliminate the fees charged by the platform, which is not a very practical solution. That being said, our results regarding which matchings are CPE also apply to the case in which $\psi = \tau = 0$, i.e., the case in which our definition of CPE is equivalent to the (first-best) notion of Pareto efficiency.

[2]The hypothesis of upper semicontinuity of the objective function with respect to (ψ, τ) guarantees the existence of an optimal (ψ, τ) for any given match μ (see proposition 6).

Definition 2. *(Constrained Pareto Efficiency) For a given* (ψ, τ)*, we say a matching* μ *is Constrained Pareto Efficient (CPE) if*

$$TS(\mu) \geq TS(\mu')$$

for all possible alternative matchings μ'.

In the next section we characterize the platforms' optimal matching μ in terms of efficiency and number of transactions, taking its commission fees (ψ, τ) as given.

4 Optimal Matching with exogenous commissions

In this section we will consider the problem of finding the optimal matching taking the commission fees, ψ and τ, as given, i.e., we will characterize the solution to the following problem:

$$\max_{\mu} \tau|S_\mu| + \psi \sum_{s_j \in S_\mu} p_j^*(\mu(s_j), \psi, \tau), \tag{4}$$

where

$$S_\mu \equiv \{s_j \in S; \mu(s_j) = b_i \in B \wedge (1-\psi)(q_{ij} - u_i) \geq c_j + \tau\}$$

corresponds to the set of sellers who end up transacting given the fees ψ and τ and matching function μ, and $p_j^*(\mu(s_j), \psi, \tau)$ is the equilibrium price given by equation (3).

We first analyze the problem when all products have the same quality, i.e., when $q_{ij} = \overline{q}$ for all $i \leq n$ and all $j \leq m$. Then we discuss how to find the optimal matching when products are differentiated.

4.1 When products are homogeneous

In this section we will characterize Constrained Pareto Efficient (CPE) matchings and profit maximizing matchings when all products have the same quality. We will show that, in this case, finding a CPE matching is trivial: one only needs to ensure that the sellers and buyers transacting are the ones with lowest outside options. A simple algorithm can be used to find the platform's optimal matching for a given $\psi \in [0, 1]$ and $\tau \geq 0$. We show that this algorithm not only maximizes the profits of the platform, but also the number of transactions that can be made, given the commissions (ψ, τ). We also present necessary and sufficient conditions under which the matching obtained through this algorithm is *not* CPE.

Suppose that all products have the same quality, i.e., $q_{ij} = \overline{q} \in \mathbb{R}_+$ for all $i \leq n$ and all $j \leq m$. In this case, Constrained Pareto Efficiency is achieved by ensuring that the customers with lowest outside options and sellers with lowest costs transact. One way to achieve this is by implementing *positive assortative matching* (*PAM*), i.e., by having the customer with the lowest outside option being matched with the seller with lowest cost, the customer with second lowest outside option being matched with the seller with second lowest cost, and so on. But there are other matchings that are also *CPE*. In fact, if we initially apply *PAM* and then apply any permutation to the assignment of those who transact under this matching, then the new matching will be *CPE*.

Proposition 1. *(Necessary and sufficient condition for* μ *to be CPE) Suppose that* $q_{ij} = \overline{q} \in \mathbb{R}_+$ *for all* $i \leq n$ *and all* $j \leq m$*. A matching* μ *is Constrained Pareto Efficient (CPE) if and only if each seller* s_j *such that* $j \leq \min\{n, m\}$ *and* $(1-\psi)(\overline{q} - u_j) \geq c_j + \tau$ *is matched to a buyer* b_i *such that* $(1-\psi)(\overline{q} - u_i) \geq c_i + \tau$.

Proof: Proof in the Appendix. ■

Corollary 1. *(PAM is CPE) Suppose that $q_{ij} = \overline{q} \in \mathbb{R}_+$ for all $i \leq n$ and all $j \leq m$. If $\mu(s_i) = b_i$ for all $i \leq \min\{n, m\}$, then μ is CPE.*

It follows directly from proposition 1 that a *CPE* matching μ maximizes the number of transactions subject to the constraint that every seller who transacts under μ must also be willing to transact with any other buyer who transacts under μ.

Corollary 2. *(CPE matchings maximize the number of transactions subject to a constraint) Suppose that $q_{ij} = \overline{q} \in \mathbb{R}_+$ for all $i \leq n$ and all $j \leq m$. If a matching μ is CPE, then*

$$\mu \in \arg\max_{\mu'} |FM_{\mu'}|$$
$$s.t. \quad (1-\psi)(\overline{q} - u_{i'}) \geq c_j + \tau \quad and \quad (1-\psi)(\overline{q} - u_i) \geq c_{j'} + \tau \quad \forall (s_j, b_i), (s_{j'}, b_{i'}) \in FM_{\mu'}$$

Proof: Proof in the Appendix. ■

In this environment we can easily find examples in which the platform has no incentives to implement a *CPE* matching.

Example 1. *(CPE matchings do not always maximize profits) For a given commission fee, $\psi \in [0, 1]$ and $\tau \geq 0$, the optimal match μ^* chosen by the platform is not necessarily CPE.*

Indeed, consider a market where the set of sellers is given by $\{s_1, s_2, s_3, s_4\}$, and the set of buyers is given by $\{b_1, b_2, b_3, b_4\}$. Suppose that $q_{ij} = 20 \ \forall i, j \leq 4$, and that sellers' costs are given by

$$(c_1, c_2, c_3, c_4) = (2, 8, 14, 16),$$

while customers' outside options are given by

$$(u_1, u_2, u_3, u_4) = (0, 3, 5, 7).$$

In a Pareto Efficient allocation, buyers b_1, b_2 and b_3 should be each matched with one of the first three sellers with highest net quality (i.e., with sellers s_1, s_2 and s_3). But if the platform chooses $(\psi, \tau) = (1/2, 0)$, then all of the Pareto efficient matchings would generate a total profit of 9.75 for the platform, whereas if the platform implemented the matching

$$s_1 : b_4, \quad s_2 : b_3, \quad s_3 : b_2, \quad s_4 : b_1$$

its profits would equal 12.25.

Intuitively, the reason why the platform may not want to implement a *CPE* matching can be explained as follows. We can write the platform's revenue for a match between a seller s_j and a customer b_i such that $(1-\psi)(\overline{q} - u_i) \geq c_j + \tau$, as

$$\psi(1-\alpha)(\overline{q} - c_j - \tau - u_i) + \frac{\psi(1 - \psi + \alpha\psi)}{1-\psi}(c_j + \tau) + \tau. \tag{5}$$

Though the first term from expression 5 is increasing in the total surplus created from the transaction between seller s_j and buyer b_i, $(\overline{q} - c_j - u_i)$, the sum of the last two terms from this expression is increasing in the opportunity cost from selling, c_j, and on τ. So the platform is not exclusively interested in maximizing the creation of value, but also in maximizing the total number of transactions so as to collect more revenues through τ, and also on facilitating the transaction of more expensive products, as that increases the final price charged by the seller, and therefore the commissions collected through ψ.

So now we provide an algorithm for finding the profit-maximizing matching μ, taking the commission fees (ψ, τ) as given. This algorithm is the same as the one presented in Boerner and Quint (2022) to find the matching that maximizes the platforms' profits: we start by matching the "worst seller" (the one with highest cost) with the "best customer" (the one with lowest outside option) and keep doing this iteratively with the agents who have not been matched yet.

Algorithm 1. *For each $B' \subseteq B$, each $s_j \in S$ and each $\psi \in [0,1]$ and $\tau \geq 0$ define*

$$F(s_j, B', \psi, \tau) \equiv \{b_i \in B'; (1-\psi)(q_{ij} - u_i) \geq c_j + \tau\},$$

i.e., $F(s_j, B', \psi, \tau)$ is the set of "feasible buyers" in B' for seller s_j, or more precisely, it is the set of buyers in B' who, if matched with seller s_j, would end up purchasing their product.

i) Initialize $j = m$ and define $B_j = B$.

ii) Compute $F(s_j, B_j, \psi, \tau)$. If $F(s_j, B_j, \psi, \tau) \neq \emptyset$, set $\mu(s_j)$ equal to $b_i \in F(s_j, B_j, \psi, \tau)$ such that $u_i \leq u_l$ for all $u_l \in F(s_j, B_j, \psi, \tau)$ (i.e., match this seller with the buyer with lowest outside option who has not been matched yet) and define $B_{j-1} = B_j \backslash \{b_i\}$, else set $\mu(s_j) = s_j$ (i.e., keep seller s_j unmatched) and define $B_{j-1} = B_j$. Then proceed to the next step.

iii) If $j > 1$, redefine $j = j - 1$ and repeat step ii), else, stop the algorithm.

To some extent, algorithm 1 does the opposite of what *PAM* does. Indeed, while *PAM* attempts to secure good matches to low-cost sellers, algorithm 1 attempts to secure good matches to high-costs sellers. The platform has incentives to adopt such an algorithm for two reasons: 1) high-cost sellers charge higher prices, so they generate more revenues to the platform in the event they transact (assuming $\psi > 0$); 2) moreover, this procedure generates more transactions (and therefore more revenues), as it frees low-cost sellers to transact with buyers who would otherwise not transact if matched with a high-cost seller.

Theorem 1. *(Algorithm 1 maximizes profits and number of transactions) Suppose that $q_{ij} = \overline{q} \in \mathbb{R}_+$ for all $i \leq n$ and all $j \leq m$. Then, given $\psi \in [0,1]$ and $\tau \geq 0$, the matching allocation obtained from algorithm 1:*

i) Maximizes the number of transactions that can be generated in the economy, i.e., it maximizes $|S_\mu|$.

ii) Maximizes the profits of the platform, i.e., it solves the maximization problem 4.

Proof: In the Appendix. ■

So we conclude that, if the products sold are homogeneous, the platform will implement algorithm 1, which does not always yield a constrained Pareto efficient allocation (see example 1).

As a corollary to theorem 1, we can find a necessary and sufficient condition under which the platform's optimal matching obtained through algorithm 1 is not *CPE*. Suppose that all sellers have at least one "*feasible buyer*", i.e., for every seller s_j there is at least one buyer c_i such that $(1-\psi)(\overline{q} - c_j) \geq u_i + \tau$. This hypothesis is without loss of generality, as the platform would never have incentives to match sellers who do not have a feasible counterpart with whom they would be willing to transact. So adding or removing those sellers from the model would not affect the platform's optimal choices or profitability. Given this hypothesis, we have that, a necessary and sufficient condition for the matching obtained through algorithm 1 to be constrained Pareto *inefficient* is that there is at least one seller such that, there is a buyer with the same index as the seller, and such a buyer is *not* willing to purchase from this seller; or there are more sellers than buyers. The hypothesis that there is at least one seller who is not willing to transact with a buyer that shares the seller's index can be interpreted as requiring that the "demand crosses the supply curve".

Corollary 3. *(The platform's optimal matching is not CPE iff "demand crosses supply" or there are more sellers than buyers) For each seller $s_j \in S$ let*

$$F(s_j, \psi, \tau) \equiv \{b_i \in B; (1-\psi)(q_{ij} - c_j) \geq u_i + \tau\},$$

i.e., $F(s_j, \psi, \tau)$ is the set of buyers, who, if matched with seller s_j, would end up purchasing their product. Without loss of generality, suppose that $F(s_j, \psi, \tau) \neq \emptyset$ for all $s_j \in S$ (i.e., all sellers have at least one "feasible buyer"). Suppose that $q_{ij} = \overline{q} \in \mathbb{R}_+$ for all $i \leq n$ and all $j \leq m$. In addition, suppose that $c_1 < c_m$. Then the matching obtained through algorithm 1 will not be CPE if and only if either one of the following conditions hold:

I) ("Demand crosses supply"): There is at least one $s_j \in S$, with $j < |B|$ such that $u_j \notin F(s_j, \psi, \tau)$, or

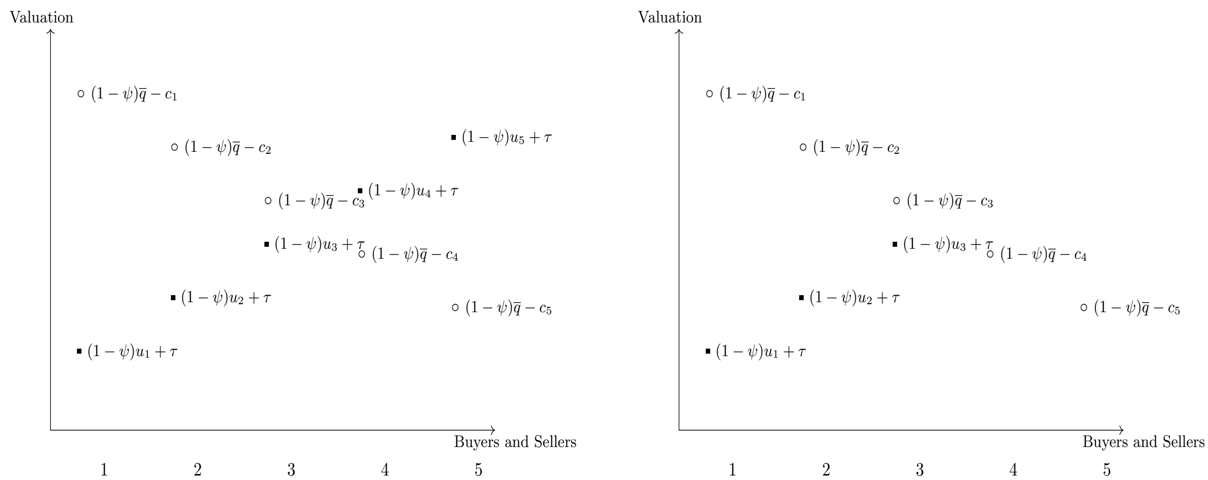
II) (There are more sellers than buyers): $|S| \geq |B|$.

Proof: *Case 1:* Suppose that there is at least one seller $s_j \in S$, such that $j < |B|$ and $u_j \notin F(s_j, \psi, \tau)$ i.e., there is a buyer with the same index as seller s_j, and this buyer would not willing to purchase from this seller. From proposition 1, in a *CPE* matching, a seller $s_{j'}$ with $j' > j$ should not transact. In particular, seller s_m will not transact under a *CPE* matching. But because $F(s_m, \psi, \tau) \neq \emptyset$, we have that seller s_m transacts under algorithm 1, so that the matching obtained through this algorithm is not *CPE*.

Case 2: Suppose that there are more sellers than buyers. Then, from proposition 1, seller s_1 transacts under a *CPE* matching, but s_m does not. But because algorithm 1 starts by matching seller s_m, the one with lowest willingness to sell, and because $F(s_m, \psi, \tau) \neq \emptyset$, seller s_m will transact under algorithm 1. ■

Figure 3 presents instances in which the conditions from corollary 3 hold, whereas figure 4 presents instances in which those conditions are not met. Intuitively, if the "supply curve crosses the demand", implementing *CPE* matchings such as *PAM* would result in sellers with high costs remaining unmatched (because of proposition 1). But the platform could be made better off by giving away the customers who were matched to low cost sellers to the unmatched sellers with high costs, as that would increase the prices paid by these consumers, and therefore, the commissions received by the platform. Doing this replacement would also free the low cost sellers to potentially form matches with buyers with high outside options, thus increasing the number of transactions, and therefore, the fees collected by the platform. Similarly, in a market with excess demand, most sellers will be able to transact regardless of how they are matched with the "most desirable" buyers (i.e., the buyers with lowest outside options), so the platform is less likely to have incentives to implement an inefficient match. But in a market with excess supply, the platform has to choose carefully how to match sellers, as it can no longer get a transaction from all of them. In this case, the platform will have incentives to prioritize matching the inefficient sellers first (i.e., the ones with higher opportunity costs) in order to inflate equilibrium prices.

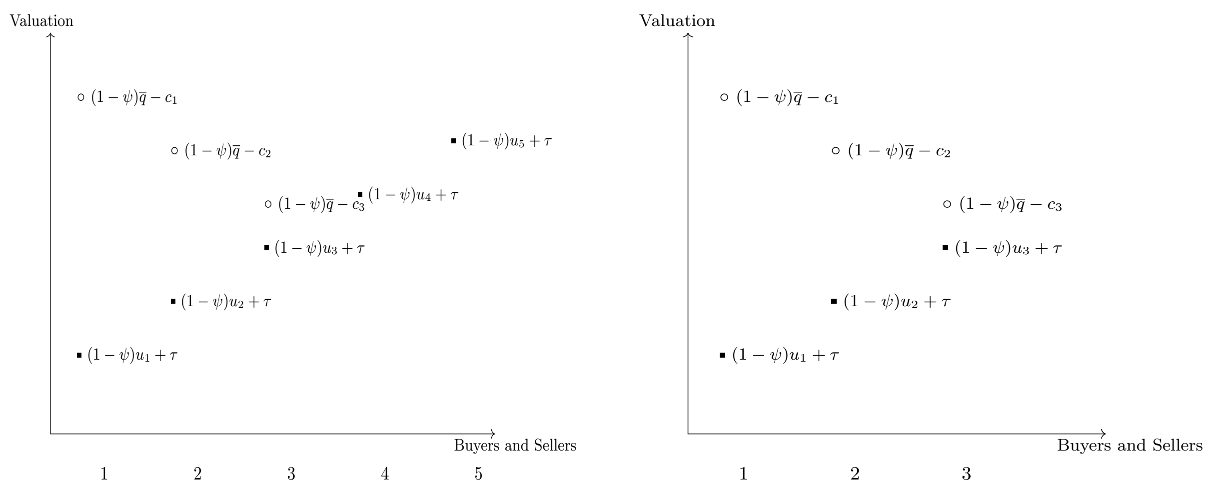
Notice that, when it comes to the transaction of digital goods, such as apps or movies from streaming services, supply always exceeds demand, as each seller can sell virtually infinite units of its product. So



(a) A case in which "*demand crosses supply*".

(b) A case with more sellers than buyers.

Figure 3: Instances in which the profit-maximizing matching obtained through algorithm 1 is *not CPE*.



(a) A case with more buyers than sellers, and the "*demand does not cross the supply*".

(b) A case in which there are as many buyers as sellers but the "demand does not cross the supply".

Figure 4: Instances in which the profit-maximizing matching obtained through algorithm 1 *is CPE*.

in these instances the platform is expected to implement a matching that is not CPE. This perhaps may explain why streaming services such as Amazon Prime also recommend paid movies to its subscribers, as illustrated in figure 5 below, even though customers usually prioritize free content.

4.2 When products are vertically differentiated

Throughout this section we assume that $q_{ij} = q_j \in \mathbb{R}_+$ for all $i \leq n$ and all $j \leq m$. Moreover, instead of ordering sellers in ascending order of their cost, we order sellers in descending order of the difference between their quality and cost, i.e., we assume that

$$q_1 - c_1 \geq q_2 - c_2 \geq \cdots \geq q_m - c_m.$$

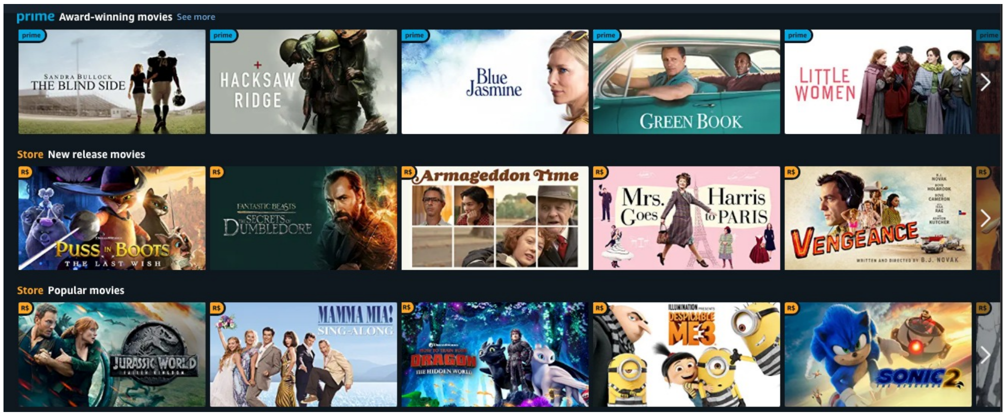


Figure 5: In this example, the streaming platform recommends paid products (the ones with yellow tag), even though subscribers probably prioritize free content (the ones with a blue tag).

In this case, *PAM* would be defined as matching buyer b_1 with seller s_1, buyer b_2 with seller s_2, and so on. With this definition, one can trivially extend the results from proposition 1 and corollaries 1 and 2, to the case in which products are vertically differentiated. So under vertical differentiation, *PAM* maximizes the surplus of buyers and sellers.

Proposition 2. *(Necessary and sufficient condition for μ to be CPE) Suppose that $q_{ij} = q_j \in \mathbb{R}_+$ for all $i \leq n$ and all $j \leq m$. A matching μ is Constrained Pareto Efficient (CPE) if and only if each seller s_j such that $j \leq \min\{n, m\}$ and $(1-\psi)(q_j - c_j) \geq u_j + \tau$ is matched to a buyer b_i such that $(1-\psi)(q_i - c_i) \geq u_i + \tau$.*

Proof: Analogous to the proof of proposition 1. ■

Corollary 4. *(PAM is CPE) Suppose that $q_{ij} = q_j \in \mathbb{R}_+$ for all $i \leq n$ and all $j \leq m$. If $\mu(s_i) = b_i$ for all $i \leq \min\{n, m\}$, then μ is CPE.*

Corollary 5. *(CPE matchings maximize the number of transactions subject to a constraint) Suppose that $q_{ij} = q_j \in \mathbb{R}_+$ for all $i \leq n$ and all $j \leq m$. If a matching μ is CPE, then*

$$\mu \in \arg\max_{\mu'} |FM_{\mu'}|$$
$$s.t. \quad (1-\psi)(q_j - u_{i'}) \geq c_j + \tau \quad and \quad (1-\psi)(q_j - u_i) \geq c_{j'} + \tau \quad \forall (s_j, b_i), (s_{j'}, b_{i'}) \in FM_{\mu'}$$

Proof: Analogous to the proof of corollary 2. ■

As shown in the previous section, *CPE* matchings (such as *PAM*) are not necessarily profit maximizing. But one can find instances in which the platform has incentives to implement a *CPE* matching. This happens if $\tau = 0$ and either: 1) all sellers have zero opportunity cost (i.e., if $c_j = 0$ for all $j \in \{1, 2, \cdots, m\}$), or 2) all sellers have the same opportunity cost (i.e., if $c_j = c \in \mathbb{R}$ for all $j \in \{1, 2, \cdots, m\}$), and they have all the bargaining power (i.e., $\alpha = 0$). Indeed, in the former case, the revenue the platform gets from a transaction between seller s_j and buyer b_i is given by

$$\psi(1-\alpha)(q_j - u_i), \tag{6}$$

whereas in the latter case, this expression reduces to

$$\psi(q_j - u_i),$$

so that in both cases the platform's revenues are proportional to the sum of value created form transactions.

Proposition 3. *(Sufficient conditions for PAM to be profit-maximizing) Suppose that $\tau = 0$, and suppose that either $c_j = 0$ for all $j \in \{1, 2, \cdots, m\}$, or $\alpha = 0$ and $c_j = c \in \mathbb{R}_+$ for all $j \in \{1, 2, \cdots, m\}$. Then, for any given $\psi \in [0, 1]$, PAM maximizes the profits of the platform.*

Proof: Proof in the Appendix. ■

Intuitively, the first part of proposition 3 holds because, if costs are zero and $\tau = 0$, the fee ψ charged by the platform is non-distortionary (i.e., it does not affect agents' decisions), in which case maximizing the creation of value allows the platform to extract more surplus. In practice, we could have $c_j = 0$ for all sellers if costs are sunk and sellers have no other means of selling their products but through the platform (e.g., sellers trying to resell tickets for concerts or sport events that they can no longer attend to, and having no other means of reselling their tickets).

Similarly, if all the sellers have the same opportunity cost, the ones with higher quality products are also the ones who are capable of generating more revenues to the platform. So the platform must ensure those sellers end up transacting in equilibrium (i.e., that the "most efficient" sellers end up transacting). Moreover, $\tau = 0$ and $\alpha = 0$ removes the platform's incentives to generate more transactions than the socially optimum, as it causes profits to be proportional to the total surplus created from transactions.

For most practical applications, however, we do not expect the hypotheses from proposition 3 to hold, so that *PAM* will usually not be profit-maximizing. Similarly, under vertical differentiation, algorithm 1 is also not guaranteed to yield a profit-maximizing outcome (see example 1, from the appendix). If, however, we were to assume that the sellers with less feasible matches (i.e., the sellers for which it is hardest to find a buyer who would purchase their product) were also the ones who, in the event they transacted, they would charge the highest prices (i.e., if the "more expensive" products were also the ones with less feasible buyers), then algorithm 1 would yield a profit-maximizing outcome.

Proposition 4. *(Sufficient conditions for algorithm 1 to maximize profits and number of transactions) Suppose that $q_{ij} = q_j \in \mathbb{R}_+$ for all $i \leq n$ and all $j \leq m$. Also suppose that sellers with higher index have less feasible matches:*

$$(1-\psi)q_1 - c_1 \geq (1-\psi)q_2 - c_2 \geq \cdots \geq (1-\psi)q_m - c_m,$$

and charge higher prices in the event they transact:

$$(1-\alpha)q_1 + \alpha\frac{c_1}{1-\psi} \leq (1-\alpha)q_2 + \alpha\frac{c_2}{1-\psi} \leq \cdots \leq (1-\alpha)q_m + \alpha\frac{c_m}{1-\psi}.$$

Then, given $\psi \in [0, 1)$ and $\tau \geq 0$, the matching allocation obtained from algorithm 1:

i) Maximizes the number of transactions that can be generated in the economy, i.e., it maximizes $|S_\mu|$.

ii) Maximizes the profits of the platform, i.e., it solves the maximization problem 4.

Proof: The proof is analogous to the proof of theorem 1, as both proofs only rely on the fact that $p_j(s_i, \psi, \tau) \geq p_{j'}(s_i, \psi, \tau)$ for all $j \geq j'$ (i.e., sellers with higher index charge higher prices in the event they transact) and $F(s_j, B, \psi, \tau) \subseteq F(s_{j'}, B, \psi, \tau)$ for all $j \geq j'$ (i.e., sellers with higher index have less feasible matches). ∎

Intuitively, proposition 4 holds because it makes sense to prioritize matching sellers who face lower demand when they are also the ones who generate more revenues in the event they transact.

Another instance in which algorithm 1 can be used to maximize the platform's revenues is when the platform only charges a flat fee $\tau > 0$ and sets $\psi = 0$. Indeed, in this case, the platform's sole objective is to maximize the total number of transactions, and the maximization of total number of transactions can be achieved if the platform prioritizes securing good matches to sellers who have less feasible buyers.

Proposition 5. *(Algorithm 1 maximizes profits when $\psi = 0$) Suppose that $q_{ij} = q_j \in \mathbb{R}_+$ for all $i \leq n$ and all $j \leq m$. Also suppose that sellers with higher index have less feasible matches:*

$$q_1 - c_1 \geq q_2 - c_2 \geq \cdots \geq q_m - c_m.$$

Then, for a given $\tau > 0$ and $\psi = 0$, the matching allocation obtained from algorithm 1 maximizes the number of transactions that can be generated in the economy, i.e., it maximizes $|S_\mu|$, which implies that it maximizes the profits of the platform (i.e., it solves the maximization problem 4).

Proof: The proof is analogous to the first part of the proof of theorem 1, as both proofs only rely on the fact that $F(s_j, B, \psi, \tau) \subseteq F(s_{j'}, B, \psi, \tau)$ for all $j \geq j'$ (i.e., sellers with higher index have less feasible matches). ∎

In practice, however, the conditions for propositions 4 and 5 to hold are rather strong. Indeed, proposition 4 requires that more expensive products are also the ones with lower demand, a condition that does not necessarily hold in practical applications. As to proposition 5, it requires that the platform only charges a flat fee per transaction and sets $\psi = 0$, which is inconsistent with the commission pricing criteria adopted in most two-sided markets. Moreover, both results do not apply to cases in which products are horizontally differentiated. In the next section we show that, for any configuration of products' quality and costs, the platform can implement the *Hungarian method*, instead, to either find a *CPE* matching or a profit-maximizing one.

4.3 When products are horizontally or vertically differentiated

When products are horizontally differentiated, it is no longer clear what Positive Assortative Matching means, as now each seller s_j has a vector of attributes $(q_{1j}, q_{2j}, \cdots, q_{nj})$, so that it is no longer clear how they should be ordered. If we were to ignore those attributes and simply ordered sellers in ascending order of their costs, i.e., if we were to assume that

$$c_1 \leq c_2 \leq \cdots \leq c_m,$$

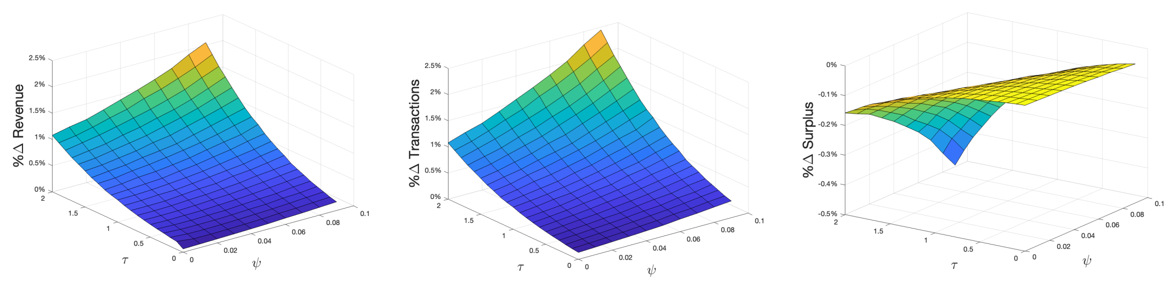
and defined *PAM* as matching buyer b_1 with seller s_1, buyer b_2 with seller s_2, and so on; then *PAM* would not necessarily be *CPE* (see example 1, from the appendix). This is because, under product differentiation, the sellers with lowest costs are not necessarily the "most efficient ones", as they may also have an average lower quality. Similarly, algorithm 1 is not guaranteed to generate a profit-maximizing outcome (see example 1 from the Appendix). But the platform can rely on the *Hungarian method*, instead, to find a *CPE* matching, as well as a profit-maximizing matching. Through this method, we are able to simulate the differences in total revenue and total number of transactions obtained when the platform implements a *CPE* matching vs. a profit-maximizing one. Details on how this algorithm works can be found in section C from the Appendix.

Our simulations indicate that *CPE* matchings can be much less profitable than the platform's optimum policy. Moreover, consistent with the theoretical results obtained for the case in which products are homogeneous, profit-maximizing matchings usually generate more transactions than *CPE* ones.

In the following simulations, we assume that costs and outside options are iid, and follow a normal distribution. More precisely, we assume that $c_j \sim N(10, 1)$ and $u_i \sim N(15, 1)$ for all $j \leq m$ and all $i \leq n$. Quality is proportional to costs, and we assume that

$$q_{ij} = \xi_{ij} + \beta c_j,$$

where $\xi_{ij} \sim N(0, 1)$ is an iid shock, and $\beta \geq 0$ is a parameter measuring how quality correlates with production costs. To implement the simulations using the Hungarian method, we used the *R* package "*RcppHungarian*" (Silverman et al. (2022)).



(a) Difference in revenues between a proft-maximizing and a *CPE* matching

(b) Difference in number of transactions between a proft-maximizing and a *CPE* matching

(c) Difference in Buyer and Seller surplus between a proft-maximizing and a *CPE* matching

Figure 6: Plots depicting the average percentage difference in profits, number of transactions and surplus between a profit-maximizing and a *CPE* matching for different levels of ψ and τ, assuming that $\beta = 3$ and that $n = m = 20$. In total, $100,000$ simulations were conducted for each combination of ψ and τ from the grid. The percentage differences were computed by taking the difference between the average of the variable of interest from the profit maximizing matching and the one from the *CPE* matching, and dividing this difference by the average obtained from the *CPE* matching.

Figure 6 depict the percentage difference between the average revenues, average number of transactions, and average surplus from buyers and sellers when implementing a profit-maximizing matching vs. a *CPE* one, when $\beta = 3$ and there are exactly 20 buyers and 20 sellers in the market. As the figure illustrate, those differences can be quite significant.

5 Optimal commissions and optimal matching

In the previous sections we assumed the commission fees ψ and τ to be exogenous. But in reality one would expect commissions to be set by the platform along with the matching function μ. In this case, the platform would choose ψ, τ and μ to maximize

$$\max_{\mu,\psi,\tau} \tau |S_{\mu,\psi,\tau}| + \sum_{s_j \in S_{\mu,\psi,\tau}} \psi p_j^*(\mu(s_j), \psi, \tau), \tag{7}$$

where

$$S_{\mu,\psi,\tau} \equiv \{s_j \in S; \mu(s_j) = b_i \in B \wedge (1-\psi)(q_{ij} - u_i) \geq c_j + \tau\}$$

corresponds to the set of sellers who end up transacting given the fees (ψ, τ) and matching function μ, and $p_j^*(\mu(s_j), \psi, \tau)$ is the equilibrium price given by equation 3.

It can be shown that a maximum to this problem always exists. Indeed, for a given matching μ, we can show that the platform's objective function is upper semicontinuous in $(\psi, \tau) \in [0,1] \times \mathbb{R}_+$. Moreover, we can limit τ to be less than or equal to a constant $\overline{\tau}$, since, for τ sufficiently high, no transactions are made, so the platform earns zero profits. Because, for a given μ, the objective function is upper semicontinuous, and because we only consider (ψ, τ) defined on the compact set $[0,1] \times [0, \overline{\tau}]$, it then follows that the problem has a maximum for a given μ, since every upper semicontinuous function defined on a compact set has a maximum (Leininger (1984)). Because the set of possible matchings is finite, we conclude that a maximum to problem 7 exists.

Proposition 6. *There exists a pair of commission rates (ψ, τ) and a matching function μ that maximizes the profits of the platform.*

Proof: In the Appendix. ■

The next lemma states that, under an optimal matching, at least one agent must be exactly indifferent between transacting and not transacting, otherwise, the platform would be able to increase its fees without reducing the number of transactions. In order to state the lemma more formally, let $\mu^*_{\psi,\tau}$ be the a matching that maximizes revenues for a given (ψ, τ). For a given matching μ, let

$$M_\mu(\psi, \tau) \equiv \{(s_j, b_i) \in S \times B; \mu(s_j) = b_i \wedge (1-\psi)(q_{ij} - u_i) \geq c_j + \tau\},$$

i.e., $M_\mu(\psi, \tau)$ represents all the matches associated with μ, excluding those in which agents do not transact. When (ψ, τ) are known and exogenous, we sometimes write M_μ instead of $M_\mu(\psi, \tau)$ to simplify the notation.

Lemma 1. *If $(\mu^*_{\psi,\tau}, \psi, \tau)$ is a solution to the platform's objective function 7, we must have $(1-\psi)(q_{ij} - u_i) = c_j + \tau$ for at least one $(s_j, b_i) \in M_{\mu^*_{\psi,\tau}}(\psi, \tau)$.*

Proof: If all agents who transact were strictly better off transacting as opposed to not transacting, the platform would be able to marginally increase ψ or τ, thus increasing its revenues per transactions, without reducing the number of transactions. ■

The next lemma states that, if products have homogenous quality and, for a pair of commission rates (ψ, τ), a matching μ^* maximizes profits, and, moreover, all transactions from this matching generate a positive surplus to buyers and sellers, then, if we were to marginally increment τ, the matching μ^* would still be profit-maximizing.

Lemma 2. *Suppose that $q_{ij} = \overline{q}$ for all $i \leq n$ and all $j \leq m$. Let $\mu^*_{\psi,\tau}$ be the profit-maximizing matching obtained through algorithm 1 when commissions are given by (ψ, τ). Suppose, in addition, that $\tau < (1-\psi)(\overline{q} - u_i) - c_j$ for all $(s_j, b_i) \in M_{\mu^*_{\psi,\tau}}(\psi, \tau)$. Then, for any $\tau' \geq \tau$ such that $\tau' \leq \min_{(s_j,b_i) \in M_{\mu^*_{\psi,\tau}}(\psi,\tau)}[(1-\psi)(q_{ij} - u_i) - c_j]$, we have that $\mu^*_{\psi,\tau'} = \mu^*_{\psi,\tau}$.*

Proof: In the Appendix. ■

If we assume that ψ is set exogenously and that $q_{ij} = \overline{q}$ for all $i \leq n$ and all $j \leq m$, we can use theorem 1 and lemmas 1 and 2 to build an algorithm to find the optimal $\tau \geq 0$ and optimal matching $\mu^*_{\psi,\tau}$. The algorithm basically finds all the points in $\tau \geq 0$ such that, after implementing the matching $\mu^*_{\psi,\tau}$, at least one of the sellers who transacts under this matching is indifferent between transacting and not transacting, as those points will be our only candidates for an optimum.

Algorithm 2. *Let $\mu^*_{\psi,\tau}$ be the matching algorithm that maximizes the platform's profits when the fees are given by (ψ, τ). For a given $\psi \in [0, 1)$, perform the following algorithm.*

i) Initialize $k = 0$ and $\tau = 0$.

ii) Compute $\mu^*_{\psi,\tau}$. If $M_{\mu^*_{\psi,\tau}} \neq \emptyset$, proceed to the next step, else, stop the algorithm.

iii) Set $k = k + 1$. Define

$$\tau_k \equiv \min_{(s_j,b_i)\in M_{\mu^*_{\psi,\tau}}} (1-\psi)(q_{ij} - u_i) - c_j$$

If $\tau_k > \tau$, redefine $\tau = \tau_k$, else (if $\tau_k = \tau$), redefine $\tau = \tau_k + \varepsilon$, where ε is arbitrarily small. Then, repeat step ii).

If $k = 0$, then any $\psi \in [0, 1]$ and any matching μ generates zero profits to the platform. If $k > 0$, choose the commission

$$\tau^* = \operatorname*{arg\,max}_{\tau\in\{\tau_1,\tau_2,\cdots,\tau_k\}} \pi(\mu^*_{\psi,\tau}, \psi, \tau)$$

and the matching μ^*_{ψ,τ^*}.

Theorem 2. *Suppose that $q_{ij} = \overline{q}$ for all $i \leq n$ and all $j \leq m$. For a given $\psi \in [0, 1)$, let τ^* and μ^*_{ψ,τ^*} be the flat fee and matching function, respectively, obtained after implementing algorithm 2. Define $S_\mu(\psi, \tau) = \{s_j \in S; \mu(s_j) = b_i \in B \wedge (1 - \psi)(\overline{q} - u_i) \geq c_j + \tau\}$. Then,*

$$(\mu^*_{\psi,\tau^*}, \tau^*) = \operatorname*{arg\,max}_{\mu,\tau} \tau|S_\mu(\psi, \tau)| + \sum_{s_j\in S_\mu(\psi,\tau)} \psi p^*_j(\mu(s_j), \psi, \tau).$$

Proof: The proof follows immediately from theorem 1 and lemmas 1 and 2. ∎

Though algorithm 2 does not simultaneously find the optimal ψ and optimal τ, one could create a grid for $\psi \in [0, 1)$, and then perform algorithm 2 for each ψ in the grid, and then select the ψ from the grid that generates the maximum revenue for the platform.

Notice that algorithm 2 does not necessarily yield the optimum policy for the platform if products are differentiated, as theorem 2 only applies to the case in which $q_{ij} = \overline{q}$ for all $i \leq n$ and all $j \leq m$. As an alternative, one could simply create a grid for both ψ and τ, and find the optimum match for each combination of (ψ, τ) within the grid. The advantage of using algorithm 2, however, is that it makes it more likely that a necessary condition for optimization is met: that at least one pair of buyer and seller is exactly indifferent between transacting and not transacting. Moreover, through this method one does not need to specify an upper bound for the values that τ can assume (notice that τ can be any positive number).

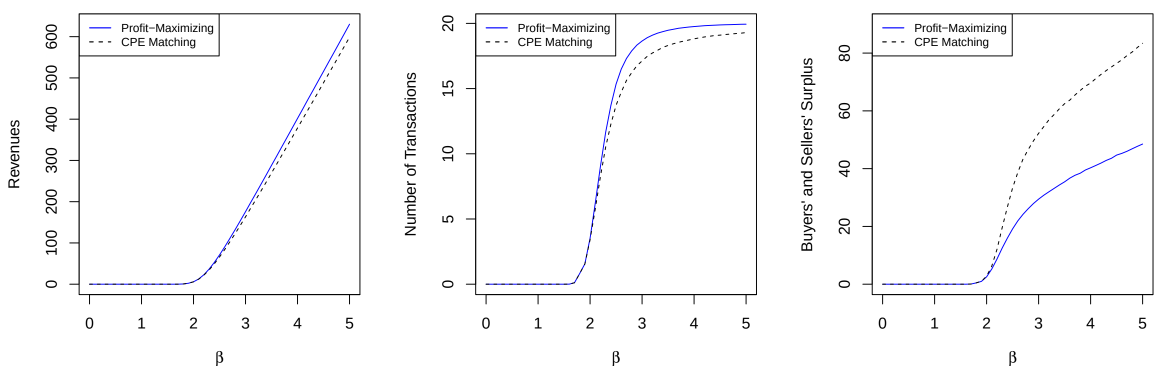
With these results we can use the Hungarian method to simulate the platform's optimal matching and fees (ψ, τ). We can then compare the platform's revenues and number of transactions when implementing the optimal policy, vs. the case in which it seeks to maximize revenues, subject to the constraint that the matching must be *CPE*.

Similar to the previous section, we assume that $c_j \sim N(10, 1)$ and $u_i \sim N(15, 1)$ for all $j \leq m$ and all $i \leq n$. Moreover,

$$q_{ij} = \xi_{ij} + \beta c_j,$$

where $\xi_{ij} \sim N(0, 1)$ and $\beta \geq 0$. The parameter β captures the extend to which quality is correlated with production costs.

Figure 7 shows the differences in revenues, number of transactions and sum of buyers' and sellers' surplus obtained when the platform attempts to implement its optimum policy by following algorithm 2 vs. the case in which it attempts to implement profit-maximizing commissions subject to the constraint that the matching must be *CPE*, by also following algorithm 2 but with the difference that, at each step of the algorithm, it implements a matching that maximizes the sum of customers' and sellers' surplus, not the platform's profits. The simulations suggest that those differences can be quite significant and they do not disappear for arbitrarily high values of β. This happens because, as β increases, the platform also increases its fees, until it reaches a point in which "demand crosses supply", which, from corollary 3, would imply that the platform has strict incentives to implement a profit-maximizing matching, as opposed to a *CPE* matching. Though corollary 3 only applies to the case in which products have homogeneous quality, our simulations indicate that the intuition of this result seems to extend to cases in which products are differentiated.



(a) Difference in revenues between a proft-maximizing and a *CPE* matching

(b) Difference in number of transactions between a proft-maximizing and a *CPE* matching

(c) Difference in Buyer and Seller surplus between a proft-maximizing and a *CPE* matching

Figure 7: Plots depicting simulated differences between a profit-maximizing and a *CPE* matching for different levels of β, assuming that $n = m = 20$ (so there can be no more than 20 transactions in total). 10,000 simulations were used to compute the averages from each β in the grid.

6 Stability

In this section we analyze the performance of Positive Assortative Matching (*PAM*) in terms of *stability* when $\psi = 0$ and $\tau \geq 0$. Our notion of stability is similar to the one studied in benchmark models of matching theory (e.g., Roth and Sotomayor (1989)), with the difference that in our case matched agents can negotiate their division of surplus (such as in Kelso and Crawford (1982)), and they must pay a search cost $r \geq 0$ to form a blocking coalition. We show that, provided that $\psi = 0$, PAM is always *stable* in the sense that no buyer is willing to pay the search cost $r \geq 0$ to form a *blocking coalition*. Stability is a desirable property in matching markets, as it implies agents have incentives to follow through the recommendations made by the matchmaker (Roth and Sotomayor (1990)), in our case, the platform. We show that, in our environment, when agents do not have incentives to follow through those recommendations, the overall number of transactions can diminish (see example 2). As the matching obtained through algorithm 1 is not necessarily stable, the platform may have incentives to implement PAM (which is stable and CPE) as opposed to a (Constrained Pareto *inefficient*) matching obtained through algorithm 1. Because all matches are stable if the search cost r is sufficiently high, this result indicates that the platform will have more incentives to implement a CPE match if it believes agents can easily find matches better than the ones recommended by the platform, e.g., by scrolling down the page of search results, or by doing a more refined search.

Suppose that the platform charges commission fees (ψ, τ) and implements the matching μ. Suppose that, after the platform implements this matching, each buyer $b_i \in B$ can pay a search cost $r \geq 0$ to be matched with a different seller in the market. If, during this second round, buyer b_i chooses to pay θ to be matched with a seller $s_j \neq \mu(s_j)$, seller s_j will then choose whether to transact with its current match $\mu(s_j)$, or transact with buyer b_i.

We could, in principle, also allow agents to pay an additional cost θ to form transactions outside the platform, so as to avoid paying the platform's commission τ. As a practical example, as we write this in 2022, Airbnb.com does not pinpoint the precise location of advertised houses before booking, so as to prevent customers from negotiating directly with the hosts without paying the platform its commissions. But if buyers are willing to engage in a search cost to find those hosts independently, they may be able to get a better deal as, in this case, agents do not have to share their surplus with the platform. The same occurs in other platforms such as eBay, which explicitly prohibits sellers from offering or referencing their personal contact information to prevent them from selling outside of the platform without paying their transaction fees.[3] Similarly, until recently Apple did not allow app developers selling digital products through the App Store to direct users to external payment methods to circumvent Apple's fees. But in 2021, after a lengthy legal battle between Apple and Epic Games, a ruling from the Northern District of California has required Apple lift those restrictions.[4]

But notice that even if platforms allowed buyers to perfectly pinpoint the identity of sellers, forming matches on their own may be contractually costly and risky, which would also be captured by θ. As an example, making a reservation through Airbnb is relatively straightforward, and both buyers and sellers are given incentives to fulfill their contractual obligations through the platform's rating system. If a customer decides to form a match without the intermediation of Airbnb, both parties may have to spend extra resources forming and agreeing to a binding contract, and they may also have to spend more on litigation in the eventuality one of them does not fulfill their side of the bargain, whereas on Airbnb unsatisfied users can simply leave a negative review in retaliation, or report the issue to Airbnb.

So throughout this section we will assume that the cost of forming transactions outside the platform, θ, is greater than τ, so that no agent has incentives to avoid paying the platforms' commissions by transacting outside of the platform. Buyers, however, may still have incentives to incur in a search cost r to seek better matches, such as by scrolling down the page of search results. We obtain similar qualitative results if transacting outside the platform was feasible and that was the only way agents could circumvent the recommendations made by the platform. In fact, in this case, the negative effects of instability to the platform would tend to be greater, as instability would result in the platform not being able to collect fees

[3] E.g., see https://pages.ebay.com/sponsored/intuit/articles/selling-fees.html

[4] E.g., see https://www.theverge.com/2021/12/8/22814147/epic-apple-app-store-injunction-paused

from transactions that did not follow its recommendation (see section D from the appendix for a thorough analysis).

We will also assume that sellers are ordered according to their "net quality" $(q_j - c_j)$ as follows

$$q_1 - c_1 \geq q_2 - c_2 \geq \cdots \geq q_m - c_m,$$

and we define PAM as the matching that assigns each seller s_j with the buyer with the same index, b_j, for all $j \leq \min\{m, n\}$.

We say a matching μ is *stable* if no buyer has incentives to search for a match different than the one recommended by the platform.

Definition 3. *(Stability) Given $\psi = 0$ and $\tau \geq 0$, and given the search cost r, a matching μ is stable if, for every $s_j \in S$ there is a payoff $w_{s_j} \in \mathbb{R}$ and for every $b_i \in B$ there is a payoff $w_{b_i} \in \mathbb{R}$ such that:*

i) (Individual rationality) $w_{s_j} \geq c_j \; \forall s_j \in S$ and $w_{b_i} \geq u_i \; \forall b_i \in B$ (i.e., agents' payoffs are greater than or equal to the their outside option). If $\mu(s_j) = s_j$, then $w_{s_j} = c_j$, and if $\mu(b_i) = b_i$, then $w_{b_i} = u_i$ (i.e., unmatched agents get their outside option).

ii) (Feasibility) If $\mu(s_j) = b_l \in B$, then $w_{b_l} + w_{s_j} \leq q_i - \tau$ if $w_{s_j} > c_j$ or $w_{b_l} > u_i$,

and there is no $(s_j, b_i) \in S \times B$ and no $(\tilde{w}_{s_j}, \tilde{w}_{b_i}) \in \mathbb{R}^2$ such that:

I) (Feasibility) $\tilde{w}_{s_j} + \tilde{w}_{b_i} \leq q_i - \tau - r$

II) (Blocking Coalition) $\tilde{w}_{s_j} > w_{s_j}$ and $\tilde{w}_{b_i} > w_{b_i}$.

Clearly, if the search cost r is sufficiently high, every matching is stable. In this case, for a given $\tau \geq 0$, the platform does not need to worry about implementing a matching that incentivizes agents to follow through its recommendations.

Theorem 3. *Suppose that the platform chooses $\psi = 0$ and $\tau \geq 0$.*

i) PAM is stable.

ii) If

$$r \geq \max_j\{q_j - c_j\} - \min_i\{u_i\} - \tau \tag{8}$$

then any matching μ is stable.

Proof: In the Appendix. ■

Theorem 3 states that stability is more easily attained if PAM is implemented. Intuitively, theorem 3 is similar to a known result in benchmark matching models without transaction or search costs, and in which agents' payoffs from a match are pre-determined and non-negotiable. The result states that, when both sides of the two-sided market have the same ordinal preferences over agents on the other side of the market and those preferences are strict, PAM is the only stable matching allocation (e.g., see Gusfield and Irving (1989)). But in or environment, it is not immediately clear which sellers are the ones most preferred by buyers and vice versa, as agents must engage in a bargaining process after the match is completed. But notice that, under the commission fee $\psi = 0$ and $\tau \geq 0$, sellers with the highest $q_j - c_j$ are the ones who can generate most surplus to be shared with buyers, so they can be interpreted as the "most preferred sellers". Similarly, buyers with lower u_i are the ones who can generate most surplus to be shared with sellers, so they can be interpreted as the "most preferred buyers". So following this intuition, in our model with search and transaction costs, a stable matching is more likely to be achieved if the platform matches the "most desirable seller" (i.e., the one with highest $q_j - c_j$) with the "most desirable buyer" (i.e., the one with lowest u_i), the "second most desirable seller" (i.e., the one with second highest $q_j - c_j$) with the "second most desirable buyer" (i.e., the one with second lowest u_i), and so on.

This result implies that policies aimed towards reducing r, such as prohibiting platforms from obfuscating search results (e.g., by prohibiting platforms from hiding shipping costs) can have a positive effect on consumer and seller surplus, by giving the platform more incentives to implement *PAM*, a matching that is *CPE*.

The next example illustrates how instability may reduce the overall profits of the platform.

Example 2. ***(To guarantee stability, the platform may have incentives to implement PAM)*** *Consider a market where the set of sellers is given by* $\{s_1, s_2, s_3, s_4\}$*, and the set of buyers is given by* $\{b_1, b_2, b_3, b_4\}$*. Sellers' quality and costs are given by*

$$(q_1, q_2, q_3, q_4) = (20, 20, 20, 20) \quad \textit{and} \quad (c_1, c_2, c_3, c_4) = (2, 8, 14, 16),$$

respectively. Customers' outside option are given by

$$(u_1, u_2, u_3, u_4) = (0, 3, 5, 7).$$

If PAM is implemented, then

$$s_1 : b_1 \quad s_2 : b_2, \quad s_3 : b_3, \quad s_4 : b_4$$

If the platform charges commission fees $\psi = 0$ *and* $\tau = 1$*, then this Pareto efficient match would generate a total of* 3 *transactions, resulting in a total profit of* $3 \times \tau = 3$ *to the platform.*

But if it implemented the match

$$s_1 : b_4, \quad s_2 : b_3, \quad s_3 : b_2, \quad s_4 : b_1 \tag{9}$$

then 4 transactions would take place, so the platform's profits would equal to $4 \times \tau = 4$*.*

However, if the search cost is given by $r = 1$*, it can be shown that the match* (9) *is unstable, as buyer* b_1 *would have incentives to form a blocking coalition with seller* s_1*, and buyer* b_2 *would have incentives to form a blocking coalition with seller* s_3*. Once these blocking coalitions are formed, buyers* b_3 *and* b_4 *remain unmatched. If they decided to pay the search cost* $r = 1$ *to be matched with sellers* s_3 *or* s_4*, they would end up with a payoff lower than their outside option, so they have no incentives to try to form a different match through the platform. As a result, in the end only two transactions take place, generating a profit of* $2 \times \tau = 2$ *to the platform.*

Meanwhile, PAM is stable (theorem 3). So the platform may be better off implementing PAM and getting a payoff of 3 *as opposed to implementing the matching* (9) *and getting a payoff of* 2*.*

7 Discussion

We build a simple theoretical model that sheds some light into two-sided platforms' incentives to steer customers towards certain products based on their willingness to pay. We show that the platform will tend to have incentives to induce more transactions than the socially optimum. We also derive some conditions under which the platform will have more incentives to adopt such a strategy. In particular, the higher the search costs, the more incentives the platform has to generate this excessive number of transactions, as in this case matchings that maximize the number of transactions will be more likely to be stable. This implies that policies directed towards reducing search obfuscation (e.g., by prohibiting platforms from hiding add on prices, such as shipping costs), can have a positive impact on sellers' and customers' surplus.

This work has some limitations that can be exploited by future research. Perhaps its biggest limitation is that, like in Hagiu and Jullien (2011) and Boerner and Quint (2022), we do not characterize the equilibrium when prices are (endogenously) set *before* the matchmaking takes place, a case that would be more consistent with many prominent platforms, such as Amazon.com. However, finding an equilibrium when prices are set before the matchmaking is probably challenging in such an environment with many buyers and many sellers, and it also requires a high level of information and rationality from sellers. As those assumptions on rationality are rather strong, in practice the platform may get a good approximation to what would be its optimal policy by treating prices as exogenous. When prices are exogenous, one can make a simple modification to algorithms 1 and 2 to derive the platform's optimal policy. In this case, instead of steering customers towards products with high costs, it will steer customers towards products with high posted prices.

While our environment is not dynamic, we partially capture the platform's reputation concerns by providing conditions under which a matching is stable. Indeed, if a matching is not stable, customers and sellers systematically receive "bad matches" in the sense that they are able to form better incentive-compatible matches on their own. So one could argue that instability reduces the probability that someone revisits the platform for future purchases. One limitation from our results on stability, however, is that we require that the platform only charges a flat fee τ per transaction (i.e., set $\psi = 0$), which is arguably not very realistic. Future research could potentially explore extensions to this result.

References

Anderson, S. P. and R. Renault (1999). Pricing, product diversity, and search costs: A bertrand-chamberlin-diamond model. *The RAND Journal of Economics 30*, 719–735.

Ausubel, L. M., C. Aperjis, and O. Baranov (2017). Market design and the fcc incentive auction.

Ausubel, L. M., P. Cramton, M. Pycia, M. Rostek, and M. Weretka (2014). Demand reduction and inefficiency in multi-unit auctions. *The Review of Economic Studies 81*(4 (289)), 1366–1400.

Azevedo, E. M. and E. Budish (2019). Strategy-proofness in the large. *The Review of Economic Studies 86*, 81–116.

Bergemann, D., B. Brooks, and S. Morris (2015, March). The limits of price discrimination. *American Economic Review 105*(3), 921–57.

Binmore, K. and J. Swierzbinski (2000). Treasury auctions: Uniform or discriminatory? *Review of Economic Design 5*, 387–410.

Boerner, L. and D. Quint (2022). Medieval matching markets. *International Economic Review n/a*(n/a).

Casner, B. (2020). Seller curation in platforms. *International Journal of Industrial Organization 72*.

Deshmukh, K., A. V. Goldberg, J. D. Hartline, and A. R. Karlin (2002). Truthful and competitive double auctions. In R. Möhring and R. Raman (Eds.), *Algorithms — ESA 2002*, Berlin, Heidelberg, pp. 361–373. Springer Berlin Heidelberg.

Diamond, P. A. (1971). A model of price adjustment. *Journal of Economic Theory 3*, 156–168.

Eliaz, K. and R. Spiegler (2011). A simple model of search engine pricing. *The Economic Journal 121*, F329–F339.

Ellison, G. (2005). A model of add-on pricing. *The Quarterly Journal of Economics 120*, 585–637.

Ellison, G. and S. F. Ellison (2009). Search, obfuscation, and price elasticities on the internet. *Econometrica 120*, 427–452.

Gusfield and Irving (1989). *The Stable Marriage Problem: Structure and Algorithms.* MIT Press.

Haghpanah, N. and R. Siegel (2019, February). Consumer-optimal market segmentation. Working Paper 3333940, Social Science Research Network (SSRN).

Hagiu, A. and B. Jullien (2011). Why do intermediaries divert search? *The RAND Journal of Economics 42*(2), 337–362.

Hannak, A., G. Soeller, D. Lazer, A. Mislove, and C. Wilson (2014). Measuring price discrimination and steering on e-commerce web sites. In *Proceedings of the 2014 Conference on Internet Measurement Conference*, IMC '14, New York, NY, USA, pp. 305–318. Association for Computing Machinery.

Kelso, A. S. and V. P. Crawford (1982). Job matching, coalition formation, and gross substitutes. *Econometrica 50*(6), 1483–1504.

Leininger, W. (1984). A generalization of the 'maximum theorem'. *Economics Letters 15*(3), 309–313.

Mattioli, D. (2012). On orbitz, mac users steered to pricier hotels. *The Wall Street Journal*.

Roth, A. E. and M. Sotomayor (1989). The college admission problem revisited. *Econometrica 57*, 559–570.

Roth, A. E. and M. A. O. Sotomayor (1990). *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis.* Econometric Society Monographs. Cambridge University Press.

Satterthwaite, M. A., S. R. Williams, and K. E. Zachariadis (2022). Price discovery using a double auction. *Games and Economic Behavior 131*, 57–83.

Silverman, J., C. Ma, and M. Buehren (2022). *R: A Language and Environment for Statistical Computing.*

Weiss, T. R. (2000). Amazon apologizes for price-testing program that angered customers. *Computerworld*.

Yang, K. H. (2022, April). Selling consumer data for profit: Optimal market-segmentation design and its consequences. *American Economic Review 112*(4), 1364–93.

Zhao, D., D. Zhang, M. Khan, and L. Perrussel (2011). Maximal matching for double auction. In J. Li (Ed.), *AI 2010: Advances in Artificial Intelligence*, Berlin, Heidelberg, pp. 516–525. Springer Berlin Heidelberg.

Zhou, B. and T. Zou (2022). Competing for recommendations: The strategic impact of personalized product recommendations in online marketplaces. *Marketing Science 0*(0), null.