The Importance of Exploration in Online Marketplaces

Siddhartha Banerjee, Zhengyuan Zhou and Ramesh Johari

Abstract—Nearly all online marketplaces face the challenge of ensuring the quality of sellers; this is particularly challenging in large marketplaces, where it may be infeasible to pre-screen all sellers. A growing trend in such settings is to use feedback from past transactions to infer the quality of new sellers. However, this process requires that existing buyers—the most valuable users of the platform—be subjected to the matching of new, untrusted sellers, potentially leading to a bad experience. This raises a natural exploration-exploitation trade-off: should the market platform match existing buyers to known good quality sellers, or alternately, use existing buyers to “explore” new, unknown sellers? The former ensures that matches are always successful, leading to good user experience; the latter, though holding the potential of discovering new high-quality participants, can lead to poor user experience. We investigate this tradeoff in a setting where existing buyers help uncover seller quality. We develop a stylized model, which captures the salient features inherent in such markets, while retaining analytical tractability. Our model uncovers a qualitative difference between exploration and exploitation, based on underlying network externalities between buyers and sellers in the market. In particular, in many settings, pure exploration achieves the optimal rate of successful matches.

I. INTRODUCTION

In the last decade, an ever increasing fraction of commerce has moved online. Online markets – platforms that enable exchange between buyers and sellers that are not necessarily co-located – are key players in this transformation. They remove search frictions, make available information (e.g., reputation), and take care of transaction details (e.g., payments). Successful online marketplaces include eBay, Lyft and Uber (transportation), Etsy (handmade goods), oDesk and Elance (online work), Airbnb (lodging), and so on.

In this work, we are interested in the question of reputation mechanisms, and how they affect market growth and revenues. There are two main ways in which an online platform can determine the quality of participants: (i) by curation/pre-screening, where the platform admits only participants with sufficiently high quality; and (ii) by using feedback from participants regarding past transactions. Though the former is often more accurate, it becomes prohibitive in terms of cost and effort when there are too many potential agents on either side of the market. In such settings, platforms rely on feedback from past transactions to infer quality – however, this leads to a natural ‘explore-exploit’ tradeoff, where the market can either match experienced buyers and sellers to ensure a good user-experience, or try out inexperienced agents to infer their quality from user-feedback, and thereby grow the pool of experienced agents.

Though such tradeoffs arise in most online marketplaces, the strategic decisions of the platform differ based on the features of the market. In this paper, we focus on one side of this problem – acquisition of high quality sellers in demand-constrained markets. We consider markets where a small number of buyers must choose from a large pool of sellers. As an example, consider online freelance marketplaces like Elance and oDesk: here the buyers are the clients bringing jobs to the marketplace, while the sellers are the freelancers looking for work. The number of freelancers on such platforms often far exceeds the number of clients bringing work.

In such demand-constrained settings where seller quality is inferred from feedback, the problem facing the marketplace can be summarized as follows: It can choose to match buyers with “trusted” sellers (i.e., those having good feedback in past transactions), preserving the user-experience for the buyers at the risk of stagnating growth. Alternatively, it can choose to match buyers with “unknown” sellers, thus potentially identifying new high-quality sellers, but at the risk of worsening the user-experience for existing buyers – and thus convincing them to move their future work to another platform. This is the question at the core of our work.

Initial thought may suggest that in such settings, the platform is always better off exploiting trusted sellers before exploring unknown sellers. Matching a buyer to a trusted seller ensures the buyer has a good experience, and ensures that the buyer brings work to the market in the future. On the other hand, exploring a new seller provides the platform information about the seller’s quality – however, it is not clear how this ‘information gain’ compensates the potential loss in future revenues due to buyers leaving the platform. This intuition is in sharp contrast to the practice followed by most platforms, where new sellers are often better treated than existing sellers. Our work takes a step towards uncovering the strategic reasoning behind this phenomenon.

A. Overview of our Results

The main insight of our results is that the explore/exploit decision is closely tied to the underlying network externalities in the marketplace. The presence of a large pool of buyers and/or trusted sellers often acts as a signaling mechanism, convincing new buyers to join the platform. We
consider a model of an online marketplace which exogenizes these externalities – in particular, we assume that the rate of new buyers grows with the number of buyers/trusted sellers on the platform, with different rates of growth.

We then ask: when would exploration be optimal for such a platform? We find sufficient conditions in terms of these externalities to ensure exploration will lead to the growth of the platform. In particular, we show the following:

- Exploration is most helpful when the cross-market externality dominates, i.e., when new buyer arrivals grow primarily due to the growth of the number of trusted sellers, then higher rates of exploration help prevent the market from collapsing.
- In contrast, when buyer arrivals are driven mainly by the number of existing buyers, then having a high rate of exploration may cause the market to collapse.
- Surprisingly, across a wide variety of settings, a pure exploration strategy also maximizes the rate of successful transactions in the market.

Though derived under a stylized model, our results make a strong qualitative case for the importance of exploration in online marketplaces. Some aspects of our analysis may arise from modeling artifacts, in particular, the assumption of exogenous externalities – however, our results and techniques point the way towards agent-based models which endogenize the externalities. Also, the assumption of demand-constrained markets allows us to isolate the role of network externalities – since such externalities are present in more general settings, our results may have wider applications.

### A. A Stochastic Model for Online Marketplaces

**Dynamic technology adoption.** A variety of models study how individuals choose to adopt new technologies. These models vary in their assumptions: for example, the observational social learning literature pioneered by Banerjee [1], Bikhchandani et al. [2], and Smith and Sorensen [3] considers a situation where individuals adopt an uncertain technology on the basis of observation of prior actions of others. Other models explicitly consider the role of network effects in causing adoption; a classic example of this is the Bass model [4], [5], [6], which proposes a fluid model of how technology adoption grows over time. The continuous time fluid analysis of our model may be seen as an extension of the Bass model to a setting with a two-sided market.

**Online learning.** Unlike the previous line of work, our setting is unique in that it contains an element of asymmetric information – the platform (and hence each buyer) does not know the quality of new sellers a priori. The resulting explore/exploit tradeoff mirrors similar tradeoffs in many learning settings, perhaps best embodied by the classical multi-armed bandit problem [7], [8], [9]. Our model is simplified from a learning perspective, as the platform learns quality of a seller after just one match. An interesting future direction is to consider the case where the platform must use multiple matches to learn the quality of a seller.

**Two-sided platforms.** A wide range of literature in economics considers the role of profit-maximizing platforms in intermediating between buyers and sellers. A particular thread of relevance to us is the literature on two-sided platforms, which aims to find optimal fee structures for each side of the market so as to maximize the profit of the intermediary [10], [11], [12]. Unlike our model, this work generally studies a static model without considering the dynamic aspect of market growth and information revelation.

**Dynamic search and matching.** Finally, our model is intended to capture markets where quality is not immediately known, and so there is a potential friction in finding the right match. Taking its roots in the labor economics literature, classic results by Diamond [13] and Mortensen and Pissarides [14] describe dynamic processes of search and matching in such markets. Our work is related to these papers, but distinct in that we consider a platform that can manipulate the matches to change the dynamics of the system.

### II. System Model

Before presenting our model, we briefly recap the setting we want to study. We focus on demand-constrained online matching-markets, where, to infer seller quality, the platform uses feedback from past transactions. This comes at a price: a buyer matched to a low quality seller may choose to leave the platform due to the bad experience. Our focus is on uncovering the strategic decisions behind high exploration rates in online marketplaces.

#### A. System Model

We consider a system evolving as a Markov chain in continuous time \( t \geq 0 \). The marketplace consists of a countable set of buyers and a countable set of sellers. All buyers are identical. Sellers on the other hand are either “good” or “bad” – we assume each seller is independently good with probability \( p \). A successful transaction or “match” is one which occurs between a buyer and a good seller. At the end of each transaction, the market receives feedback whether it was successful or not.

At time \( t \), the set of buyers is denoted as \( N_B(t) \). We assume that each buyer makes multiple visits to make transactions (or jobs), with visits following a Poisson process at rate \( \lambda \). For example, in freelance marketplaces, the set of buyers are clients who are registered on the platform, and a visit corresponds to a client bringing a job to the platform. Assuming all her matches are successful, a buyer leaves the platform at an intrinsic departure rate of \( r_B \), i.e., the time each buyer spends on the platform is exponentially-distributed with mean \( 1/r_B \). From standard properties of Poisson processes, this implies that each buyer brings on average \( \lambda/r_B \) jobs assuming all her matches are successful. However a buyer also departs immediately upon having an unsuccessful match. Finally new buyers arrive to the platform following a Poisson process at rate \( \Lambda_B \), which is a function

\[ 1 \text{More specifically, the number of jobs } K \text{ a buyer brings (assuming all matches are successful) obeys } \mathbb{P}[K = k] = p^k(1 - p), \text{ where } p = r_B^{-1}/\lambda. \]
of the state of the platform; we discuss this relationship in more detail below.

On the other side of the market, at time $t$, we have a large pool of sellers. Only a subset of size $N_S(t)$ of these are trusted, i.e., good sellers identified by the platform via feedback from past successful transactions. Trusted sellers leave the system independently after an $\text{Exponential}(rs)$ time. In addition, we assume we have an infinite supply of unknown sellers, each of whom is good with probability $\rho$. This assumption is a stylized reflection of the demand-constrained nature of the market: enough unknown sellers exist to meet the demand of arriving buyers.

How do buyers and sellers match? We focus on uniform, static matching policies with exploration rate $\epsilon \in [0, 1]$. In such a policy, upon bringing a job to the marketplace, a buyer is matched to a new seller with probability $\epsilon$, and to a trusted seller with probability $1 - \epsilon$. We refer to the $\epsilon = 1$ case as pure exploration, and $\epsilon = 0$ as pure exploitation. We assume that if no trusted seller exists in the system, then the buyer is matched to a new sellers. Our focus in this paper is on the case of pure exploration.

Note that in general, we may have to worry about availability of sellers, e.g., if trusted sellers are all “taken up” by other buyers. To avoid this issue we make a crucial simplifying assumption: we assume that all matches are instantaneous, so that there is “blocking” of trusted sellers. This is a simplifying assumption, but because it increases the supply of trusted sellers, it biases our model against exploration.

When a buyer is matched to a trusted seller, it is guaranteed to result in a successful match, and the seller now becomes trusted. When a buyer is matched to a new (untrusted) seller, a successful match happens with probability $\rho$, resulting in the seller becoming trusted; an unsuccessful match happens with probability $1 - \rho$, resulting in the effective removal of the seller from the platform. Further, we assume the buyer also departs from the platform in this case, having just received a negative experience. Note that this assumption again biases our model against exploration – in reality, buyers may leave only after multiple bad experiences rather than just one.

We assume that the platform is interested in maximizing the number of successful matches formed. This is reasonable because in many online marketplaces, revenue is directly dependent on the engagement of the users (i.e., some fixed percentage per transaction). We use $J(t)$ to denote the total number of successful matches generated until time $t$, given the initial conditions $N_S(0)$ and $N_B(0)$ (the implicit dependence of $J$ on the initial conditions is suppressed).

Another critical feature of online marketplaces is the presence of network externalities in agent arrivals – buyers and sellers are attracted to the marketplace due to the presence of agents already in the marketplace. We assume that new buyers arrive at a rate which can depend on $N_S$ and $N_B$. In particular, we assume a linear dependence, wherein the rate of buyer arrivals obeys:

$$\Lambda_B = c_S N_S + c_B N_B.$$

The parameter $c_S$ models an exogenous cross-market externality, where new buyers are attracted by the presence of existing trusted sellers. On the other hand, $c_B$ models an externality wherein new buyers are attracted by the presence of existing buyers in the marketplace. The latter is similar to models where long queues at a firm act as a signal of quality for uninformed agents [15]; the former is more relevant for online marketplaces which can advertise their trusted-seller pool as a signal of quality.

All our system parameters, unless mentioned, are assumed to be strictly positive (i.e., $> 0$). For convenience, we summarize our system parameters in the following table:

<table>
<thead>
<tr>
<th>Notation</th>
<th>Definition of Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
<td>Probability that an unknown seller is good</td>
</tr>
<tr>
<td>$r_s$</td>
<td>Departure rate of sellers</td>
</tr>
<tr>
<td>$r_b$</td>
<td>Departure rate of buyers</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>Rate at which buyers bring jobs</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>Exploration rate (Probability a match is an exploration)</td>
</tr>
<tr>
<td>$\Lambda$</td>
<td>Arrival rate of buyers</td>
</tr>
<tr>
<td>$c_S, c_B$</td>
<td>Arrival-rate coefficients ($\Lambda = c_S N_S + c_B N_B$)</td>
</tr>
</tbody>
</table>

Two interesting questions arise from this model: First, if the market always explores, under what conditions will the market grow unbounded (a desired outcome for a market designer) or decay (an undesired outcome)? Second, under the conditions that the market grows unbounded, which static policy (i.e., $\epsilon$) maximizes the asymptotic growth rate of the total revenue? We now address these questions.

### III. RESULTS

#### A. Formal System Dynamics

We first introduce some formalism for our system dynamics. Let

$$N(t) = [N_B(t) N_S(t)]^T$$

denote the state of the system – $N(t)$ is a continuous-time Markov process, taking values in the non-negative integer lattice $\mathbb{Z}_+^2$. We define $\overline{p} = (1 - \rho)$, i.e., the probability that an exploration leads to a bad match. Now, for $n_S > 0$, the transition rates obey:

$$
\begin{align*}
\begin{bmatrix} n_B \\ n_S \end{bmatrix} &\rightarrow \\
\begin{bmatrix} n_B + 1 \\ n_S \end{bmatrix} &\text{at rate } c_B n_B + c_S n_S \\
\begin{bmatrix} n_B - 1 \\ n_S \end{bmatrix} &\text{at rate } r_B n_B + \lambda n_B c_B \overline{p} \\
\begin{bmatrix} n_B \\ n_S + 1 \end{bmatrix} &\text{at rate } \lambda n_B \epsilon \rho \\
\begin{bmatrix} n_B \\ n_S - 1 \end{bmatrix} &\text{at rate } r_S n_S
\end{align*}
\tag{1}
$$

In addition to these transitions, there is a stream of events corresponding to successful ‘exploitation’ matches – wherein a visiting buyer is matched to a trusted seller, resulting in a good match. In state $[n_B \ n_S]^T$, these exploit events occur at rate $\lambda n_B (1 - \epsilon)$. Though they do not change the state of the system, they are important for estimating the rate of successful matches – we return to this issue in Theorem 2. Note also that state $[0 \ 0]^T$ is an absorbing state – this corresponds to the marketplace collapsing.

For states $[n_B \ 0]^T$, $n_B > 0$, i.e., when there are no trusted sellers on the platform, exploitation is not feasible. A natural option is to revert to pure exploration in such a
state – this corresponds to the transition:
\[
\begin{bmatrix} n_B \\ 0 \end{bmatrix} \rightarrow \begin{bmatrix} n_B \\ 1 \end{bmatrix}
\]
at rate \(\lambda n_B \epsilon\rho\)

In order to make the analysis easier, we make the following simplification: we assume that the transitions in Equation (1) hold even at \(\begin{bmatrix} n_B & 0 \end{bmatrix}^T\). We henceforth refer to this as the boundary approximation, and denote the resulting process by \(\tilde{N}(t)\). Our subsequent theoretical results are developed under the boundary approximation. Note that the boundary approximation is exact when \(\epsilon = 1\), i.e., under pure exploration. On the other hand, for any \(\epsilon < 1\), the boundary approximation only serves to slow down the rate of departure of buyers – in the extreme case, under \(\epsilon = 0\), it results in buyers departing at their natural rate, without performing any transactions.

A consequence of the boundary approximation is that the dynamics admit the following differential equation:
\[
\frac{d}{dt} \begin{bmatrix} \mathbb{E}[N_B(t)] \\ \mathbb{E}[S(t)] \end{bmatrix} = \begin{bmatrix} c_B - r_B - \lambda \epsilon \rho & c_S \\ \lambda \epsilon \rho & -r_S \end{bmatrix} \begin{bmatrix} \mathbb{E}[N_B(t)] \\ \mathbb{E}[S(t)] \end{bmatrix}.
\]

Let \(A_\epsilon \triangleq \begin{bmatrix} c_B - r_B - \lambda \epsilon \rho & c_S \\ \lambda \epsilon \rho & -r_S \end{bmatrix}\). Solving the above differential equation, we get:
\[
\mathbb{E}[\tilde{N}(t)] = e^{A_\epsilon t} \mathbb{E}[	ilde{N}(0)],
\]
(2)

Let \(\beta_1(A_\epsilon) \geq \beta_2(A_\epsilon)\) be the two eigenvalues of \(A_\epsilon\) – it is easy to check (via the characteristic polynomial) that both are real. Now if both eigenvalues are strictly negative, the above equation indicates that the expected number of agents in the marketplace goes to zero. This can be used to get a high probability result for the collapse of the marketplace. The absorption time of the Markov chain, i.e., the time to market-collapse, is denoted by \(\tau_{n_0} = \inf_{t > 0} \{ \tilde{N}(t) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}^T \} \). Then we have:

**Lemma 1:** If \(\beta_2(A_\epsilon) < 0\), then for any \(t > 0\),
\[
\mathbb{P}[\tau_{n_0} > t] = O(e^{-|\beta_1|t}),
\]
and also, \(\mathbb{E}[\tau_{n_0}] = O(1)\).

**Equation:** If \(\beta_2(A_\epsilon) < 0\), the solution of the simplified dynamics \(\tilde{N}(t)\) is given by:
\[
\tilde{N}(t) = \begin{bmatrix} a_{11} e^{\beta_1 t} + a_{12} e^{\beta_2 t} \\ a_{21} e^{\beta_1 t} + a_{22} e^{\beta_2 t} \end{bmatrix},
\]
for some constants \(a_{11}, a_{12}, a_{21}, a_{22}\), which depend on system parameters and \(\tilde{N}(0)\), but are independent of \(t\). Now we have:
\[
\mathbb{P}[\tau_{n_0} > t] = \mathbb{P} \left[ \tilde{N}_B(t) + \tilde{N}_S(t) > 0 \right] = \mathbb{P} \left[ \tilde{N}_B(t) + \tilde{N}_S(t) > 0 \right]_{n_0} \\
\leq \mathbb{E} \left[ \tilde{N}_B(t) + \tilde{N}_S(t) \right]_{n_0} \\
= a_1 e^{\beta_1 t} + a_2 e^{\beta_2 t} = O(e^{-|\beta_1|t}),
\]
where \(a_1 = a_{11} + a_{21}, a_2 = a_{12} + a_{22}\). Moreover, we have:
\[
\mathbb{E} \left[ \tau_{n_0} \right] = \int_0^\infty \mathbb{P}[\tau_{n_0} > t] dt \\
= \frac{a_1}{e^{|\beta_1|}} + \frac{a_2}{e^{|\beta_2|}} = O(1)
\]

Thus, under the conditions of Lemma 1, the market collapses within a constant time (on average), independent of the initial conditions. On the other hand, \(\beta_1(A_\epsilon) > 0\) is sufficient to ensure that the market does not collapse with some positive probability:

**Lemma 2:** If \(\beta_1(A_\epsilon) > 0\), then for any \(n_0 \neq 0\):
\[
\mathbb{P}[\tau_{n_0} < \infty] < 1
\]

The proof of this result is somewhat technical – due to shortage of space, we only give an outline here, and defer the complete proof to our technical report [16]. To prove the system does not collapse, we consider the evolution of Markov chains to show that \(S(t)\) escapes to infinity with positive probability. This gives our result.

The above lemmas show that \(\beta_1 > 0\) forms a sharp threshold for the survival of the marketplace. Returning to the original dynamics, we can now get conditions for when the market survives under pure exploration:

**Theorem 1:** For any non-zero initial state \(n \in \mathbb{Z}_+^2\), the market survives under pure exploration if and only if one of the following hold:
- If \(c_B < r_B + \epsilon\rho\) and \(\rho > \frac{r_S}{r_B + \epsilon\rho} (1 + \frac{c_B - c_S}{\lambda})\), and \(\epsilon\) satisfies:
  \[
  \epsilon > \frac{r_S - r_B}{c_B - c_S}
  \]
- If \(c_B < r_B + \epsilon\rho\) and \(\rho > \frac{r_S}{r_B + \epsilon\rho} (1 + \frac{c_B - c_S}{\lambda})\), and \(\epsilon\) satisfies:
  \[
  \epsilon (1 - \frac{r_S + r_B}{c_B - c_S}) < \frac{c_B - c_S}{\lambda}
  \]
- If \(c_B > r_B + r_S\), and we choose:
  \[
  \epsilon \leq \frac{c_B - c_S}{\lambda (r_B - r_S)}
  \]

**Proof:** Recall that \(A_\epsilon = \begin{bmatrix} c_B - r_B - \lambda \epsilon \rho & c_S \\ \lambda \epsilon \rho & -r_S \end{bmatrix}\), and \(\beta_1(A_\epsilon)\) denotes the leading eigenvalue of \(A_\epsilon\). From Lemmas 1 and 2, we have that \(\beta_1(A_\epsilon) > 0\) is a necessary and sufficient condition for \(\tilde{N}(t)\) to survive with some positive probability. Now, from the characteristic polynomial of \(A_\epsilon\), we have:
\[
\beta_1(\epsilon) = \frac{(c_B - r_B - r_S - \lambda \epsilon \rho)^2}{2} + \frac{\sqrt{(r_B - c_B - r_S + \lambda \epsilon \rho)^2 + 4 \lambda c_S \rho}}{2}
\]

Let us define \(a = (c_B - r_B - r_S - \lambda \epsilon \rho)\) and \(b = \sqrt{(r_B - c_B - r_S + \lambda \epsilon \rho)^2 + 4 \lambda c_S \rho}\) (note that \(b \geq 0\)). Now as in the theorem statement, we consider the following cases:
Let $b_t = (\epsilon - 1) A_t - \epsilon B_t$ be the cumulative number of good matches up to time $t$ in the simplified dynamics. In settings where $\beta_1(A) > 0$, we have:

$$\beta_1(A) = \max \left\{ \beta_1(A_1), \beta_1(A_2) \right\}.$$  

Furthermore, in these settings, $\beta_1(A_3)$ is maximized by pure exploration (i.e., $\epsilon = 1$) when one of the following holds:

- **Condition 1**: $r_B > c_B$, and $\rho > \frac{r_B - c_B}{r_s + \epsilon \beta}$
- **Condition 2**: $r_B < c_B$, and $\rho > 1 - \frac{r_B - c_B}{r_s + \epsilon \beta}$

In all other cases, $\beta_1(A_3)$ is maximized by following pure exploration (i.e., $\epsilon = 1$).

Proof: Recall that a good match occurs when a buyer is matched to a good seller – a trusted seller in case of exploration, and a newly discovered one in case of exploitation. When the system is in state $N(t)$, transactions occur at rate $\lambda N_B(t)$, of which $(1 - \epsilon)$ fraction are exploitations, and $\epsilon \rho$ are successful exploitations. Thus, we have:

$$\frac{d}{dt} E[M(t)] = \lambda (1 - \epsilon (1 - \rho)) E[N_B(t)].$$

From Equation (2), we have that the expected number of buyers over time obeys $E[N_B(t)] = a_{11} e^{\beta_1 t} + a_{12} e^{\beta_2 t}$, where $a_{11}, a_{12}$ are functions of the system parameters and the initial state $\tilde{N}(0)$. Substituting this into the above equation, and setting $\tilde{M}(0) = 0$, we get:

$$E[\tilde{M}(t)] = a_0 + \tilde{a}_1 e^{\beta_1 t} + \tilde{a}_2 e^{\beta_2 t},$$

where constants $a_0, \tilde{a}_1$ and $\tilde{a}_2$ are again functions of system parameters (including $\epsilon$), but not of $t$. Now, whenever $\beta_1 > 0$, from the discrete Laplace principle, we get:

$$\lim_{t \to \infty} \frac{1}{t} \log E[\tilde{M}(t)] = \beta_1(A).$$

Thus, conditioned on the marketplace growing, to maximize the rate of good matches it suffices to maximize $\beta_1(A)$. The rest of the proof follows from straightforward computations – in particular, we show for $\epsilon \in [0, 1]$, under Conditions 1 or 2, $\beta_1(A)_{\epsilon}$ is increasing in $\epsilon$, and when neither condition holds, then $\beta_1(A)_{\epsilon}$ is decreasing in $\epsilon$. Due to shortage of space, we defer the details to [16].

The above theorem essentially shows that when the marketplace grows, then the rate of growth of successful matches is closely tied to the rate of growth of buyers in the system – in particular, the asymptotic log-rates of the two are the same. Furthermore, this rate is maximized either by pure exploration or pure exploitation. In particular, pure exploitation is optimal only when $r_B < c_B$ and the percentage of good sellers $\rho$ is below a certain threshold – in all other cases, pure exploration is optimal.

Note that the optimality of pure exploration carries over to the process $N(t)$, as the two processes are identical for pure exploration, and for all other $\epsilon$, the boundary approximation only records spurious successful matches when $N_B(t) = 0$. We confirm this behavior in Figure 1 by simulating the process $N(t)$ for $\epsilon = 0, 1$ in the two regimes.

IV. SIMULATIONS

We verify our theoretical results using simulations – in particular, we test the validity of the boundary approximation, and the results of Theorem 2. Due to shortage of space, we present some representative plots in Figure 1 (simulation parameters are specified in the plot titles) – for more extensive plots, and simulation details, refer our technical report [16].

V. DISCUSSION AND CONCLUSIONS

Our work takes an important step towards uncovering the strategic rationale behind exploration in demand-constrained online marketplaces. In such markets, new sellers are often matched to buyers preferentially, compared to existing trusted sellers – this contradicts the intuition that exploitation is more beneficial to the platform. Our main message is that such behavior is rational, and in fact, often optimal, depending on the nature of underlying network externalities. Intuitively, when the externality is driven by the number of sellers in the system, exploration improves the signal of quality for an online marketplace, attracting new buyers. We formally study this phenomenon in a stylized model with exogenous externalities.

Because our model is so stylized it clearly has some drawbacks that represent avenues for future work. In particular,
there are two modeling assumptions that are problematic when considered relative to real markets. First, we have assumed that transactions are *instantaneous*. This has several important consequences, most notably that it does not allow for the possibility that newly arriving buyers may be *blocked* from matching to an existing seller. Our techniques can be extended to allow for holding-times – however the resulting analysis is more complicated. Refer [16] for details.

The second key modeling assumption is that externalities in our model are exogenous, i.e., not driven by individual agent behavior. To have a more accurate model, we should consider agent incentives in joining the platform, and leaving the platform: this would *endogenize* the origin of the externalities. Our results, under exogenous externalities, strongly suggest that the explore/exploit decisions are closely tied to the nature of the externality, so understanding the origins of this externality is important to design better matching schemes. This is our primary direction of future investigation.

**REFERENCES**


