The Sunk Cost Fallacy in Reverse Auctions

Yu Wu*
Stanford University
Stanford, CA

Hang Ung*
Ecole Polytechnique
Palaiseau, France

Christina Aperjis
HP Labs
Palo Alto, CA

ABSTRACT

We empirically study buyer behavior in an online outsourcing website where sealed bid auctions are held with bids arriving over time. We focus on when buyers terminate their requests and how they behave when choosing the winning bid. We find that buyers are more likely to choose the last bid; and all other positions (i.e., the first bid, the second bid, until the penultimate bid) are chosen with approximately the same frequency. We provide a simple probabilistic model that captures this behavior. The key characteristic of this model is that buyers are more likely to stop when the most recent bid is the best so far. This feature is related to the sunk cost fallacy: once a buyer has waited for some time, she has an escalating tendency to continue waiting until a bid that is better than all prior bids arrives. A buyer is unwilling to recall early bids, because that would make her perceive the time since the arrival of early bids as "wasted", even though the time cost has already been incurred at the time of the decision.

Keywords

Electronic Commerce, Reverse Auctions, Online Auctions, Outsourcing, Sunk Cost Fallacy

Categories and Subject Descriptors

J.4 [**Computer Applications**]: Social and Behavioral Sciences

1. INTRODUCTION

In the last decade, the Internet has not only become a vast shopping mall for consumers and companies, but has also helped various forms of auctions reach unprecedent scales. Examples include the English auction, which has been popularized by eBay, the generalized second-price auctions in online advertising, "penny" auctions, and reverse/outsourcing auctions. In order to design these systems optimally, it is important to understand how buyers and sellers behave.

This paper studies buyer behavior in vWorker[1], an online outsourcing auction platform that focuses on computer programming. In vWorker, a company or an individual (the *buyer*) outsources an IT project through a reverse auction mechanism: any individual (the *bidder* or *seller*) can submit sealed bids to any of the requests posted by buyers. For each request, bids arrive over time and the buyer can decide at any time to terminate the auction and select a bid. Our goal is to study the behavior of the buyer with respect to the timing of her decisions.

The fact that a buyer generally prefers to find a good deal without having to wait or search too long creates the following tradeoff: terminating the auction early gives speed – avoiding for instance, to delay a larger project; while waiting might increase the quality and/or decrease the price of the best bid. Since humans typically experience "entrapment" in waiting situations [11], one can ask whether buyers in online reverse auctions are also prone to a form of psychological bias with respect to the decision of terminating the auction.

Our study begins by observing statistical patterns in the number of bids received by requests at the time they were closed by the buyers (hereinafter *the number of bids*), and by assessing the relation between a bid's position (*i.e.*, if it was the first bid of the request, the second, or on the contrary, the last one before the buyer closed the request) and its empirical probability to be chosen. Two key observations arise from the data. First, the number of bids follows a geometric distribution. Second, for any given number of bids n, the last bid (whose position is n) is chosen more frequently while all other positions are chosen with approximately the same frequency.

We show that this behavior cannot arise from simple models of rationality. We then provide a simple probabilistic model that captures this behavior. In this model, after the arrival of each bid the buyer stops with some probability, depending on whether the most recent bid is the best so far: the buyer stops with a larger probability if the most recent bid is the best so far. Specifically, this model is characterized by two probabilities: the probability of stopping when the most recent bid is the best so far, denoted by p, and the probability of stopping when the most recent is not the best so far, denoted by q. We refer to this model as the pq *stopping rule.*

The key characteristic of the pq stopping rule is that $p > q$,

*Work performed in part while an intern at HP Labs.

[1]`http://www.vworker.com`, previously known as Rent A Coder (`http://www.rentacoder.com`)

that is, the buyer is more likely to stop when the most recent bid is the best so far. This behavior is related to the *sunk cost fallacy.* Sunk costs are costs that have already been incurred and cannot be recovered, and thus should not influence an agent's decisions. The sunk cost fallacy arises when sunk costs affect actors' decisions; evidence from behavioral economics suggests that it is often the case [1]. The sunk cost fallacy has been observed in a variety of settings, where the cost can be in terms of money, effort or time. The latter occurs in the context of vWorker.

A buyer that has already spent time waiting for a better bid has incurred some cost. A rational buyer would realize that this cost is sunk and would not let it affect her decision. On the other hand, an irrational buyer feels that this time would be wasted if she terminates the request when the most recent bid is not the best so far. Thus, that a buyer may be more likely to close a request when the most recent bid is the best so far is a form of sunk cost fallacy.

The paper is organized as follows. After reviewing related work in Section 2, we describe vWorker and our data set in Section 3. Section 4 presents our main observations from the data. Section 5 describes the pq stopping rule, which is consistent with the observations from the data. Section 6 discusses several aspects of the pq stopping rule with respect to the sunk cost fallacy.

2. RELATED WORK

Since the emergence of e-commerce, specific questions related to the structure of online markets and the behavior of their users have drawn a lot of interest from researchers. Various models are attempting to describe the economics of crowdsourcing (e.g., [5, 8]) while a few empirical studies have considered crowdsourcing markets like Taskcn [15] and the Amazon Mechanical Turk [10]. Outsourcing platforms have been considered with respect to management issues. For instance, Gefen and Carmel have studied vWorker (then called Rent A Coder) to investigate the flat-world effect in offshore outsourcing [7]. They pointed out that despite the higher prices, buyers tend to prefer domestic service providers, especially those they previously worked with.

In Taskcn, one of the largest online outsourcing websites in China, workers submit sample solutions to the task to compete for winning it. Yang et al. [15] showed that experienced workers in Taskcn learn to submit later in order to get a better chance of winning, because the time a worker takes before submission could be a signal of the effort spent on it. However, this effect does not arise in vWorker, because coders do not submit sample solutions in order to enter the auction. We suggest that buyer behavior on vWorker is related to the sunk cost fallacy.

Evidence of the sunk cost fallacy has been provided by various disciplines across the social sciences. Through a series of experiments, Arkes and Blumer [1] showed that people have a "*great tendency to continue an endeavor once an investment in money, effort, or time has been made*", and called this behavior the "*sunk cost effect*". Before that, similar behavior had been identified: people tend to stick to the previous investment decisions that bring negative consequences [12]. This escalation of commitment to a failing course of action is further studied by Whyte [14] and Brockner [4] from the organizational and social psychology perspective. More generally, the sunk cost fallacy can be related to loss aversion and prospect theory [9, 13]. The sunk cost fallacy often emerges when decision makers are facing a sequence of similar situations. In this paper, we study how this effect influences a buyer's stopping decision in a reverse auction with sequential bids. Recently, Augenblick [2] showed that bidding behavior in Swoopo can be explained by the sunk cost fallacy. However, Swoopo uses penny auctions (a relatively new auction format), whereas in this paper we study reverse auctions (an auction format that is widely used in procurement both online and offline).

3. VWORKER

3.1 The Auction Process

vWorker is an outsourcing auction website which connects buyers – companies or individuals – with workers, some working as freelancers, others acting on behalf of their firms. Since our data is from Rent A Coder, which was specializing in IT projects (such as websites, web browser add-ons, or small pieces of software), in what follows we refer to workers as coders.

A buyer starts an auction by posting a request, describing the project. Coders can browse through all the open requests by category and bid. A bid consists of a price and generally includes a text where the coder provides information on her experience and/or explains how she intends to execute the project. Coders cannot observe the bids that a request has already received, i.e., this is an auction with sequential sealed bids. Moreover, the buyer does not know the number of bidders in advance.

At any point in time, a buyer can terminate the auction, either by choosing a bid (if any bid was received), or by canceling her request. In both cases, the request is then closed to new bids. When a bid is selected, an escrow is secured from the buyer and the coder can start working. Once the work is done, the coder is paid by the buyer according to the price specified in the bid. Requests not closed by the buyers expire after 3 months.

3.2 Dataset

Our data set consists of all requests and bids posted on the website between the website's inception (November 2000) and October 2009. For our present purposes, we discarded the requests which were cancelled or had expired, as well as those with only one bid. Indeed, those requests are either not relevant for studying the termination and selection behavior, or in the case of single-bid requests, might introduce noise in the data. For instance, it could be the case that a buyer has already decided to contract an acquaintance and uses vWorker only to secure the transaction. The resulting data set contains 207,000 requests and 2.4 million bids. We also have information on the 380,000 users (buyers and coders) who were active during that period.

The data set was stored in a MySQL database. Analysis and simulation were conducted with the software R.

4. OBSERVATIONS

Since our objective is to understand when buyers terminate the requests and how they choose the winners, there are two characteristics of each request we should pay particular attention to: how many bids it received and which bid was chosen by the buyer. For each request we index the bids in the chronological order of arrival. We use the following notation:

- $R(n)$ is the number of requests with n bids,
- $r(n, k)$ is the number of requests with n bids for which the kth bid is chosen,

- $\pi(n,k) = r(n,k)/R(n)$ is the proportion of requests with n bids for which the kth bid is chosen.

We also call $\pi(n,k)$ the (empirical) "probability" that a buyer chooses the kth bid in a request with n bids. We make two observations on $R(n)$ and $\pi(n,k)$.

OBSERVATION 1. *$R(n)$ follows a geometric distribution in n.*

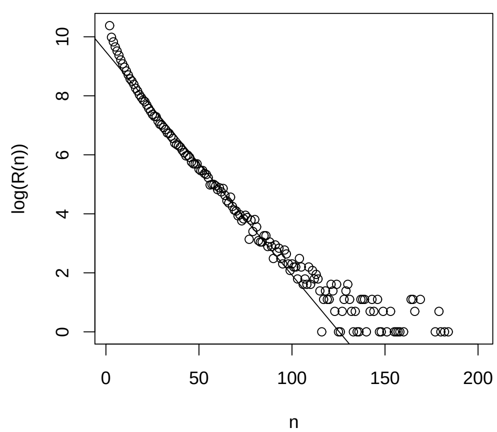

Figure 1: $R(n)$, the number of requests that have n bids decreases in n. The trend is approximately exponential, as shown by the linear regression *log*$(R(n)) \sim n$ for $2 \le n \le 100$ (solid line).

Observation 1 says that the number of requests with n bids is geometrically decreasing in n. Figure 1 shows $\log(R(n))$ as a function of n. We see that the curve of dots is close to a straight line in the range $2 \le n \le 100$. The curve gets flat and noisy after $n > 100$, possibly because of the large variance of the tail distribution. Restricting our attention to $2 \le n \le 100$, we get the following regression formula for *log*$(R(n))$:

$$\log(R(n)) = 9.4916 - 0.0757 \cdot n, \quad 2 \le n \le 100 \qquad (1)$$

The R^2 of the regression is 0.99, hence the geometric distribution fits the data very well.

OBSERVATION 2. *For any $n > 1$:*

(a) $\pi(n,n)$ is significantly larger than $\pi(n,k)$ for all $k < n$.

(b) $\pi(n,k)$ is approximately the same for all $k < n$.

In other words, buyers are much more likely to accept the last bid, whereas all positions before the last are selected with approximately the same frequency.

Figure 2 (*top*) is a histogram of requests with $n = 25$ bids, showing $\pi(25,k)$ as a function of k. We see that the last bar is significantly higher than previous bars. Moreover, all bars with $k \le 24$ have approximately the same heights. We also note that there is a weak increasing trend of $\pi(n,k)$ in k for large k.

Observation 2 is more generally shown on Figure 2 (*bottom*). The grey line is $1/n$, i.e., the probability that a given position would be chosen if all positions were chosen with equal probabilities. Circles represent $\pi(n,n)$, the proportion of requests where the last bid was chosen. Triangles $\bigtriangledown$

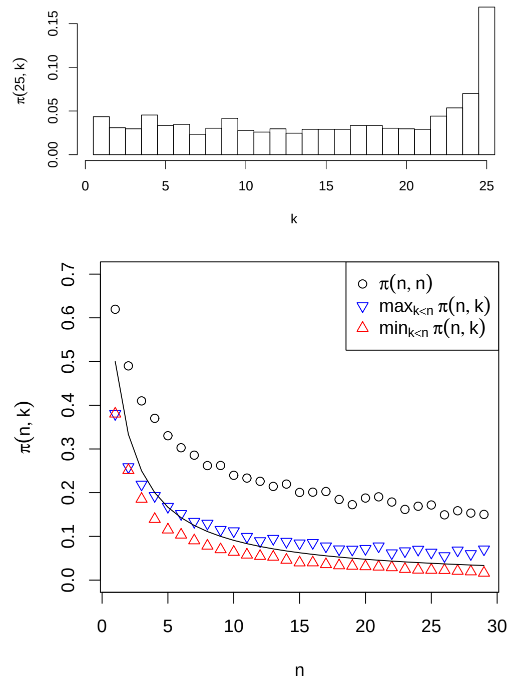

Figure 2: The empirical probability $\pi(n,k)$, as illustrated by the histogram in the case where $n = 25$ (*top*); and as visualized more generally by $\pi(n,n), \max_{k<n} \pi(n,k)$ and $\min_{k<n} \pi(n,k)$ for all $2 \le n \le 25$ (*bottom*).

and $\bigtriangleup$ are respectively $\max_{k<n} \pi(n,k)$ and $\min_{k<n} \pi(n,k)$. Figure 2 provides support for Observation 2(a) since circles are always significantly above the triangles, and also for Observation 2(b) in that the differences between (each pair of) $\min_{k<n} \pi(n,k)$ and $\max_{k<n} \pi(n,k)$ are relatively small.

To verify Observation 2(b) statistically, for each n we conduct the Pearson test on the null hypothesis that $\{\pi(n,k), k < n\}$ are equal. Given the weak increasing trend when k is close to n, we test the following series of hypotheses

$$H_0^{n,m} : \{\pi(n,k), k = 1, \cdots, n-m\} \text{ are equal.}$$

where $m = 1, 2, 3$ and 4.

Results of the tests show that for most values of n, we cannot reject the null hypotheses $H_0^{n,3}$ and $H_0^{n,4}$. Yet $H_0^{n,1}$ is rejected at the 0.05 significance level for all n and $H_0^{n,2}$ is rejected for some n. These results support Observation 2 for $k < n-2$, and are also consistent with the weak increasing trend we observe in Figure 2 (top). We are going to discuss the trend later in in Section 6.

5. MODELS

In this section, we search for a model that is consistent with the observations in Section 4. Before considering specific models, we introduce a general framework that applies throughout this section.

We associate each bid with the utility that the buyer would derive by selecting it. We assume that this utility depends on the bid's attributes (e.g., price, bidder's reputation). We note that a bid's utility is subjective, since different buyers may value price and reputation differently. However, since a bid is only addressed to one buyer (that is, the buyer that posted the request), in what follows we sometimes say *the utility of the bid* to refer to the utility that the buyer would derive by selecting it.

We will refer to the bid that the buyer selected as the "best bid." In other words, the fact that a bid is selected by a buyer implies that it maximizes the buyer's utility among all available bids.

The following assumptions will be used throughout this section.

ASSUMPTION 1. *For any request, the utilities of bids are independent of the order that the bids arrive.*

ASSUMPTION 2. *Buyers can observe a new bid immediately after it arrives.*

Assumption 1 says that later bids are neither better nor worse than early bids on average. In an open bid auction where bidders can observe previous bids, like in eBay, there is dependence between a bid's price (or more generally utility) and the arrival order. In particular, a bidder would probably not bid if she knows that her bid is worse than any previous bid. As a result, later bids are better. On the other hand, this dependence is largely reduced in vWorker, because bids are sealed. In fact, the data show that there is no (decreasing) trend in the prices of bids for a particular request as they arrive over time.[2] This suggests that Assumption 1 is reasonable. Note that the independence in Assumption 1 is unconditional, therefore it does not conflict with Observation 2, which says that the last bid is more likely to be the best bid, because the word "last bid" itself exhibits conditionality.

Assumption 2 is motivated by the fact that buyers in vWorker have the option to receive instantaneous notifications on new bid arrivals for their requests. With such assistance, a buyer can observe a new bid for her request soon after the bid arrives without checking the vWorker website every minute. We are going to revisit Assumption 2 in Section 6.

5.1 Deterministic Models

First we consider a deterministic model in which the buyer has prior knowledge about the market and decides when to terminate the request by comparing (each time a bid arrives) the benefit and cost of waiting. In order to quantitatively analyze the problem, we assume that for a given request the utilities of the bids are independently and identically drawn from some distribution known by the buyer. We assume that the buyer's objective is to maximize her utility from the best bid minus the disutility of waiting.

As shown in Appendix A, if (1) the utilities of bids are drawn from a uniform distribution, (2) the arrival process of bids is Poisson and (3) buyers have linear waiting costs, then the optimal strategy of the buyer is a simple threshold strategy: there exists a threshold such that when the utility of a new bid exceeds this threshold, the buyer stops the request and chooses this new bid. In this case, no matter when the request is terminated, the best bid (which is chosen by the buyer) is always the last bid.

Though the threshold strategy is concise and natural, it does not explain the data. We find that there are very few buyers in our dataset that always choose the last bid, therefore threshold strategies are probably not widely used among buyers in vWorker.

Besides this particular case which leads to a threshold strategy, we have also assessed our deterministic model though numerical simulations by using different combinations of utility distributions, bid arrival processes and waiting cost functions. We have not been able to find a simple deterministic model that gives rise to Observations 1 and 2. Details can be found in Appendix A.

Another potential model for the buyer's behavior is that when a bid that is better than all previous bids arrives (say bid A), the buyer waits a bit longer before making a decision in order to see if an even better bid will arrive soon. If either no new bids arrive within some time interval or only worse bids arrive, then the buyer accepts A. This model implies that the buyer waits for a longer time before closing the request when the last bid is not the best so far. However, the data shows that this is not the case, suggesting this is not a prevailing behavior in the buyer population of vWorker.

Since a well fitted deterministic model would require a high level of model complexity, in the following sections, we consider two simple probabilistic models that attempt to explain the data.

5.2 Probabilistic Model: q Stopping rule

The following stopping rule, with only one parameter, is probably the simplest model we could consider:

q stopping Rule: A buyer terminates her request with probability q each time a new bid arrives.

Under this stopping rule, we have that a request has n bids with probability $(1-q)^{n-2}q$ (we only study requests that have at least two bids). Thus, the number of requests with n bids follows the geometric distribution in n, consistent with Observation 1. Further, since the buyer terminates the request only based on number of bids in the request, then according to Assumption 1, each existing bid has the same probability to be the best one and therefore all $\pi(n,k)$'s are equal for all $k = 1, \cdots, n$, which is consistent with Observation 2(b). However, Observation 2(a) is not satisfied, that is, the probability of accepting the last bid is not higher that the probability of accepting an earlier bid.[3]

We conclude that the q stopping rule can explain some of our observation from the data, but not why buyers are more likely to choose the last bid.

5.3 Probabilistic Model: pq Stopping Rule

Though the previous models have some difficulties explaining all three features in our observations, they provide some insights for finding a better model.

- In Section 5.1, we considered a threshold utility with which buyers always stop at a bid better than all previous bids and always accept the last bid. Conversely, the observation that buyers are more likely to accept the last bid (Observation 2(a)) actually suggests that buyers are more likely to terminate the request when the most recent bid is better than all previous bids.

- The geometric distribution in Observation 1 is obtained if buyers stop at each step with a constant probability q: the probability that a request has n bids is

[2]This is based on one and two-sided t-tests. The null hypothesis is that the sequence of price differences between the bids of a particular request has zero mean.

[3]The q stopping rule cannot explain Observation 2(a), even if we allow q to vary over time.

$q(1-q)^{(n-2)}$, which follows exactly the geometric distribution with parameter q. In this case, buyers are equally likely to accept each bid, including the last one.

Based on these insights, we propose the following model.

***pq* Stopping Rule:** When a buyer observes a new bid, she stops the request

- with probability p if the new bid is the best bid so far,
- with probability q if the new bid is not the best so far.

If she does not stop the request, she continues to wait.

Under the pq stopping rule, we can compute the probability that a request has n bids as a function of p and q.[4] We can also compute the probability that the kth bid is chosen in requests with n bids. We use these formulas together with the empirical probabilities from Observations 1 and 2 to estimate p and q. In fact, we find that p is not a constant and is increasing in n, the number of bids currently in the request. Let p_n be the probability that the buyer terminates the request when the most recent bid (the nth bid) is better than all previous bid. Then we have the following estimates:

$$\hat{q} = 0.0696, \quad \hat{p}_n = 0.1236 + 0.0085n \qquad (2)$$

Details of the estimation process can be found in Appendix B.

We observe that $\hat{p}_n > \hat{q}$, that is, the buyer is more likely to stop the request when the most recent bid is the best so far. In other words, buyers tend to wait for new bids until the new one is better than all previous bids. This is closely related to the sunk cost fallacy, which is described as a "great tendency to continue an endeavor once an investment in money, effort, or time has been made." Although the time already spent on waiting is sunk, it seems that buyers still take into account the waiting costs when they decide whether to stop.

Moreover, $\hat{p}_n$ (or $\hat{p}_n/\hat{q}$) is increasing in n, which means that the effect of the sunk cost fallacy becomes more significant as the number of bids in the request increases. To understand this monotonicity, we note that the data shows the average and median timespan of requests with n bids is increasing in n. Therefore the more bids in the request, the longer the buyer has been waiting, which implies the cost to stop and choose an early bid is larger under the sunk cost fallacy.

5.4 Verification for the pq Model

The pq stopping rule directly captures both Observation 1 and 2. In most cases, especially when the number of bids currently received by a request is large, a new bid is not likely to be better than all previous bids. Then, buyers close their requests with probability q and the number of requests with n bids is roughly geometric in n. Moreover, all positions before the last have approximately the same probability to be chosen. However, in the special case when the new bid indeed surpasses all previous ones, the buyer is more likely to stop and accept the last bid. Thus, overall the last bids are more likely to be chosen.

We compute the expected values of $\pi(n, k)$ for the case that the total number of requests is 464, 668 (which is the number of requests in the data set). Let $p_1 = 0.4853$ be the empirical probability that a request is terminated at the first bid in the data set. Then, let $p_n = \hat{p}_n$ for $n \geq 2$ and $q = \hat{q}$.

[4]We note that throughout this paper (except for Section 5.2), q refers to the probability that a buyer stops the request when the most recent bid is not the best so far.

Denote by $\hat{r}(n, k)$ the expected number of requests that have n bids for which the kth bid is chosen. We compute $\hat{r}(n, k)$ using the formulas in Appendix B. Then, we can compute $\hat{R}(n) = \sum_{k=1}^{n} \hat{r}(n, k)$ and $\hat{\pi}(n, k) = \hat{r}(n, k)/\hat{R}(n)$. Figures 3 and 4 compare the estimated values with the values from the data.

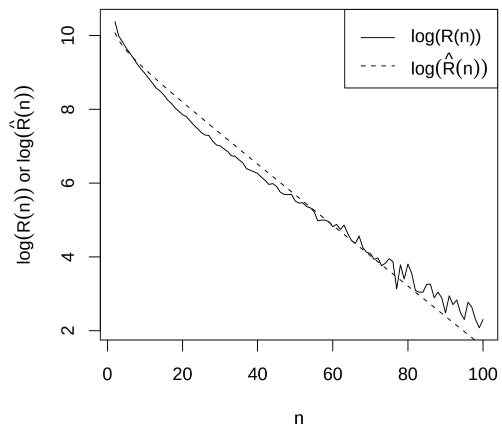

Figure 3: The expected number of requests decreases exponentially as a function of the number of bids n, both for the expected value $\log(\hat{R}(n))$ (dashed) under pq stopping rule and the original data $\log(R(n))$ (solid).

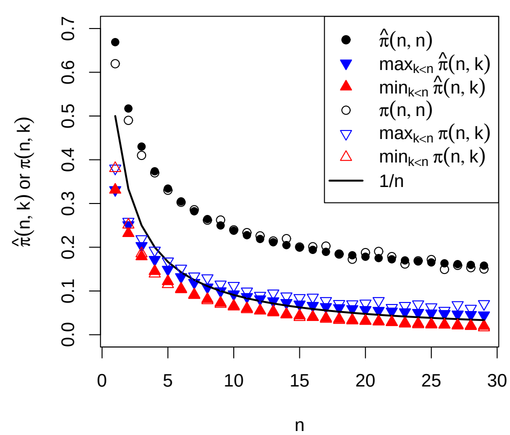

Figure 4: This figure shows the features of the expected value $\hat{\pi}(n, k)$ (solid), along with the corresponding values of $\pi(n, k)$ (hollow). The pq model captures the features of $\pi(n, k)$ very well.

In Figure 3, we can see that $\log(\hat{R}(n))$ is almost a straight line in n and is very close to $\log(R(n))$. In Figure 4, for any n, $\hat{\pi}(n, n)$ is significantly larger than $\max_{k<n} \hat{\pi}(n, k)$ and the difference between $\max_{k<n} \hat{\pi}(n, k)$ and $\min_{k<n} \hat{\pi}(n, k)$ is small. All these values are approximately the same as the observed values in original data. The only apparent difference is that $\max_{k<n} \hat{\pi}(n, k)$ is smaller than $\max_{k<n} \pi(n, k)$ when n is large. This is because we do not model the increasing trend of $\pi(n, k)$ for large n and k close to n when we estimate $\hat{p}_n$ and $\hat{q}$. Except for this, the pq model captures all the observations of Section 4.

6. DISCUSSION

In the previous section, we have seen that the pq stopping rule, a simple probabilistic model, captures the behavior of buyers on vWorker. The key characteristic of the pq stopping rule is that the buyer is more likely to stop when the most recent bid is the best so far.

This characteristic is reminiscent of the sunk cost fallacy. Consider that a buyer who has already spent time waiting for bids has indeed incurred some cost. A rational buyer considers this cost as sunk, ignoring it in her decision to continue or stop. An irrational buyer will however prefer not to close the request at a moment when the most recent bid is not the best so far, for she considers that the time elapsed since the best bid arrived would then be "wasted". Conversely, when the most recent bid is the best, she perceives that "it was worth waiting" because the sunk costs of waiting have been balanced off. Thus, buyers following the pq stopping are subject to the sunk cost fallacy, by having a higher probability to close a request when a new bid is the best so far. In the remainder of this section, we first discuss the weak increasing trend of $\pi(n,k)$ — which is not explained by the pq stopping rule — and then relate our results to the secretary problem.

6.1 The Weak Increasing Trend of $\pi(n,k)$

We have seen that even though the value of $\pi(n,k)$ is approximately the same for all $k < n$, it exhibits an increasing trend as k approaches n. In this section, we discuss two potential explanations for this trend.

First, this can be explained by the sunk cost fallacy. We mentioned before that buyers are less likely to stop if the most recent bid is not the best one because otherwise their waiting is "wasted". Similarly, the buyer may feel less of a sunk cost by accepting the penultimate bid than by accepting the first bid. This suggests that the buyer is more likely to select later bids than early bids. Interestingly, the weak increasing trend arises only for the last few bids (see Figure 2(top) and tests in Section 4), which suggests that the perceived sunk cost associated with a bid that arrived m bids ago is approximately constant for $m \geq 4$. We note that we could consider a generalization of the pq model, where the buyer stops with probability p_m when the best bid arrived m bids ago.

Another possible explanation for the weakly increasing trend in $\pi(n,k)$ is that Assumption 2 is violated. A buyer may be experiencing delay in observing new bids and, as a result, observe multiple new bids at the same time. In particular, even if a buyer is using the instantaneous notification feature of vWorker, she may be offline when some bids arrive (e.g., during the night or off-business hours). Then, the last time she checks the request (before she stops it), she may actually see several new bids at the same time. Essentially, any of these new bids is like the "last bid" for the buyer, a possibility that is not considered in the pq model. Hence, the violation of Assumption 2 increases the probability that the last few bids are chosen. The weak increasing trend of $\pi(n,k)$ in k can be explained by the fact that later bids are more likely to be observed at the same time with the last bid.

To verify this explanation, we study the interarrival times between bids. As an example, we now restrict attention to all requests with $n = 25$ bids (as in Figure 2(top)), while following results also hold for general n. We define Δ_k to be the time (in hours) between the arrival of the kth bid and the last bid. A small Δ_k suggests that the buyer may have observed the last $n-k$ bids at the same time. Table 1 shows the number of requests for which the kth ($k = 22, 23, 24$) bid is chosen under various constraints on Δ_k. The first column ($\Delta_k > 0$) shows all the requests for which the kth bid was chosen, and thus exhibits an increasing trend. However, as the constraint becomes more restrictive (i.e., as we move towards the columns to the right in Table 1), this trend gradually disappears. Moreover, we note that when $1 \leq k \leq 21$, the average number of requests with $n = 25$ bids and the kth bid chosen is 50, numbers in the last column of Table 1 have come back to this average level. These results suggests that the increasing trend of $\pi(n,k)$ in k arises because buyers see multiple new bids at the same time.

	$\Delta_k > 0$	$\Delta_k > 1$	$\Delta_k > 2$	$\Delta_k > 4$	$\Delta_k > 6$
k=22	70	60	57	50	48
k=23	85	72	65	62	48
k=24	111	80	69	55	46

Table 1: Number of requests with $n = 25$ bids for which the kth ($k = 22, 23, 24$) bid is chosen with different constraints on Δ_k.

6.2 Optimal Stopping Problems

Optimal stopping has been investigated in various stochastic decision contexts, and the secretary problem [6] might be the most famous example. In the classical secretary problem, a manager interviews sequentially n applicants for a single secretary position and would like to make the best choice. Applicants can be ranked from best to worst, yet this ranking is unknown to the interviewer, and the arrival order of the applicants is independent of their rankings. After each interview, the applicant is accepted or rejected. If this applicant is accepted, the whole process is terminated and this applicant get the position. If the applicant is rejected, she cannot be recalled. The goal of the process is to find the best secretary: the utility of the company is 1 if the chosen secretary is the best one, and 0 otherwise. The optimal stopping rule is to interview and reject first n/e applicants and then accept the next applicant who is better than all previous applicants.

We observe that the pq stopping rule resembles the optimal stopping rule in the secretary problem in two aspects. First, cost is incurred as the buyer explores the distribution of the bids. In the secretary problem, the interviewer passes the first n/e applicants at the cost that she may miss the best applicant, while in vWorker, buyers need to balance between higher waiting cost and better bids that may come later. Second, and more importantly, the buyer in vWorker (resp., the interviewer in the secretary problem) is more likely to (resp., only) stop the process when the new bid (resp., applicant) is better than all previous ones. This happens in the secretary problem since recall is not allowed. In vWorker where recall is allowed, it is the "sunk" waiting costs that reduce the probability that buyers choose early bids.

In experimental studies of the secretary problem, the optimal policy is rarely adopted. Decision makers tend to overestimate the absolute quality of the early applicants and stop the screening too early [3], giving insufficient consideration to late applicants. When recall is allowed, Zwick et al. [16] showed that decision makers also search too little in comparison to the optimal model, but they search too much if search costs also apply. Our results suggest that — because of the sunk cost fallacy — buyers in vWorker may search too much when good bids arrive early on.

7. ACKNOWLEDGEMENTS

We gratefully acknowledge Ian A. Ippolito, CEO of Exhedra Solutions, Inc, the parent company of vWorker, for providing the dataset. We also thank the anonymous reviewers for their insightful suggestions.

8. REFERENCES

[1] H. R. Arkes and C. Blumer. The psychology of sunk cost. *Organizational Behavior and Human Decision Processes*, 35(1):124–140, Feb. 1985.

[2] N. Augenblick. Consumer and producer behavior in the market for penny auctions: A theoretical and empirical analysis. Working Paper, 2011.

[3] J. N. Bearden, A. Rapoport, and R. O. Murphy. Sequential observation and selection with Rank-Dependent payoffs: An experimental study. *Management Science*, 52:1437–1449, Sept. 2006.

[4] J. Brockner. The escalation of commitment to a failing course of action: Toward theoretical progress. *The Academy of Management Review*, 17(1):39–61, Jan. 1992.

[5] D. DiPalantino and M. Vojnovic. Crowdsourcing and all-pay auctions. In *Proceedings of the 10th ACM conference on Electronic commerce*, EC '09, pages 119–128, 2009.

[6] P. R. Freeman. The secretary problem and its extensions: A review. *International Statistical Review / Revue Internationale de Statistique*, 51(2):189–206, 1983.

[7] D. Gefen and E. Carmel. Is the world really flat? a look at offshoring at an online programming marketplace. *MIS quarterly*, 32(2):367–384, 2008.

[8] J. J. Horton and L. B. Chilton. The labor economics of paid crowdsourcing. In *Proceedings of the 11th ACM conference on Electronic commerce*, EC '10, pages 209–218, 2010.

[9] D. Kahneman and A. Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291, Mar. 1979.

[10] A. Kittur, E. H. Chi, and B. Suh. Crowdsourcing user studies with mechanical turk. In *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, CHI '08, pages 453–456, 2008.

[11] J. Z. Rubin and J. Brockner. Factors affecting entrapment in waiting situations: The Rosencrantz and Guildenstern effect. *Journal of Personality and Social Psychology*, 31(6):1054–1063, June 1975.

[12] B. M. Staw. Knee-deep in the big muddy: a study of escalating commitment to a chosen course of action. *Organizational Behavior and Human Performance*, 16(1):27–44, June 1976.

[13] A. Tversky and D. Kahneman. Loss aversion in riskless choice: A Reference-Dependent model. *The Quarterly Journal of Economics*, 106(4):1039–1061, Nov. 1991.

[14] G. Whyte. Escalating commitment in individual and group decision making: A prospect theory approach. *Organizational Behavior and Human Decision Processes*, 54(3):430–455, Apr. 1993.

[15] J. Yang, L. A. Adamic, and M. S. Ackerman. Crowdsourcing and knowledge sharing: strategic user behavior on taskcn. In *Proceedings of the 9th ACM conference on Electronic commerce*, EC '08, pages 246–255, 2008.

[16] R. Zwick, A. Rapoport, A. K. C. Lo, and A. V. Muthukrishnan. Consumer sequential search: Not enough or too much? *Marketing Science*, 22(4):503–519, Oct. 2003.

APPENDIX

A. THRESHOLD STRATEGY

Our deterministic model is based on the following assumptions.

Assumption 3. *The utility of each bid is drawn from a distribution F.*

Assumption 4. *For each request, at t units of time after the request is posted, bid arrival process follows a Poisson process with parameter $\lambda(t)$.*

Assumption 5. *Each buyer incurs a waiting cost $c(t)$ if t time units elapse until she selects the best bid.*

First we show that the threshold strategy is the optimal strategy under the assumptions that F is the uniform distribution on $[a, b]$, $\lambda(t) = \lambda$, and $c(t) = ct$.

Consider an experienced buyer who knows or has a very good estimate of the bid utility distribution F. Suppose after n bids, the current minimal utility of the bids is α and maximal utility is β. Since the waiting cost is constant c, and the Poisson process is memoryless with mean interarrival time λ, the expected waiting cost until the next bid arrives is always equal to $c\lambda$.

The expected maximum utility at the time the next bid comes is

$$\beta \cdot \frac{\beta - a}{b - a} + \int_\beta^b \frac{x}{b-a}\, dx = \frac{1}{b-a}\left(\frac{1}{2}b^2 + \frac{1}{2}\beta^2 - a\beta\right) \quad (3.1)$$

and the expected benefit of waiting is

$$\frac{1}{b-a}\left(\frac{1}{2}b^2 + \frac{1}{2}\beta^2 - a\beta\right) - \beta = \frac{(b-\beta)^2}{2(b-a)} \quad (3.2)$$

Therefore, the buyer's myopic decision, based on the comparison between $\frac{(b-\beta)^2}{2(b-a)}$ and $c\lambda$ is to terminate the request if and only if she observes a (new) bid with utility $\beta > \beta^*$ where the threshold $\beta^* = b - \sqrt{2c\lambda(b-a)}$. Since this threshold is independent of t, it is the optimal myopic strategy at any time. Moreover, $\frac{(b-\beta)^2}{2(b-a)} - c\lambda$ is decreasing in β, hence this rule is not only myopic optimal but also a global optimal termination rule. This stoping rule has the following properties:

(1). This is a threshold stopping rule based on the current maximum bid quality β.

(2). The buyer always accepts the last bid.

(3). Since the bid comes independently, a buyer stops at the arrival of each bid with constant probability $\rho = \beta^*/b$, which is actually dependent on a, b, c and λ.

This outcome of the threshold strategy resembles Observation 2(a) in that buyers only accept the last bid. Also, buyers close the request with a constant probability ρ at the arrival of each new bid, which implies that the number of requests with n bids follows the geometric distribution in n. However, the threshold strategy cannot explain our observation that buyers choose any bid prior to the last one with approximately the same probability. And more importantly, are there buyers who always chooses the last bid?

Clearly, not all buyers always choose the last bid, since not all requests end up with last bid chosen. Our assumption that buyers are balancing bid utility and waiting cost suggests that buyers are well informed of the parameters of the auction, like the distribution of bid utilities, arrival rates, and so on. These parameters are usually not directly accessible but may be learned through a number of requests. Therefore, experienced buyers are most likely to use the threshold strategy. We study buyers with at least 200 requests, and considering that any experienced buyers may be inexperienced at the beginning, for each of them we focus on those requests after this buyer's first 100 requests. We compute the frequencies that these buyers choose the last bid in their post-100th requests, and the results are shown in the following figure.

The above figure summarizes the discussion by showing the histogram of the frequencies of choosing the last bid among experienced buyers. The histogram has a peak around

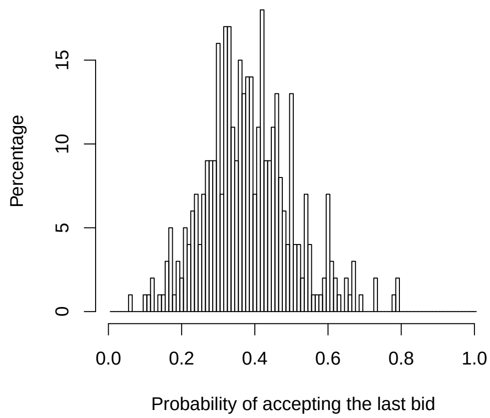

Figure 5: The histogram of the frequencies of choosing the last bid for buyers with at least 200 requests. There is no experienced buyer that chooses the last bid with frequency higher than 0.8.

0.3 to 0.4. Besides, among 347 buyers that have at least 200 requests, none of them choose the last bid with probability higher than 80%. Thus even the most experienced buyers do not use the threshold strategy.

Other cases of the deterministic models can be explored by assuming different distributions F for bids, time-dependent rates for the Poisson arrival process, and waiting costs which are not constant. In general, it is not possible to determine analytically an optimal strategy. Thus, we use numerical simulations to study the outcomes of each combination of parameters. We found that the deterministic models can often capture one or two out of the three features described in Observation 1 and 2 but fail in the other(s). The performance of the deterministic model is not satisfactory unless we introduce more parameters, which makes the model less general and thus less convincing. Even though there may exist models of rationality that explain the observations, the explanatory power of such models will be weak (per unit of complexity).

B. ESTIMATION OF p AND q FOR THE pq STOPPING RULE

Consider an arbitrary request, let random variable N be the number of the bids it receives and random variable B be the index of the bid that is chosen. $B = 1$ means the first bid is chosen and $B = N$ means the last one is chosen. Recall the pq stopping rule: if the most recent bid is the best bid, stop the request with probability p, otherwise stop with probability q. The probability that the kth bid is the best one among first k bids is $1/k$, therefore the probability that the request is not terminated at the kth bid given it is not terminated before the kth bid is

$$f_k \triangleq P(N > k | N \geq k) = \frac{1-p}{k} + \frac{k-1}{k}(1-q) = (1-q) - \frac{p-q}{k} \quad (3)$$

We assume that if $p - q$ is small, then $f_k \approx 1 - q$ and $\lim f_k = 1 - q$ as k goes to infinity. We will come back to this assumption later.

Consider the event $\{N = n, B = b\}$, i.e., the buyer terminates the request at the nth bid and chooses the bth ($b \leq n$) bid. In that case, the bth bid is the best one in the first n bids, and

$$P(N = n, B = b) = \begin{cases} n^{-1} \cdot f_1 \cdots f_{b-1} \cdot (1-p)(1-q)^{n-b-1} q & \text{, if } b < n \\ n^{-1} \cdot f_1 \cdots f_{n-1} \cdot p & \text{, if } b = n \end{cases} \quad (4)$$

Therefore we have

$$\frac{P(N = n, B = b)}{P(N = n, B = b-1)} = \begin{cases} f_{b-1}/(1-q) & \text{, if } b < n \\ f_{n-1} p/(1-p) q & \text{, if } b = n \end{cases} \quad (5)$$

With the assumption of f_k, the above ratio is approximately 1 when $b < n$ and $(1-q)p/(1-p)q$ when $b = n$.

We note that this result is similar to Observation 2: given the request is terminated at the nth bid, the probabilities that the bth bid is chosen are approximately the same for all $b < n$ and are significantly lower than the probability that the last bid is chosen. Further more, we note that (3) implies a request is terminated with more than k bids with probability $P(N > k) = \prod_{i=1}^{k} f_i$. Since $f_k \approx 1 - q$ as for all k, $P(N > k)$ is approximately exponential in k and N follows a geometric distribution.

Given (5), we can estimate parameters p and q by

$$\begin{cases} \dfrac{P(N = n, B = n)}{P(N = n, B = n-1)} \approx \dfrac{(1-q)p}{(1-p)q}, \ n \geq 2 \\ P(N \geq k) \approx (1-q)^{k-2}, \ k \geq 2 \end{cases} \quad (6)$$

Specifically, to estimate q, we use $\#(N \geq k)/\#(N \geq 2)$ as an estimate for $P(N \geq k)$, where $\#(\mathrm{A})$ is the count of event A in our data set. (6) implies $\log(P(N \geq k))$ is linear in k with zero intercept and slope $\log(1-q)$. The linear regression gives

$$\log \frac{\#(N \geq k)}{\#(N \geq 2)} = -0.06588 - 0.08138 \cdot n, \quad 2 \leq n \leq 100 \quad (7)$$

Therefore we get $\log(1 - \widetilde{q}) = -0.08138$ and the estimated $\widetilde{q} = 0.0782$.

The estimated $\widetilde{p}$ can also be obtained by (6). To reduce the variance of our estimation, we use $\#(N = n, B < n)$ instead of $\#(N = n, B = n-1)$ because it has more samples. Since

$$\begin{aligned} P(N = n) &= f_1 \cdots f_{n-1}(1 - f_n) \\ P(N = n, B = n) &= n^{-1} f_1 \cdots f_{n-1} p \\ P(N = n, B = n-1) &= n^{-1} f_1 \cdots f_{n-2}(1-p) q \end{aligned}$$

then

$$\frac{P(N = n, B = n-1)}{P(N = n, B < n)} = \frac{1-p}{(n-1) f_{n-1}}$$

and

$$\begin{aligned} \frac{\# \{N = n, B = n\}}{\# \{N = n, B < n\}} &= \frac{P(N = n, B = n)(1-p)}{P(N = n, B = n-1)(n-1) f_{n-1}} \\ &= \frac{p}{(n-1)q} \end{aligned}$$

Therefore, given the estimated $\widetilde{q}$, we have

$$\widetilde{p} = (n-1)\widetilde{q} \cdot r_n \quad (8)$$

where $r_n = \# \{N = n, B = n\} / \# \{N = n, B < n\}$

However, the data shows that r_n's are different for different n, which implies p is not a constant but a function of n. Define p_n as the probability that a buyer terminates the request when the request has n bids and the nth bid is better than all previous ones. Then we can approximate p_n from (8). It turns out that the estimated probability $\widetilde{p}_n$ is approximately linearly increasing in n, with the following relation

$$\widetilde{p}_n = 0.13876 + 0.00956n, \quad 2 \leq n \leq 30.$$

This linear relation obviously does not hold for large n since $\widetilde{p}_n$ is upper bounded by 1, yet it approximates the observed $\widetilde{p}_n$ well when n is not too large, and the R^2 of the above regression is 0.95.

Finally, we go back to the assumption that $f_k \approx 1 - q$. Since the estimated $\widetilde{p} = \widetilde{p}_n = 0.13876 + 0.00956n$ is roughly linear in n, $(p - \widetilde{q})/n$ is approximately a nonzero constant. Therefore we need to adjust our estimates $\widetilde{p}$ and $\widetilde{q}$. It follows from (3) and (6) that the adjusted estimate $\hat{q}$ is still a constant. Then we can assume the adjusted estimate $\hat{q} = \lambda\widetilde{q}$, then the adjusted $\hat{p}_n$ satisfies

$$\frac{\hat{p}_n}{(n-1)\hat{q}} = r_n = \frac{\widetilde{p}_n}{(n-1)\widetilde{q}}$$

and therefore $\hat{p}_n = \lambda\widetilde{p}_n = 0.13876\lambda + 0.00956\lambda n$. Plug this into 3, we got

$$1 - \widetilde{q} \approx f_n = (1 - \hat{q}) - \frac{1}{n}(\hat{p}_n - \hat{q}) \approx 1 - \lambda\widetilde{q} - 0.00956\lambda$$

Thus $\lambda = 0.89$ and therefore

$$\hat{q} = \lambda\widetilde{q} = 0.0696, \quad \hat{p}_n = \lambda\widetilde{p}_n = 0.1236 + 0.0085n$$