MDP Project: One-shot Secretary Problem with Recall and Uncertain Employment

Connor Colombe

April 27, 2023

Abstract

In this paper we present and analyze a variant of the classical secretary problem: The one-shot secretary problem with recall and uncertain employment. In this variant, our objective is still to obtain the best candidate, but the structure varies in two important ways: When we send an offer to a candidate, whether they accept it or not is subject to a probability q that is a function of how recently we interviewed that candidate. We may send an offer to any candidate we have seen so far but we only have a single offer to send over the entire problem. We analyze this problem and present the optimal policy as well as asymptotic results.

1 Introduction

The classical Secretary problem is as follows: Suppose there are N people applying for a job at a company and each has a unique ranking $v \in [N]$. The order in which the candidates are to be interviewed is a uniformly random permutation. The candidates are then interviewed one by one. After each interview, a candidates *relative* rank amongst the previously interviewed candidates is revealed. The company can then hire them on the spot or reject them. Once rejected, the candidate cannot be hired in the future. The company is concerned only with hiring the person with the absolute lowest rank of the N candidates. The problem is then to determine a strategy that maximizes the probability of hiring the absolute lowest rank candidate.

In this paper we present an interesting modification of the secretary problem. After any k of the N candidates have been interviewed already, the decision maker may now extend an offer to any one of the k previously interviewed candidates. But, the decision maker only has the capacity to extend a *single* offer for the duration of the interview process (for fear of seeming desperate). However, there is the additional assumption that the offer is subject to some probability of acceptance by the candidate $q(r)$. Here r is the number of time-periods since the courted candidate was interviewed. We make the assumptions that $q(0) = 1$ (if the candidate is offered the job immediately after their interview, they accept) and $q(\cdot)$ is a non-increasing function (the longer we wait to offer someone a job, they are not more likely to accept). We will now give a full MDP formulation of our problem and then dive into analysis.

2 Problem and Model

Let the time indices of our model be $k \in [N]$, one for each potential candidate interviewed. Our state space be given by the pair $(k, r) \in \mathbb{Z}_+ \times \mathbb{Z}_+$ where k is the number of candidates interviewed so far and $r \in \{0, 1, \ldots, k-1\}$ is the the number of interviews that have occurred since interviewing the current relative rank one candidate. In state (k, r), we have two possible actions: send an offer the candidate of current lowest relative rank or to not send an offer and interview the next candidate. The reward for offering a job and the absolute lowest rank candidate accepting it is one and all other rewards are zero.

Based on this description, we can write down our value function. Let $J_k(r)$ be defined as the probability of obtaining the best candidate given we are in state (k, r). It is easy to reason that

$$J_N(r) = q(r).$$

For all other states, we have the option of sending an offer or waiting, which implies

$$J_k(r) = \max\left\{\frac{k}{n}q(r), \frac{1}{k+1}J_{k+1}(0) + \frac{k}{k+1}J_{k+1}(r+1)\right\} \quad \forall k \in [n-1] \tag{1}$$

where the first term represents sending an offer to the current relative best candidate. There is a k/n probability that this candidate is the true best candidate and the probability they accept is $q(r)$. In the second term, we have the case that the relative best candidate is the next one interviewed and the case where they are not, in which the relative best candidate is the same and it has now been $r + 1$ time periods since they were interviewed. For the purposes of this project, we will assume that $q(r) = p^r$ where $p \in [0, 1]$ is some constant.

The remainder of the paper will consist of presenting and proving the form of the optimal policy based on p, deriving asymptotic policy results, and highlighting interesting numerical results that the theory predicts.

3 Analysis

We will break our analyses into a series of cases based on the value of p. The first two analysis are simple.

3.1 Case 1: $p = 0$

Here we have the form of our $q(r)$ function is

$$q(r) = \begin{cases} 1 & r = 0 \\ 0 & \text{else.} \end{cases}$$

In this case, we get the classical secretary problem. To see why, note that we can only ever ask the person that was most recently interviewed (since every previous person will accept an offer with probability $p = 0$), and we can only ask one person. Moreover, since our objective is to maximize the probability of finding the overall best candidate, this is exactly the formulation of the classical secretary problem.

It is well known that the optimal policy in the classic problem is to interview some number of people $f(N)$ during a "sampling phase", and then offer the job to the next person who is interviewed that is better than the previous best. Asymptotically, we have that $f(N) \to N/e$ and the probability of successfully acquiring the best candidate tends to $1/e$.

In Figures 1 and 2 we plot the optimal policy and the probability of winning in each state for $N = 100$. Here N is large enough to see the asymptotic policy result. In figure 1 we see that the decision maker waits until after they have observed N/e people before considering sending any offers.

3.2 Case 2: $p = 1$

The other straightforward case is when $p = 1$. In this case, previous candidates will always accept an offer. The intuitive policy that comes to mind here is to wait until all candidate have been interviewed and then offer the job to the, now revealed, best candidate. This strategy wins with probability 1. We can show this this using Bellman's equations (1) and induction.

Lemma 1 (Optimally of Waiting)**.** *Let $p = 1$ and Bellman's equations be given by (1). Then for any state (k, r) where $k < N$ the optimal policy is to wait. In state (N, r), the optimal policy is to send an offer to the candidate that was interviewed r time periods ago, which is the best candidate.*

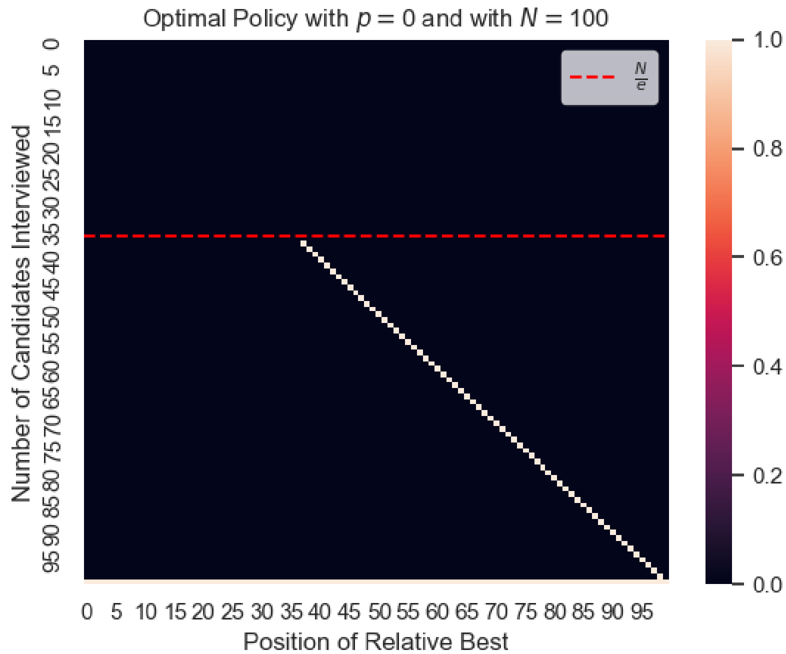

Figure 1: The optimal policy for each state (k, r). Here a white square means to send an offer and a black square means to wait.

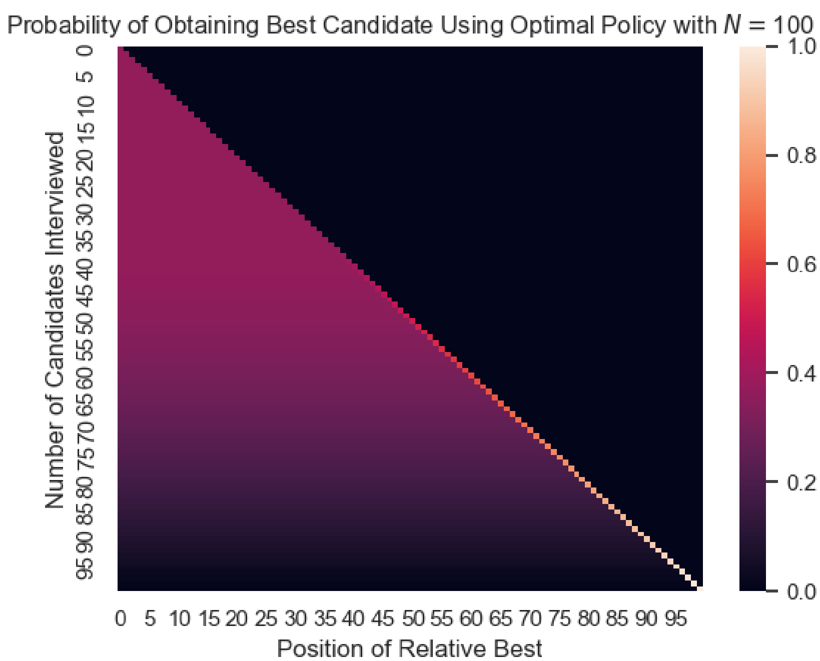

Figure 2: The the probability of winning from each state in the classic variant of the problem.

Proof. When $k = N$, Bellman's equations imply that $q(r) = p^r = 1$ for all r and thus $J_N(r) = 1$ for all r. In this case, we send an offer to the best candidate and obtain them with certainty. Now for the inductive step. We assume that $J_k(r) = 1$ for all $r \in \{0, \dots, k-1\}$. Now suppose we are in state $(k-1, r)$. We are faced with the maximization problem

$$J_{k-1}(r) = \max\left\{\frac{k-1}{N}, \quad \frac{1}{k}J_k(0) + \frac{k-1}{k}J_k(r+1)\right\}$$

By our inductive hypothesis, this simplifies to

$$\begin{aligned} J_{k-1}(r) &= \max\left\{\frac{k-1}{N}, \quad \frac{1}{k}J_k(0) + \frac{k-1}{k}J_k(r+1)\right\} \\ &= \max\left\{\frac{k-1}{N}, \quad \frac{1}{k}(1) + \frac{k-1}{k}(1)\right\} \\ &= \max\left\{\frac{k-1}{N}, \quad 1\right\} \end{aligned}$$

which for all r implies that the optimal policy is to wait to send an offer and that $J_{k-1}(r) = 1$ completing the proof. □

Note that lemma 1 implies that $J_1(0) = 1$ which is the probability of obtaining the best candidate over the entire game. We have included the experimental results that come from solving Bellman's equations via backwards induction for $N = 100$ in Figures 3 and 4. The experimental results match those predicted by the theory and our intuition.

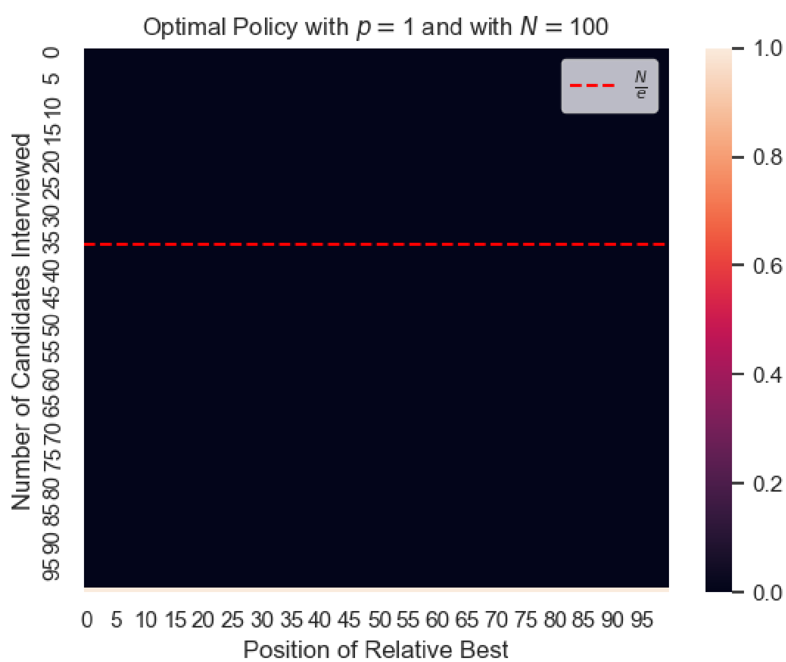

Figure 3: The optimal policy for each state (k, r). Here a white square means to send an offer and a black square means to wait. If you look closely, you can see the final row is colored in.

3.3 Case 3: $0 < p < 1$

Now we move onto the non-trivial instances of the problem. As with most finite horizon MDP's we begin to analyze the problem via backwards induction. We will explicitly work out the first few cases to give a sense of what form the optimal policy takes and then generalize and prove the results.

In the base case, we are in state (N, r) for $r \in \{0, \ldots, N-1\}$. We have that

$$J_N(r) = p^r$$

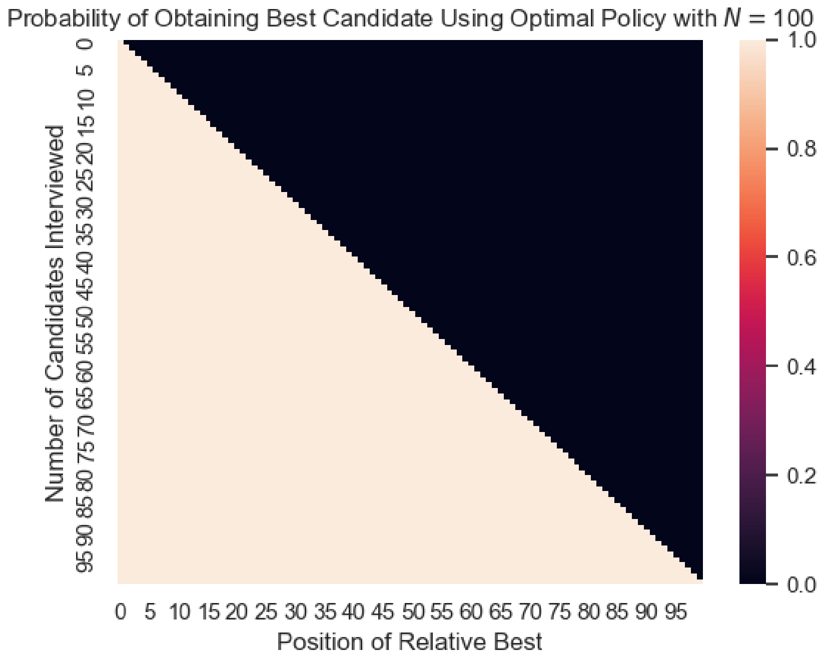

Figure 4: The the probability of winning from each state in the classic variant of the problem.

for all r. The intuition here is, if we have seen all the candidates, then we know where the true best candidate is. Therefore, we send them an offer regardless of when we interviewed them. If $k = N - 1$, then Bellman's equation (1) suggests we face the optimization problem

$$
\begin{aligned}
J_{N-1}(r) &= \max\left\{\frac{N-1}{N}p^r, \frac{1}{N}J_N(0) + \frac{N-1}{N}J_N(r+1)\right\} \\
&= \max\left\{\frac{N-1}{N}p^r, \frac{1}{N} + \frac{N-1}{N}p^{r+1}\right\}.
\end{aligned}
$$

From this, we can deduce that we will send an offer to the current best candidate if and only if

$$
\frac{N-1}{N}p^r > \frac{1}{N} + \frac{N-1}{N}p^{r+1}
$$

which, solving for r, yields the inequality

$$
r \le \log_p\left(\frac{1}{1-p}\cdot\frac{1}{N-1}\right) = r^*_{N-1}. \tag{2}
$$

Here r^*_{N-1} is the threshold r for which we send an offer or not in state $(N-1, r)$, the intuition here is that if r is too large, we are better off "being optimisitic" that the true best candidate is yet to come and we can procure them with a high probability rather than taking a low probability shot at the current best candidate, regardless of the large probability that they are the best candidate. What we have shown here is that there is some threshold r for which the policy swaps from "send an offer" to "wait". Based on this, we have that

$$
J_{N-1}(r) = \begin{cases} \frac{N-1}{N}p^r & r \le r^*_{N-1} \\ \frac{1}{N} + \frac{N-1}{N}p^{r+1} & \text{else} \end{cases} \tag{3}
$$

Now we consider the case $k = N - 2$. Once more, from (1) we have that

$$\begin{aligned} J_{N-2}(r) &= \max\left\{\frac{N-2}{N}p^r, \frac{1}{N-1}J_{N-1}(0) + \frac{N-2}{N-1}J_{N-1}(r+1)\right\} \\ &= \max\left\{\frac{N-2}{N}p^r, \frac{1}{N} + \frac{N-2}{N-1}J_{N-1}(r+1)\right\}. \end{aligned}$$

We can analyze this in two cases. The case where $r \leq r^*_{N-1} - 1$ and when $r > r^*_{N-1} - 1$. The rationale behind splitting our analysis into these two particular cases is that: in each, the functional form of Bellman's equations (1) will vary.

3.3.1 Case 1: $r > r^*_{N-1} - 1$

Here Bellman's equations give us

$$\begin{aligned} J_{N-2}(r) &= \max\left\{\frac{N-2}{N}p^r, \frac{1}{N-1}J_{N-1}(0) + \frac{N-2}{N-1}J_{N-1}(r+1)\right\} \\ &= \max\left\{\frac{N-2}{N}p^r, \frac{1}{N} + \frac{N-2}{N-1}\left(\frac{1}{N} + \frac{N-1}{N}p^{r+2}\right)\right\}. \end{aligned}$$

Similarly to before, we can deduce that we should send an offer to the current best candidate if and only if

$$p^r(1-p^2) \geq \frac{1}{N-1} + \frac{1}{N-2}$$

which upon solving for r yields $r \leq \log_p\left(\frac{1}{1-p^2}\left(\frac{1}{N-1} + \frac{1}{N-2}\right)\right)$. How does this compare to our previous value of r^*_{N-1}? It turns out, that for any $N > 2$ and $p \in (0, 1)$, this is always less than r^*_{N-1}.

3.3.2 Case 2: $r \leq r^*_{N-1} - 1$

Now the functional form of Bellman's equation becomes

$$\begin{aligned} J_{N-2}(r) &= \max\left\{\frac{N-2}{N}p^r, \frac{1}{N-1}J_{N-1}(0) + \frac{N-2}{N-1}J_{N-1}(r+1)\right\} \\ &= \max\left\{\frac{N-2}{N}p^r, \frac{1}{N} + \frac{N-2}{N-1}\left(\frac{N-1}{N}p^{r+1}\right)\right\} \\ &= \max\left\{\frac{N-2}{N}p^r, \frac{1}{N} + \frac{N-2}{N}p^{r+1}\right\}. \end{aligned}$$

Which as before, we can manipulate to obtain the maximum r for which we will send an offer, and we find

$$r \leq \log_p\left(\frac{1}{1-p} \cdot \frac{1}{N-2}\right).$$

Furthermore comparing this to (2), we can see it greater than r^*_{N-1}. This implies that for all $r \leq r^*_{N-1} - 1$, it is optimal to send an offer. Putting the results of these two cases together, we find that the policy threshold r^*_{N-2} is given by

$$r^*_{N-2} = \log_p\left(\frac{1}{1-p^2}\left(\frac{1}{N-1} + \frac{1}{N-2}\right)\right).$$

At this point we have a few conjectures. Firstly, we conjecture that r^*_{N-k} is a decreasing sequence in k. Secondly, we conjecture that the functional form of r^*_{N-k} is $r^*_{N-k} = \log_p\left(\frac{1}{1-p^k} \cdot \sum_{i=1}^{k} \frac{1}{N-i}\right)$. Before diving into these proofs lets look at how these conjectures hold up to a few numerical experiments (see Figure 5).

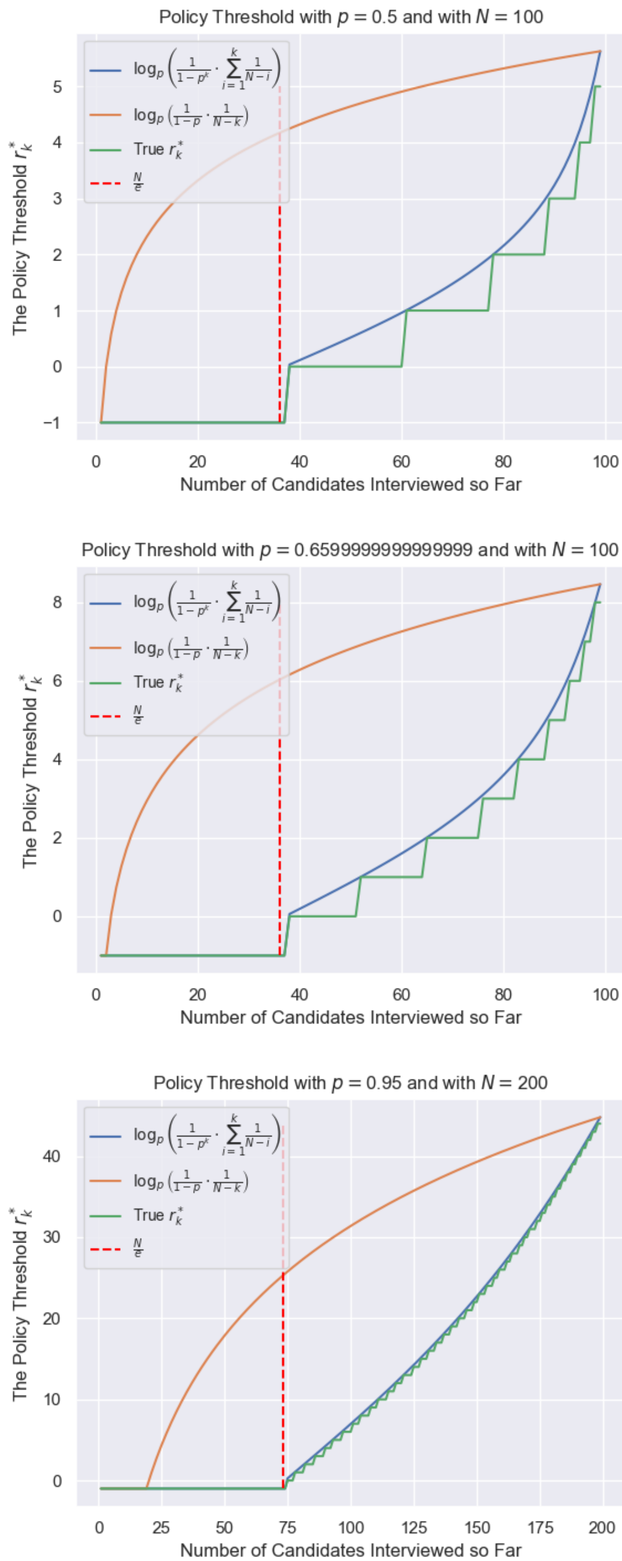

Figure 5: Experimental results examining at how our conjectures hold up to the exact numerical results.

Based on these results, we have strong evidence for our conjecture that $r^*_{N-k} = \log_p \left(\frac{1}{1-p^k} \cdot \sum_{i=1}^{k} \frac{1}{N-i} \right)$, (a corollary of which is that r^*_{N-k} is decreasing in k) we can now turn our conjectures into propositions and prove them! We will begin with a seemingly unrelated but helpful lemma.

Lemma 2. *Let k be a positive integer less than N and $\alpha \in (0,1)$. If $p < \frac{N\alpha}{N\alpha+1}$, then*

$$\log_p\left(\frac{1}{1-p}\cdot\frac{1}{N-k}\right) > \log_p\left(\frac{1}{1-p^k}\cdot\sum_{i=1}^{k}\frac{1}{N-i}\right) \tag{4}$$

for all $1 \leq k \leq N(1-\alpha)$

Proof. Taking (4), we can exponentiate each side ($p \in (0,1)$ so the inequality flips) to obtain the inequality

$$\frac{1}{1-p^k}\cdot\sum_{i=1}^{k}\frac{1}{N-i} > \frac{1}{1-p}\cdot\frac{1}{N-k}.$$

Rearranging , we get

$$(N-k)\sum_{i=1}^{k}\frac{1}{N-i} > \frac{1-p^k}{1-p}$$

We can now expand both sides since the RHS has the form of a geometric series. Now we have

$$1 + \frac{N-k}{N-k+1} + \cdots + \frac{N-k}{N-k+(k-1)} > 1 + p + \cdots p^{k-1}. \tag{5}$$

When does this inequality hold? Well we are comparing a harmonic-ish series to a geometric one, the terms in the harmonic series ought to get smaller "slower" than the terms in the geometric series which decay at a constant rate. By this logic, if the fist non-one term on the LHS is larger than the corresponding term on the RHS, then every term on the LHS will be larger than its corresponding term on the RHS and the inequality will hold. When does this happen?

We need to solve the inequality $\frac{N-k}{N-k+1} > p$. Solving for k yields $k < N - p/(1-p)$, and then plugging in $p < \frac{N\alpha}{N\alpha+1}$ implies that $\frac{N-k}{N-k+1} > p$. From here, we claim that $\frac{N-k}{N-k+m} > p^m$. We can show this my induction on m.

We already covered the base case, so suppose $\frac{N-k}{N-k+m} > p^m$. Multiply both sides by p we have

$$\frac{N-k}{N-k+m}p > p^{m+1}. \tag{6}$$

We can write

$$\frac{N-k}{N-k+m}\alpha = \frac{N-k}{N-k+m+1} \implies \alpha = \frac{N-k+m}{N-k+m+1}.$$

Using our base case, we have $p < \frac{N-k}{N-k+1} < \frac{N-k+m}{N-k+m+1} = \alpha$. Combining this with equation (6), yields

$$\frac{N-k}{N-k+m+1} = \frac{N-k}{N-k+m}\alpha > \frac{N-k}{N-k+m}p > p^{m+1} \tag{7}$$

proving the inequality (5) for $p < \frac{N\alpha}{N\alpha+1}$ and since all the prior steps were reversible, completing the proof.

□

The motivation behind this lemma is that we will have different potential candidates for r^*_{N-k} for each k and this lemma will allow us to easily determine the correct value of r^*_{N-k} for the relevant k. One important observation before moving onto the main results of the paper. The lemma above was conditioned on $p < \frac{N\alpha}{N\alpha+1}$. Note that for fixed $\alpha \in (0,1)$, for sufficiently large N, this can be made arbitrarily close to 1. That is, if we assume N is large then we can set α close to 0 and have the result hold for nearly all $1 \leq k < N$ and any almost any p.

Now for another helpful lemma!

Lemma 3. *Let $p \in (0,1)$, $N \geq 4$ be a positive integer and $1 \leq k \leq N-2$ be an integer. The function*

$$f(k) = \log_p \left(\frac{1}{1-p^k} \cdot \sum_{i=1}^{k} \frac{1}{N-i} \right)$$

is strictly decreasing in k.

Proof. We want to show that for any integral $k \in [1, N-2]$, we have that

$$\log_p \left(\frac{1}{1-p^k} \cdot \sum_{i=1}^{k} \frac{1}{N-i} \right) > \log_p \left(\frac{1}{1-p^{k+1}} \cdot \sum_{i=1}^{k+1} \frac{1}{N-i} \right)$$

If we exponentiate each side we get the inequality

$$\frac{1}{1-p^{k+1}} \cdot \sum_{i=1}^{k+1} \frac{1}{N-i} > \frac{1}{1-p^k} \cdot \sum_{i=1}^{k} \frac{1}{N-i}.$$

We can further manipluate both sides in the following ways

$$\begin{aligned}
\frac{1}{1-p^{k+1}} \cdot \sum_{i=1}^{k+1} \frac{1}{N-i} &> \frac{1}{1-p^k} \cdot \sum_{i=1}^{k} \frac{1}{N-i} \\
(1+p+\cdots+p^{k-1}) \sum_{i=1}^{k+1} \frac{1}{N-i} &> (1+p+\cdots+p^k) \sum_{i=1}^{k} \frac{1}{N-i} \\
(1+p+\cdots+p^{k-1}) \left(\sum_{i=1}^{k} \frac{1}{N-i} + \frac{1}{N-k-1} \right) &> (1+p+\cdots+p^{k-1}+p^k) \sum_{i=1}^{k} \frac{1}{N-i} \\
(1+p+\cdots+p^{k-1}) \frac{1}{N-k-1} &> p^k \sum_{i=1}^{k} \frac{1}{N-i} \\
\frac{1+p+\cdots+p^{k-1}}{p^k} &> (N-k-1) \sum_{i=1}^{k} \frac{1}{N-i} \\
\frac{1}{p^k} + \frac{1}{p^{k-1}} + \cdots \frac{1}{p} &> \frac{N-k-1}{N-k} + \frac{N-k-1}{N-k+1} + \cdots + \frac{N-k-1}{N-1}
\end{aligned}$$

Now notice that there are k terms on either side. Each term on the LHS is greater than one since $p \in (0,1)$ and each term on the RHS is less than one since the denominator is larger. Since each step was reversible, the initial inequality holds and $f(k)$ is a decreasing function, as desired. □

We will now state and prove both

- that there exists a threshold r^*_{N-k} for which $r \leq r^*_{N-k}$ implies the optimal action is the send an offer to the current best candidate, otherwise the optimal action is to wait
- the functional form of $J_{N-k}(r)$ for relevant k.

Proposition 1. *Consider the modified Secretary problem with N people and probability $p \in (0,1)$. Let $\alpha \in (0,1)$ and define*

$$r_{N-k} = \log_p \left(\frac{1}{1-p^k} \cdot \sum_{i=1}^{k} \frac{1}{N-i} \right). \tag{8}$$

If $p < \frac{N\alpha}{N\alpha+1}$ and $1 \leq k \leq N(1-\alpha)$ then $r^*_{N-k} = r_{N-k}$. Moreover if $r^*_{N-k+1} \geq 0$, the functional form of J_{N-k} is given by

$$J_{N-k}(r) = \begin{cases} \frac{N-k}{N} p^r & r < r^*_{N-k} \\ \frac{N-k}{N} \left(p^{r+k} + \sum_{i=1}^{k} \frac{1}{N-i} \right) & \text{else.} \end{cases} \tag{9}$$

Proof. We will proceed by induction on k. We have already seen in equation (2) that the threshold for $k = 1$ is given by $r^*_{N-1} = \log_p\left(\frac{1}{1-p} \cdot \frac{1}{N-1}\right)$. We also found that $r^*_N = N - 1$ and that

$$J_{N-1}(r) = \begin{cases} \frac{N-1}{N} p^r & r \leq r^*_{N-1} \\ \frac{1}{N} + \frac{N-1}{N} p^{r+1} & \text{else} \end{cases} \tag{10}$$

completing the base case. Now suppose we are given that $r^*_{N-k+1} = r_{N-k+1} \geq 0$ and that

$$J_{N-k+1}(r) = \begin{cases} \frac{N-k+1}{N} p^r & r < r^*_{N-k+1} \\ \frac{N-k+1}{N} \left(p^{r+k-1} + \sum_{i=1}^{k-1} \frac{1}{N-i}\right) & \text{else.} \end{cases}$$

Consider Bellman's equations for the situation where we have interviewed $N - k$ people. We have

$$J_{N-k}(r) = \max\left\{\frac{N-k}{N} p^r, \frac{1}{N-k+1} J_{N-k+1}(0) + \frac{N-k}{N-k+1} J_{N-k+1}(r+1)\right\}$$

Since we are given $r^*_{N-k+1} \geq 0$, then for all r then we can use our inductive hypothesis to show that

$$J_{N-k}(r) = \max\left\{\frac{N-k}{N} p^r, \frac{1}{N} + \frac{N-k}{N-k+1} J_{N-k+1}(r+1)\right\} \tag{11}$$

We now must break this into cases on r as we did in earlier sections. If $r \leq r^*_{N-k+1} - 1$, then (11) becomes

$$\begin{aligned} J_{N-k}(r) &= \max\left\{\frac{N-k}{N} p^r, \frac{1}{N} + \frac{N-k}{N-k+1} J_{N-k+1}(r+1)\right\} \\ &= \max\left\{\frac{N-k}{N} p^r, \frac{1}{N} + \frac{N-k}{N-k+1}\left(\frac{N-k+1}{N} p^{r+1}\right)\right\} \\ &= \max\left\{\frac{N-k}{N} p^r, \frac{1}{N} + \frac{N-k}{N} p^{r+1}\right\} \end{aligned}$$

If we now solve for the maximum r for which the optimal action is to send an offer, we find

$$r \leq \log_p\left(\frac{1}{1-p} \cdot \frac{1}{N-k}\right) = r_1.$$

Now, if we consider the case that $r > r^*_{N-k+1} - 1$, equation (11) becomes

$$\begin{aligned} J_{N-k}(r) &= \max\left\{\frac{N-k}{N} p^r, \frac{1}{N} + \frac{N-k}{N-k+1} J_{N-k+1}(r+1)\right\} \\ &= \max\left\{\frac{N-k}{N} p^r, \frac{1}{N} + \frac{N-k}{N-k+1}\left(\frac{N-k+1}{N}\left(p^{(r+k-1)+1} + \sum_{i=1}^{k-1} \frac{1}{N-i}\right)\right)\right\} \\ &= \max\left\{\frac{N-k}{N} p^r, \frac{1}{N} + \frac{N-k}{N}\left(p^{r+k} + \sum_{i=1}^{k-1} \frac{1}{N-i}\right)\right\} \\ &= \max\left\{\frac{N-k}{N} p^r, \frac{N-k}{N(N-k)} + \frac{N-k}{N}\left(p^{r+k} + \sum_{i=1}^{k-1} \frac{1}{N-i}\right)\right\} \\ &= \max\left\{\frac{N-k}{N} p^r, \frac{N-k}{N}\left(p^{r+k} + \sum_{i=1}^{k} \frac{1}{N-i}\right)\right\} \end{aligned} \tag{12}$$

Once more, if we now solve for what is the maximum r for which the optimal action is to send an offer, we find

$$r \leq \log_p\left(\frac{1}{1-p^k} \cdot \sum_{i=1}^{k} \frac{1}{N-i}\right) = r_2. \tag{13}$$

Now, by lemma 2, we have that $r_2 < r_1$, which implies that the threshold for sending an offer r^*_{N-k} exists and is given by $r^*_{N-k} = \log_p\left(\frac{1}{1-p^k} \cdot \sum_{i=1}^k \frac{1}{N-i}\right)$ as desired. It then follows from the definition that for $0 \leq r \leq r^*_{N-k}$ we have $J_{N-k}(r) = \frac{N-k}{N} p^r$.

It remains to show the functional form of $J_{N-k}(r)$ for $r > r^*_{N-k}$. If we look back at Bellman's equation (11), we need to know what to plug in for $J_{N-k+1}(r+1)$. Recall that when deriving r^*_{N-k}, we assumed $r > r^*_{N-k+1} - 1$. But by lemma 3 we have that $r^*_{N-k} < r^*_{N-k+1}$. It then follows that for $r^*_{N-k} \in [r^*_{N-k+1} - 1, r^*_{N-k+1}]$ and therefore if $r > r^*_{N-k}$ then $r > r^*_{N-k+1} - 1$. Using this we now can use our inductive hypothesis for the form of $J_{N-k+}(r+1)$ and carry out the calculations we used to get equation (12). We then have the completed functional form for $J_{N-k}(r)$,

$$J_{N-k}(r) = \begin{cases} \frac{N-k}{N} p^r & r < r^*_{N-k} \\ \frac{N-k}{N}\left(p^{r+k} + \sum_{i=1}^k \frac{1}{N-i}\right) & \text{else.} \end{cases}$$

as desired. □

To summarize what we have just done: we derived that the optimal policy has the form of a threshold for each number of candidates interviewed for when to send an offer to the current best candidate, we found an expression the policy thresholds r^*_{N-k}, we showed this is a decreasing function in k, and we derived an expression for the value function $J_{N-k}(r)$. We will now turn to proving some asymptotic results before showing numerical results and concluding remarks.

4 Asymptotic Results

In the previous section we proved that these policy thresholds r^*_{N-k} are strictly decreasing in k. The value of r^*_{N-k} dictates how many candidates back in time we should consider sending an offer to if they are the best so far with 0 being the most recent. A natural question to ask is then, when is the first time $r^*_{N-k} < 0$ implying for larger k, we never send an offer. This can be thought of as the "sampling" phase as in the classic secretary problem. For the remainder of this section we will think of N as being arbitrarily large.

Jumping right into it, we would like to solve the inequality

$$r^*_{N-k} = \log_p\left(\frac{1}{1-p^k} \cdot \sum_{i=1}^k \frac{1}{N-i}\right) < 0 \tag{14}$$

Like much of the work in this paper, we will split it into cases, now based on the size of p.

4.1 Small p

If p is not too close to 1, we can use some approximations. If we exponentiate both sides of (14), we obtain the inequality

$$\frac{1}{1-p^k} \cdot \sum_{i=1}^k \frac{1}{N-i} > 1. \tag{15}$$

Now if p is not too large, then p^k approaches zero very quickly for modestly sized k. Thus we will approximate $1/(1-p^k) \approx 1$. We can also utilize an approximation to the harmonic series $\sum_{i=1}^N \frac{1}{i} \approx \log(N)$. With these two approximations, we can write the inequality as

$$1 < \frac{1}{1-p^k} \cdot \sum_{i=1}^k \frac{1}{N-i} \approx 1 \cdot \log\left(\frac{N}{N-k}\right)$$

which implies the inequality $e < N/(N-k)$ which has solution

$$\boxed{k > N\left(1 - \frac{1}{e}\right)}.$$

But recall that this is k as measured from N, the number of candidates seen before any offer could be made is $N-k=N/e$ which is the same threshold as in the classical secretary problem!

4.2 Big p

If p^k does not decay too quickly, meaning if p is quite close to 1, then we need to be a bit more careful in our analysis. Starting with equation (15), and using the harmonic series approximation as before, we have

$$\frac{1}{1-p^k}\cdot\log\left(\frac{N}{N-k}\right)>1$$

This can be rearranged to get a similar inequality to that in the previous case

$$\frac{N-k}{e^{p^k}}>\frac{N}{e}. \tag{16}$$

Here we can see that additional influence of the p^k term. If this were approximately zero, we would recover the small p results. Unfortunately we couldn't come up with a clean approximation for k here. But for p large enough, we would expect the threshold k to be smaller than in the small p case. In other words, for p sufficiently close to 1, we would expect to see the decision maker have a longer sampling phase than the classic N/e length sampling phase. If p^N is sufficiently close to 1, then we see no sampling phase and the decision maker will wait until seeing every candidate.

We can actually determine what p needs to be for the optimal policy to have a sampling phase that is an α fraction of all candidates (in the classical problem $\alpha=1/e$). Let $1/e<\alpha<1$. If we plug in $k=(1-\alpha)N$ for k in equation (16), and then solve for p, we find

$$p>\exp\left[\frac{\log(1+\log\alpha)}{(1-\alpha)N}\right]. \tag{17}$$

For an example, if $N=100$ and we want the optimal policy to sample $\alpha=1/2$ fraction of the candidates before sending an offer, we will need p to be $p=\exp\left[\frac{\log(1-\log 2)}{50}\right]\approx 0.976649$. If we look at Figure 6 can see this this agrees with numerical results.

One interesting further bit of analysis that we could do but will leave as a future topic is, given N, to determine the minimum p for which the optimal sampling phase begins to deviate from the classic $1/e$ sampling fraction.

We will not go into an asymptotic analysis of the value function here as it is not easily approximated[1]. We leave this as a future topic for analysis. We will see in the next section that for most $p\in(0,1)$, the modified problem only marginally outperforms the strategy of the classic secretary problem. However, for p sufficiently large we see a dramatic improvement in the probability of obtaining the best candidate when compared to the classic problem (see Figure 10).

5 Numerical Results

We numerically solved for the value functions and optimal policy in each state for a variety of candidate pool size N and $p\in(0,1)$. Some cases highlighting the theoretical results can be seen below in Figures 7, 8, and 9.

Based on these figures, we can clearly see that there is a threshold r^*_{N-k} that is decreasing in k, as the theory predicts. Furthermore, we see the asymptotic sampling phase threshold of N/e for the smaller p values and then for p close to 1, we see it begin to deviate, as predicted. We also see that the value function is increasing in each state for p, and decreasing in r, which equation (9) from the theory predicts.

[1] I didn't have enough time to do it justice

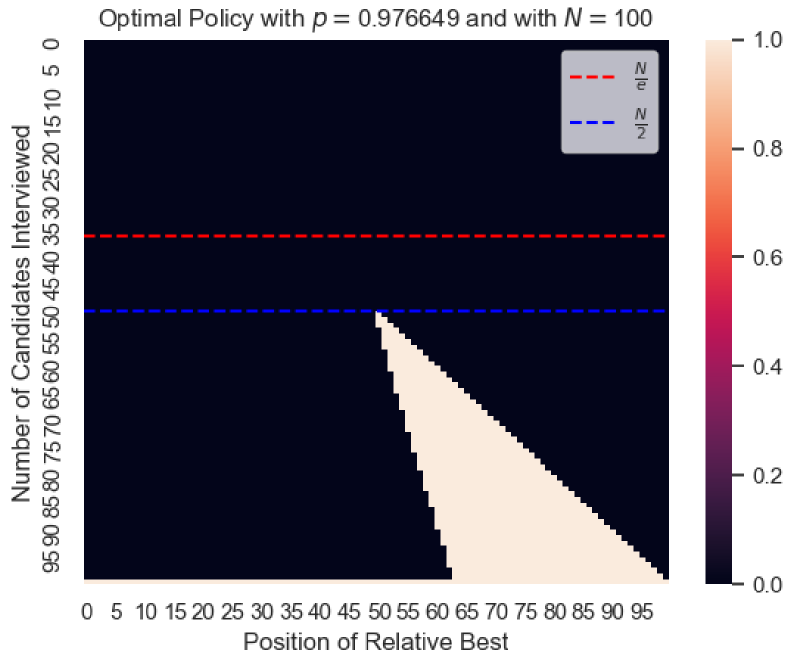

Figure 6: The optimal policy for $p = 0.976649$ and $N = 100$.

Lastly, we will look at how the probability of ending up with best candidate $J_1(0)$ compared to the classical secretary problem depends on p for large N. We plotted the numerical results for $J_1(0)$ for $N = 50$ and $N = 100$ below in Figure 10. We also included the numerical results from a closely related problem: secretary problem with recall and uncertain employment. It is the same as our variant but we now can send an offer each time period and the game ends when we an offer is accepted or we have exhausted the candidates with no accepted offer.

We can see that in our case, unless p is sufficiently large, we only marginally outperform the optimal policy in the classical problem. However, for large enough p, we see we can substantially outperform the classical result. This is because we can still keep a likely best candidate in consideration longer and still explore the space to see if something better arises.

6 Conclusion and Future Work

In this paper we looked at the one-shot secretary problem with recall and uncertain employment. We assumed the special form of probability function $q(r) = p^r$. We then looked at several cases depending on the value of p. For $p = 0$ and $p = 1$ we found the problem either reduced to a known problem (the classic secretary problem) or a trivial problem (wait until we see all the candidates and pick the best). In the case of $p \in (0, 1)$ we were able to show that the optimal policy has the form of: for any $N - k$ number of candidates seen, if the best candidate was one of the r^*_{N-k} most recent candidates, send them an offer, otherwise wait. We were then able to derive and prove an explicit form for r^*_{N-k} based on some modest assumptions. We then derived and proved a closed form for $J_{N-k}(r)$, given that we have not yet reached this policy threshold. We then moved on to analyzing each of these components asymptotically and found an approximation for the value of k for which we should first consider sending an offer to the current best candidate, marking the end of the "sampling phase". For large p this approximation is not as easily approached but we were able to determine: in order to have a α fraction sampling phase, how big must p be? We then demonstrated empirical evidence for our theoretical work by solving the problem exactly in Python using dynamic programming. The data nicely supports our claims.

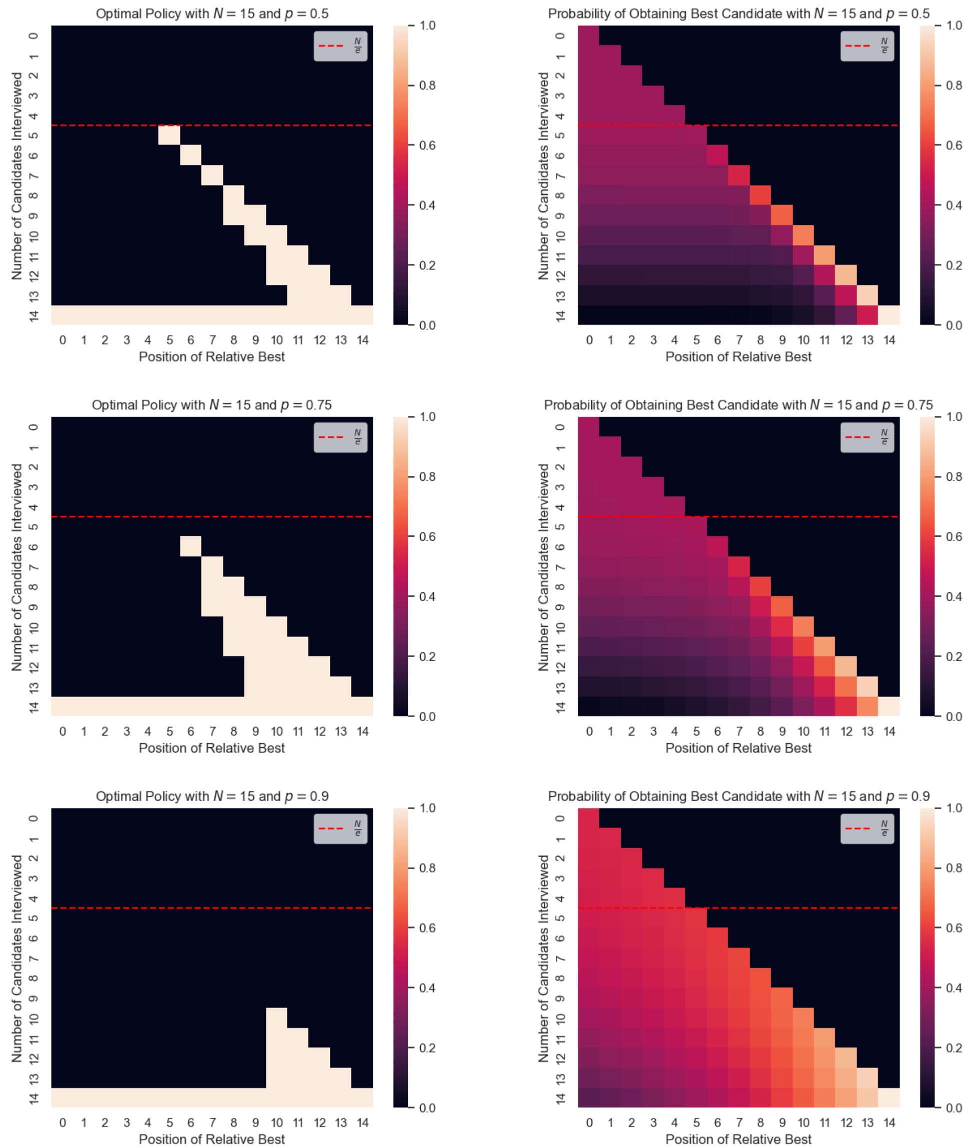

Figure 7: Numerically solving for the optimal policy and the value function for each state when $N = 15$ for several p values.

In the future, it would be interesting to find an approximation for $J_1(0)$ based on equation (9). Also, if we look at Figure 10, in the left panel we can see there appears to be some constant offset for the probability of success in our problem (orange line) and the classical probability (red line) for moderately sized p. It would be interesting to investigate if this is a general trend and if so can we approximate the offset. We would also like to explore different forms for $q(r)$. While our chosen form seemed quite natural, there may be alternative forms that give rise to interesting behavior.

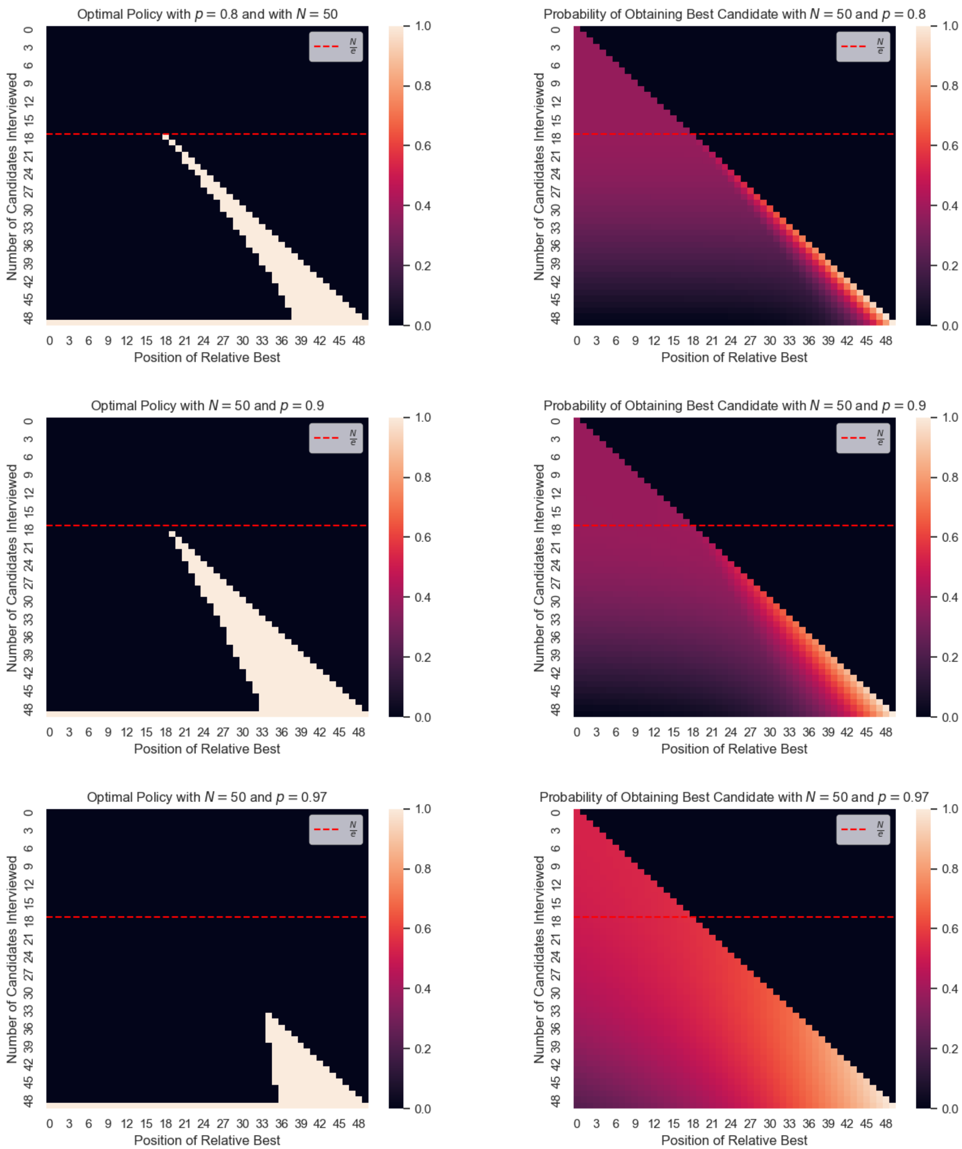

Figure 8: Numerically solving for the optimal policy and the value function for each state when $N = 50$ for several p values.

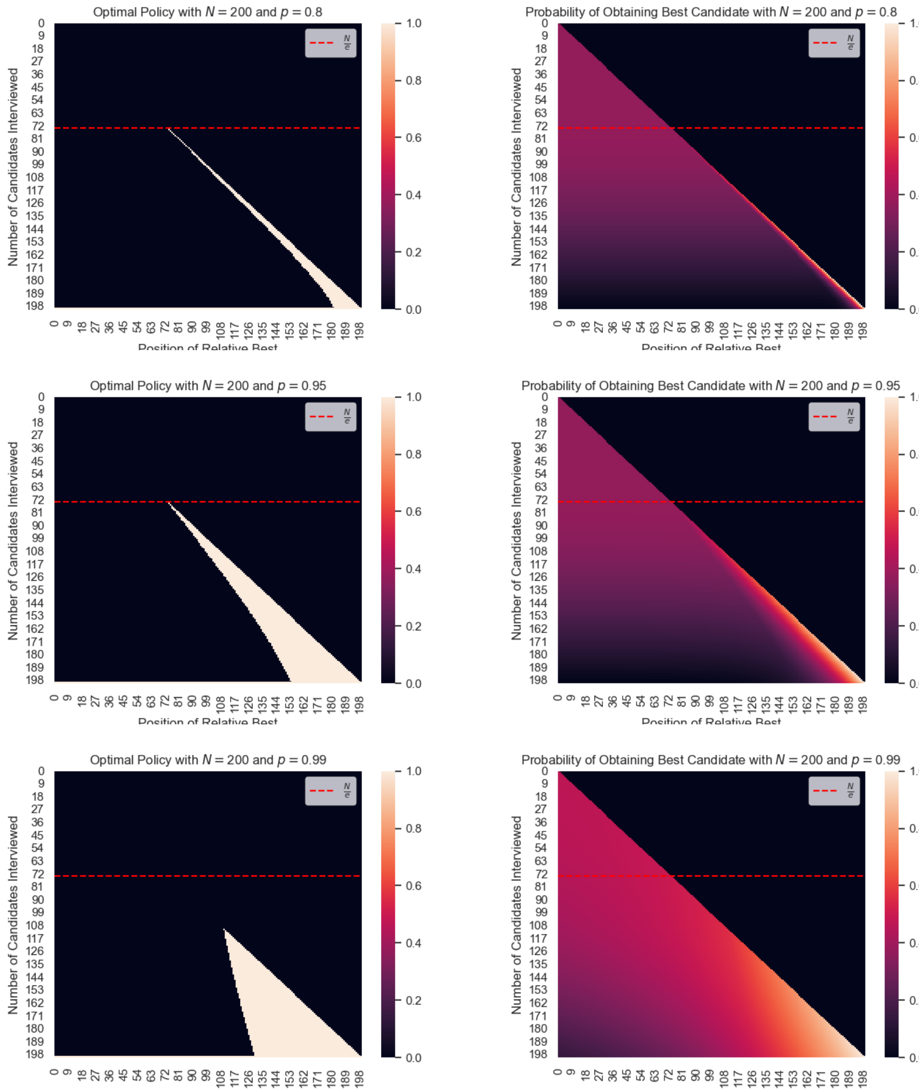

Figure 9: Numerically solving for the optimal policy and the value function for each state when $N = 200$ for several p values.

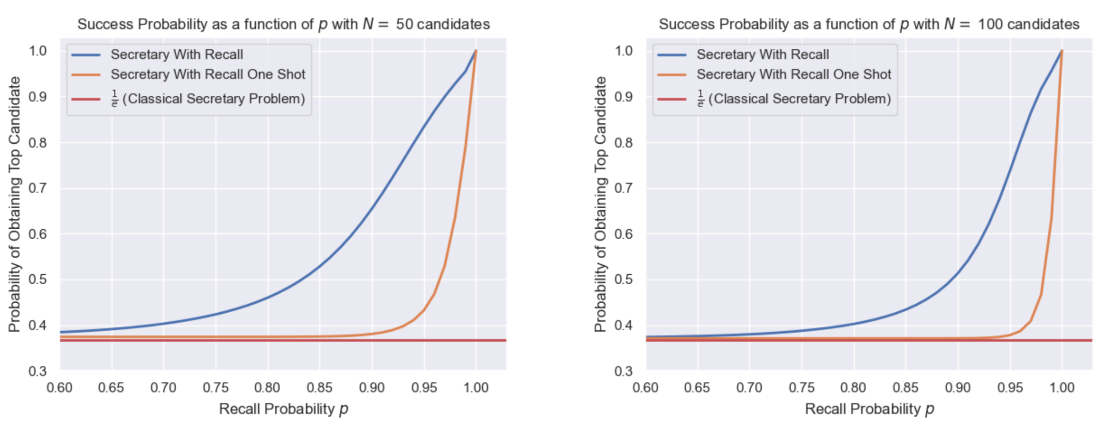

Figure 10: The probability of successfully obtaining the best candidate in each variant. The variant we studied in this paper can be seen in orange.