Revisiting the Secretary Problem

Steven J. Brams
Department of Politics
New York University
New York, NY 10012
USA
steven.brams@nyu.edu

D. Marc Kilgour
Department of Mathematics
Wilfrid Laurier University
Waterloo, Ontario N2L 3C5
CANADA
mkilgour@wlu.ca

Abstract

The venerable Secretary Problem asks how a decision maker (DM) should select one of *n* candidates, who come up randomly, when the only information available is each candidate's strict rank within the set of previous applicants. Moreover, DM may select only the current candidate. Our illustrative case has $n = 9$ candidates, for which the *Standard Method* is to reject outright the first 3 candidates and then choose the first of the 4^{th} through 8^{th} candidates who is better than any of the first three, and the 9^{th} candidate if none of the 4^{th} through 8^{th} candidates satisfies this condition. We compare the Standard Method with two other selection methods that change the conditions under which DM decides:

Reserve Method. Same as the Standard Method, except that it assumes the best of the first 3 candidates is held in reserve and chosen if none of the 4^{th} through 8^{th} (Version A) or the 4^{th} through 9^{th} (Version B) candidates is better.

Score Method. Each candidate receives a score between 0 and 1; scores are known to be uniformly distributed. DM decides, on each round, a numerical threshold and selects a candidate that exceeds it. If none does, the 9^{th} candidate is selected. We assume that the DM uses thresholds that maximize the expected score of the selected candidate.

The Standard Method gives DM a probability of 41% of selecting the best candidate, whereas the Reserve Methods substantially raise this probability to 70 – 74%—depending on whether Version A or B is used—while the Score Method raises it to 55%. Thus, the other methods, especially the Reserve Methods, outperform the Standard Method in selecting the best candidate, but the Score Method requires seeing significantly fewer candidates—an average of 4.3, compared with an average of 6.3 for the Standard Method and 6.3 and 6.7 for the two versions of the Reserve Method.

We also discuss a third criterion, the average rank of the candidate selected, which is about 1.5 for the Reserve Methods and the Score Method but 2.9 for the Standard Method. In sum, the superiority of the alternative methods over the Standard Method reflects the severity of the restrictions placed on the original Secretary Problem, suggesting it is time to revisit the assumptions of this method and consider realistic alternatives.

1. Introduction

The Secretary Problem became famous when Martin Gardner (1960) discussed it in his February 1960 *Scientific American* column. A decision maker (DM) must select one candidate for a vacant position under the following conditions:

- There are n candidates, where n is known.
- DM can strictly rank any subset of the candidates from best to worst but has no additional information about candidate quality.
- The candidates are interviewed sequentially and in random order, where every one of the $n!$ orders is equiprobable.
- Immediately after being interviewed, a candidate must be selected or rejected and, if rejected, cannot be selected later.

DM's objective is to maximize the probability of selecting the best candidate.

The Secretary Problem has other names, such as the marriage problem, because it can be viewed as choosing a spouse under the condition that a rejected candidate cannot be reconsidered later. For an overview of the problem, its solution, and extensions, see Wikipedia (2023), Freeman (1983), and Ferguson (1989). It continues to appear in different formulations in the literature: Bruss (1984) posits n to be a random variable and so not known, as assumed here; Hahn et al. (2022) use it as a context for principal-agent information transmission; and Li and Toda (2022) incorporate certain costs in their model. Each of the latter two papers discusses some of the considerable recent literature, which has not stopped growing since Gardner first popularized the problem and its striking solution.

The procedure that has been demonstrated to solve the problem, which we call the *Standard Method*, is to reject outright the first $r < n$ candidates and then choose the first subsequent candidate who is better than any of the first r. If this condition is not met after candidate $n - 1$ is interviewed, the n^{th} candidate must be selected.

In addition to studying the Standard Method in general, we consider in detail the case of $n = 9$ candidates, where the Standard Method prescribes rejecting $r = 3$ candidates out of hand and then selecting the first candidate preferred to any of those 3. As we show, the best candidate is selected with probability 0.406. As n approaches infinity, the proportion of candidates that DM rejects initially tends toward $1/e$, or about 0.37, which is also, surprisingly, the limiting probability that DM accepts the best candidate.

We compare the Standard Method with two alternatives that significantly alter features of the problem. In preparation, we ask two further questions about the Standard Method. First, how many candidates are likely to be interviewed? Second, what is the probability of ending up with a good—but not necessarily the best—candidate? We address the first question with an expected-value calculation and the second with an ordinal comparison of candidates.

Initially, it would appear unwise, as prescribed by the Standard Method, for DM to reject outright candidates, one of whom may turn out to be the best. As our first alternative method, we propose the *Reserve Method*, in which the best of the first 3 rejected candidates is kept in reserve and turned to later if and only if none of candidates 4 – 8 (or 4 – 9)—there are two variants—is better. Note that the Reserve Method requires that the rules be revised so that the best of the first 3 candidates, automatically rejected under the Standard Method, is available for selection if no later candidate is better.

The first variant of the *Reserve Method* is that the best of the first 3 candidates is selected if none of candidates 4 – 8 is better—instead of defaulting to the selection of the 9th candidate, as under the Standard Method. But if a better candidate does turns up among the 4th through 8th candidates, that candidate will be selected, exactly as under the Standard Method.

The second variant of the *Reserve Method* is identical to the first, except that the reserve candidate is selected only if none of candidates 4 – 9 is better. In effect, there is a runoff between the reserve candidate and the 9th candidate, with the better one selected instead of automatically selecting the 9th candidate, as under the first variant.[1]

Our second alternative method, the *Score Method*, "cardinalizes" the problem. Each candidate is tested; the test scores are assumed uniformly distributed on a 0-1 scale of increasing quality.[2]

We compare three methods—Standard, Reserve (variants A and B), and Score according to three criteria:

(1) Probability of selecting the best candidate;
(2) Expected number of candidates to be interviewed; and
(3) Expected quality of the candidate selected.

The usual standard is (1), as already discussed, but (2) is important for a DM whose time to assess the candidates is valuable. Criterion (3) is ordinal for the Standard and Reserve Methods but cardinal (i.e., quantitative) for the Score Method. Obviously, (1) and (3) are related.

---

[1] Smith and Deely (1975) calculate how many candidates DM must evaluate, and whose rankings he must remember, before reaching a prespecified probability of finding the best among *all* candidates. This rule differs from ours, wherein DM stops when a candidate is better than any candidate rejected at the start; if there is no such candidate, the better of the candidate who remains and the best of the rejected candidates is selected.

[2] This approach was taken by Sakaguchi (1961); see also Wikipedia (2023), where it is called the dynamic programming method. The first candidate whose score exceeds a threshold is selected, where the threshold may depend on the round. We find thresholds, which do indeed depend on the round, that are optimal in the following sense: They maximize the expected score of the selected candidate. These thresholds decline as the number of remaining candidates decreases, which is to say that DM becomes less fussy as the pool of candidates shrinks.

All three methods, because of their sequential nature, are vulnerable to the premature choice of a candidate. This happens when a candidate, who is good enough to be selected, precedes a candidate who is the best. This is especially problematic for the Standard Method, which rejects 1/3 of the candidates out of hand.

After discussing the setup of the Standard Method in section 2, we analyze and compare the three methods in section 3. The Standard Method finds the best candidate with a probability of 41%; this probability is 70% and 74% for Reserve Methods A and B and 55% for the Score Method.

On the other hand, the Score Method requires significantly fewer candidates be assessed than the other two methods (an average of 4.3, compared with 6.3 for the Standard Method and 6.3 and 6.7 for the two versions of the Reserve Method). The Reserve and Score Methods are comparable in expected quality, however; they select a candidate with an average rank of about 1.5 versus 2.9 for the Standard Method. In the concluding section, we weigh the advantages and disadvantages of each method.

2. The Standard Method

First, we show how to set the parameter r of the Standard Method, in which the first r of the n candidates are assessed but not selected. Instead, DM selects the first of candidates $r + 1, r + 2, \ldots, n - 1$ who is preferred to all the first r candidates, or candidate n if none of these candidates is preferred to all the first r candidates.[3]

We choose r to maximize the probability that the best candidate is selected. To calculate this probability, note that any candidate j is best with probability $1/n$. Fix r, and let q_j be the conditional probability that candidate j is selected given that candidate j is best overall. Clearly, $q_1 = \ldots = q_r = 0$. Moreover, $q_{r+1} = 1$, because if candidate $r + 1$ is best overall, then she is certainly better than any of the first r candidates. Given that candidate $r + 2$ is best, she is selected whenever any of the first r candidates is preferred to candidate $r + 1$, which occurs with probability $r/(r + 1)$. For $j > r + 1$, given that candidate j is the best candidate, she is selected when the best preceding candidate is among the first r, which occurs with probability $r/(j - 1)$.

Thus, when there are n candidates, the (unconditional) probability that the best candidate is selected is

$$Q(r) = \frac{1}{n}\left[0 + 0 + \cdots + 0 + 1 + \frac{r}{r+1} + \frac{r}{r+2} + \cdots + \frac{r}{n-1}\right]$$

for any $r = 1, 2, \ldots n - 1$. This expression can be written

[3] For a different perspective on this method when n is not known, see Bruss (1984). But n will be known if the number of candidates is large, and DM (assumed male) must limit himself to seeing no more than n candidates (assumed female). This gender distinction is for convenience only; we could as well make DM female and the candidates male.

$$Q(r) = \frac{r}{n}\sum_{i=r}^{n-1}\frac{1}{i} = \frac{r}{n}[H(n-1) - H(r-1)]$$

where $H(j)$ is the sum of the first j terms of the harmonic series, $H(j) = \sum_{i=1}^{j}\frac{1}{i}$.

For specific values of n, it is easy to evaluate $H(j)$ and thus $Q(r)$. For $n = 9$, the table below shows that the value of r that maximizes $Q(r)$ is $r = r^* = 3$.

r	1	2	3	4	5	6	7	8
$Q(r)$	0.3020	0.3817	0.4060	0.3931	0.3525	0.2897	0.2083	0.1111

Hence, the probability that the Standard Method selects the best candidate in this representative case is $Q(r^*) = 0.4060$, where $r^* = 3$.[4]

3. Comparing the Three Methods

Standard Method

We have already shown that, in our example with $n = 9$, the Standard Method should be applied with $r = 3$. In this case, the probability that the best candidate is selected is $Q = \frac{1}{9}\sum_{j=1}^{9} q_j = \frac{1}{9}\left(0 + 0 + 0 + 1 + \frac{3}{4} + \frac{3}{5} + \frac{3}{6} + \frac{3}{7} + \frac{3}{8}\right) = \frac{341}{840}$, or Q = 0.4060.

Our analysis of the Standard Method begins with the determination of p_j, the (unconditional) probability that the j^{th} candidate is selected. Because the first r candidates are rejected out of hand, it is clear that $p_j = 0$ for $j = 1, 2, \ldots, r$. Set $n = 9$ and $r = 3$, and consider the 4th round. The probability that the 4th candidate is the best so far is the probability that she is better than any of the first 3 candidates, which is ¼, so $p_4 = ¼$.

Now consider when the 5th candidate is selected. Two conditions must be satisfied. First, the 4th candidate must not have been selected. Second, the 5th candidate must be the best of the first 5. Thus, of the five candidates now under consideration, the 5th must be first-ranked, and the second-ranked must be one of candidates 1, 2, and 3. Of the 5! permutations of {1, 2, 3, 4, 5}, 3 × 3! satisfy these conditions. Thus, the probability that the 5th candidate is selected is $p_5 = 3 \times 3!/5! = 3/20$.

By similar reasoning, it follows that, for $j = 4, 5, \ldots 8$,

[4] Substituting $H(n - 1) - H(r - 1) \approx \ln n - \ln r$ into $Q(r)$ and treating r as a continuous variable permits maximization of $Q(r)$ using calculus, resulting in the well-known formula $r^* \approx n/e$, where e is the base of the natural logarithms.

$$p_j = \frac{3 \times (j-2)!}{j!} = \frac{3}{j(j-1)}$$

In particular, the probability that the 8th candidate, or some prior candidate, is chosen is

$$1/4 + 3/20 + 1/10 + 1/14 + 3/56 = 5/8.$$

Hence, the probability that the Standard Method selects the 9th candidate, which occurs if and only if no previous candidate was selected, is $p_9 = 1 - 5/8 = 3/8$. In the table below, we show the selection probabilities for candidates 4 – 9.

Candidate	4	5	6	7	8	9
Selection Probability	$p_4 = 1/4 =$ 0.250	$p_5 = 3/20$ = 0.150	$p_6 = 1/10$ = 0.100	$p_7 = 1/14$ = 0.071	$p_8 = 3/56$ = 0.054	$p_9 = 3/8$ = 0.375

Note that the "best so far" rule for choosing a candidate would select the 9th candidate with probability 1/24 = 0.042, as can be seen by substituting $j = 9$ in the formula for p_j above. The 9th candidate is selected if she either meets the "best so far" condition or if the best candidate is one of the first 3, which occurs with probability 1/3. Summing these probabilities gives $p_9 = 1/24 + 1/3 = 3/8$. For general values of n and r satisfying $0 < r \le n - 2$, it can be verified that $p_1 = \ldots = p_r = 0$, $p_j = r/(j(j-1))$ for $j = r + 1, \ldots, n - 1$, and $p_n = r/(n-1)$. If $0 < r = n - 1$, then $p_1 = \ldots = p_{n-1} = 0$ and $p_n = 1$.

Another measure of the quality of a selection method is $E[C]$, the expected number of candidates that DM evaluates before selecting one. In our example with $n = 9$ and $r = 3$, we note that

(i) three candidates are rejected outright at the beginning;
(ii) if candidate j is selected, where $j = 4, 5, 6, 7$, or 8, then DM must have evaluated exactly j candidates; and
(iii) if candidate 9 is selected, then DM must have evaluated exactly 8 candidates, because candidate 9 is selected exactly when all previous candidates are rejected.

It follows that

$$E[C] = \sum_{j=4}^{8} j \cdot p_j + 8 \cdot p_9 = 4 \cdot \frac{1}{4} + 5 \cdot \frac{3}{20} + 6 \cdot \frac{1}{10} + 7 \cdot \frac{1}{14} + 8 \cdot \frac{3}{56} + 8 \cdot \frac{3}{8} = \frac{879}{140}$$

so that $E[C] = 6.2786$. On average, DM must see about 6.3 candidates before selecting one of them – almost 70% of the 9 candidates.

For general values of n and r satisfying $0 < r \le n - 1$, it is easy to verify that

$$E[C] = r[1 + H(n-2) - H(r-1)]$$

where, as before, $H(j)$ is the j^{th} partial sum of the harmonic series, $H(j) = \sum_{i=1}^{j} \frac{1}{i}$ and $H(0) = 0$.

Our third measure of the quality of the Standard Method is the expected quality of the candidate selected. Because we have only ordinal information, we use $E[R]$ to measure expected quality, where R is a random variable representing the rank of the candidate selected, and $E[R]$ is its expected value. For instance, if the method manages to select the best overall candidate, then $R = 1$. We already know that Standard Method finds the best overall candidate less than half the time, so $R > 1$ must occur frequently.

To find the expected rank of the candidate selected by the Standard Method, we find the conditional probability, t_{jk}, that $R = k$, that is, that the selected candidate has rank k (among all candidates), given that the selected candidate is the j^{th} candidate. For instance, set $n = 9$ and $r = 3$, and consider candidate $j = 4$. If candidate 4 is selected, then she must be preferred to candidates 1, 2, and 3, but she may in fact be ranked as low as 6th among all candidates. Thus, $t_{4k} > 0$ whenever $k = 1, 2, \ldots$, or 6.

In contrast, $t_{8k} > 0$ only when $k = 1$ or 2, because candidate 8 can be selected only if she is better than all 7 preceding candidates. In contrast, if candidate $j = 9$ is selected, her rank may be 1st, 2nd, …, or 9th.

For fixed $j = 4, 5, 6, 7$, or 8, we have noted that $t_{jk} > 0$ for $k = 1, 2, \ldots, 10 - j$. There are $\frac{3}{j(j-1)} \cdot 9!$ (equally probable) rankings of the candidates in which the Standard Method selects candidate j. (Of candidates 1, 2, … j, candidate j must be top-ranked and candidate 1, 2, or 3 must be second-ranked.) If $k = 1, 2, \ldots, 10 - j$, candidate j is k^{th} in exactly $\frac{(9-j)!}{(10-j-k)!} \cdot \frac{3}{j-1} \cdot (9-k)!$ of these rankings. It follows that the conditional probability $R = k$, that is, that candidate j is ranked k^{th}, given that j is chosen, is

$$t_{jk} = \frac{j \cdot (9-j)! \cdot (9-k)!}{(10-j-k)! \cdot 9!}, j = 4, 5, 6, 7, 8; k = 1, 2, \ldots, 10 - j.$$

Now consider $j = 9$. The Standard Method selects candidate 9 if and only if the top-ranked among candidates 1, 2, …, 8 is one of candidates 1, 2, and 3. There are $\frac{3}{8} \cdot 9!$ rankings satisfying this condition. For each $k = 1, 2, \ldots, 9$, candidate 9 is ranked k^{th} in exactly $\frac{3}{8} \cdot 8!$ Of these, it follows that

$$t_{9k} = \frac{1}{9}, k = 1, 2, \ldots, 9.$$

The values of t_{jk} are given in Table 1.

j	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = 7$	$k = 8$	$k = 9$	p_j

4	$\frac{56}{126}$	$\frac{35}{126}$	$\frac{20}{126}$	$\frac{10}{126}$	$\frac{4}{126}$	$\frac{1}{126}$	--	--	--	$\frac{1}{4}$
5	$\frac{70}{126}$	$\frac{35}{126}$	$\frac{15}{126}$	$\frac{5}{126}$	$\frac{1}{126}$	--	--	--	--	$\frac{3}{20}$
6	$\frac{56}{84}$	$\frac{21}{84}$	$\frac{6}{84}$	$\frac{1}{84}$	--	--	--	--	--	$\frac{1}{10}$
7	$\frac{28}{36}$	$\frac{7}{36}$	$\frac{1}{36}$	--	--	--	--	--	--	$\frac{1}{14}$
8	$\frac{8}{9}$	$\frac{1}{9}$	--	--	--	--	--	--	--	$\frac{3}{56}$
9	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{3}{8}$

Table 1: Standard Method: Conditional probability that $R = k$, given that candidate j is selected

Of course, the unconditional probability that candidate j is selected and that she is k^{th}-ranked is $p_j \cdot t_{jk}$. These values are shown as decimals in Table 2.

j	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = 7$	$k = 8$	$k = 9$
4	.1111	.0694	.0397	.0198	.0079	.0020	--	--	--
5	.0833	.0417	.0179	.0060	.0012	--	--	--	--
6	.0667	.025	.0071	.0012	--	--	--	--	--
7	.0556	.0139	.0020	--	--	--	--	--	--
8	.0476	.0060	--	--	--	--	--	--	--
9	.0417	.0417	.0417	.0417	.0417	.0417	.0417	.0417	.0417
Pr{k}	.4060	.1976	.1083	.0687	.0508	.0437	.0417	.0417	.0417

Table 2: Standard Method: Unconditional probability that candidate j is selected and $R = k$

From Table 2, one can calculate $E[R] = \sum_{k=1}^{9} k \cdot \Pr\{k\} = 2.9167$. Thus, under the Standard Method the candidate selected is, on average, just slightly better than the 3^{rd}-ranked candidate.

For general values of n and r satisfying $0 < r \leq n - 1$, it can be seen that

$$t_{jk} = \frac{j \cdot (n-j)! \cdot (n-k)!}{(n+1-j-k)! \cdot n!}, j = r+1, \dots, n-1; k = 1, 2, \dots, n+1-j;$$

$$t_{nk} = \frac{1}{n}, k = 1, 2, \dots, n.$$

Reserve Method

The Reserve Method is almost identical to the Standard Method. The only difference is that the best of the first 3 candidates is held in reserve and will be selected if none of candidates 4 – 8 is selected.[5] In fact, there are two variants of the Reserve Method, which we earlier called Version A and Version B. We next compare them with each other as well as with the Standard Method.

Reserve Method A: If none of candidates 4 – 8 is selected, the reserve candidate (the best of candidates 1, 2, and 3) is selected without assessing candidate 9.

Reserve Method B: If none of candidates 4 – 8 is selected, then candidate 9 is assessed, and the preferred of candidate 9 and the reserve candidate is selected.

Observe that Reserve Method B requires one more assessment than Reserve Method A. It selects candidate 9 if she is the best candidate overall (not just the best of candidates 1, 2, and 3), rather than being the default candidate if none of candidates 4 – 8 is better than candidates 1, 2, or 3. Below we denote the reserve candidate by R.

We now proceed to calculate these probabilities, Q^A and Q^B, analogous to $Q(r)$ for the Standard Method, for the two Reserve Methods. First, for each candidate who may be selected, we calculate the conditional probability that she is selected, given that she is the best candidate overall. These probabilities are tabulated below.

j	4	5	6	7	8	9	R
q_j^A	1	$\frac{3}{4}$	$\frac{3}{5}$	$\frac{3}{6}$	$\frac{3}{7}$	0	1
q_j^B	1	$\frac{3}{4}$	$\frac{3}{5}$	$\frac{3}{6}$	$\frac{3}{7}$	$\frac{3}{8}$	1

For the Standard Method, each candidate is best overall with probability 1/9, so candidate R is best overall with probability 1/3. Therefore, the probability that the best candidate is selected under Reserve Method A is

$$Q^A = \frac{1}{3} q_R{}^A + \frac{1}{9} \sum_{j=4}^{9} q_j{}^A = \frac{1}{3} \cdot 1 + \frac{1}{9}\left(1 + \frac{3}{4} + \frac{3}{5} + \frac{3}{6} + \frac{3}{7} + 0\right) = \frac{2560}{3780}, \text{ or } Q^A = 0.6976.$$

Similarly, the probability that the best candidate is selected using Reserve Method B is

$$Q^B = \frac{1}{3} q_R{}^B + \frac{1}{9} \sum_{j=4}^{9} q_j{}^B = \frac{1}{3} \cdot 1 + \frac{1}{9}\left(1 + \frac{3}{4} + \frac{3}{5} + \frac{3}{6} + \frac{3}{7} + \frac{3}{8}\right) = \frac{207}{280}, \text{ or } Q^B = 0.7393.$$

For general values of n and r satisfying $0 < r \leq n - 2$, it can be verified that

$$Q^A = \frac{1}{r} + \frac{r}{n}[H(n-2) - H(r-1)]$$

[5] This seems to be the method that the German astronomer, Johannes Kepler, used after his first wife died of cholera in 1611. Over the next two years he interviewed 11 candidates for a new wife and eventually chose the 5th, apparently having kept her in reserve so he could return to her if none of the later candidates proved better or was no longer available (Ferguson, 1989).

$$Q^B = \frac{1}{r} + \frac{r}{n}[H(n-1) - H(r-1)]$$

Next, we note that the probability of selecting the j^{th} candidate under Reserve A, p_j^{A}, and the parallel probability under Reserve B, p_j^{B}, are quite different from their probabilities in the previous table. For example, compare $q_9^{\text{B}} = 3/8$ and $p_9^{\text{B}} = 1/24$: The former is the probability that none of candidates 4 – 8 is better than candidates 1, 2, and 3 (i.e., candidate 9 is the best so far, compared with candidates 4 – 8), whereas the latter is the probability that candidate 9 is better than all the earlier candidates (the best overall, including candidates 1, 2, and 3).

j	4	5	6	7	8	9	R
p_j^{A}	$\frac{1}{4}$	$\frac{3}{20}$	$\frac{1}{10}$	$\frac{1}{14}$	$\frac{3}{56}$	0	$\frac{3}{8}$
p_j^{B}	$\frac{1}{4}$	$\frac{3}{20}$	$\frac{1}{10}$	$\frac{1}{14}$	$\frac{3}{56}$	$\frac{1}{24}$	$\frac{1}{3}$

For general values of n and r satisfying $0 < r \le n - 2$, it can be verified that

$$p_1^{\text{A}} = \ldots = p_r^{\text{A}} = 0,\ p_j^{\text{A}} = r/(j(j-1)) \text{ for } j = r+1, \ldots, n-1, \text{ and } p_R^{\text{A}} = r/(n-1)$$

Similarly,

$$p_1^{\text{B}} = \ldots = p_r^{\text{B}} = 0,\ p_j^{\text{B}} = r/(j(j-1)) \text{ for } j = r+1, \ldots, n, \text{ and } p_R^{\text{B}} = 1/r$$

The expected number of candidates that DM must assess before selecting a candidate is our second criterion for comparing the different methods. We calculate $E[C^{\text{A}}]$ for Reserve Method A and $E[C^{\text{B}}]$ for Reserve Method B, just as we calculated $E[C]$ for the Standard Method. Note that the maximum number of assessments for Method A is 8 and for Method B is 9. For Method A we have

$$E[C^A] = \sum_{j=4}^{8} j \cdot p_j{}^A + 8 \cdot p_R{}^A = 4 \cdot \frac{1}{4} + 5 \cdot \frac{3}{20} + 6 \cdot \frac{1}{10} + 7 \cdot \frac{1}{14} + 8 \cdot \frac{3}{56} + 8 \cdot \frac{3}{8} = \frac{879}{140}$$

so that $E[C^{\text{A}}] = 6.2786 = E[C]$. On average, DM must see about 6.3 candidates before selecting one, exactly the same as the Standard Method. For Reserve Method B,

$$E[C^B] = \sum_{j=4}^{9} j p_j{}^B + 9 p_R{}^B = 4 \cdot \frac{1}{4} + 5 \cdot \frac{3}{20} + 6 \cdot \frac{1}{10} + 7 \cdot \frac{1}{14} + 8 \cdot \frac{3}{56} + 9 \cdot \frac{1}{24} + 9 \cdot \frac{1}{3} = \frac{1543}{280}$$

so that $E[C^B] = 6.6536 > E[C^A] = E[C]$. Thus, on average, DM must see almost 6.7 candidates before making a selection using Reserve Method B. Clearly, the higher probability Q for Method B comes at the cost of a higher value of $E[C]$.

For general values of n and r satisfying $0 < r \leq n - 2$, it can be verified that

$$E[C^A] = r[1 + H(n-2) - H(r-1)] = E[C]$$

$$E[C^B] = r[H(n-1) - H(r-1)] + \frac{n}{r}$$

Our third criterion for comparing the different methods is the expected rank of the candidate each method selects, denoted $E[R]$. For the Reserve Methods, we must revise Tables 1 and 2. These straightforward revisions are shown in the Appendix and are summarized by $t_{jk}^A = t_{jk}^B = t_{jk}$ for $j = r + 1, r + 2, \ldots, n - 1$; $k = 1, 2, \ldots n + 1 - j$, $t_{R1}^A = (n - 1)/n$, $t_{R2}^A = 1/n$, $t_{92}^B = 1$, and $t_{R2}^B = 1$.

We have already seen that the Reserve Methods dramatically increase the probability of selecting the best candidate, which equals $Q^A = 0.6976$ under Reserve Method A and $Q^B = 0.7396$ under Reserve Method B (compared with $Q = 0.4060$ under the Standard Method). These improvements are reflected in equally dramatic improvements in the expected rank of the selected candidate. Specifically, $E[R^A] = 1.4583$ under Reserve Method A and $E[R^B] = 1.4167$ under Reserve Method B, compared to $E[R] = 2.9167$ for the Standard Method. Thus, the Reserve Methods reduce the expected rank of the selected candidate from almost 3 to less than 1.5, and nearly double the probability of selecting the best candidate, from 0.41 to about 0.70. Method B does a little better than Method A in quality, but the price is an increase in the expected number of candidates.

Score Method

Using the Score Method, like the Standard Method, requires DM to make an up-or-down assessment of each candidate, without holding anybody in reserve. However, it relies not on relative comparisons of candidates but a cardinal evaluation of each one, compared to an absolute standard. How high should a candidate's score be to justify her selection? On average, how good will the selected candidate be?

The Score Method associates each of the $n > 1$ candidates with a numerical score, or measure of quality, which we assume to be a number between 0 and 1. The score of candidate j, denoted x_j, is the realization of a random variable, X_j, that is uniformly distributed on (0, 1). Candidates' scores are assumed to be independent. DM must decide whether or not to select a candidate immediately after her score is known. Comparisons with previous candidates, as under the Standard and Reserve Methods, are irrelevant.

The Secretary Problem with scores was addressed earlier (Sakaguchi, 1961), but we provide a complete development for comparison with the Standard and the Reserve

Methods, based on the aforementioned criteria. Because the problem is now cardinal, our "average quality" criterion can be measured cardinally.

The Score Method is the implementation of a *threshold procedure*, defined by an n-vector, $\mathbf{a} = (a_1, a_2, \ldots, a_{n-1}, a_n)$, where $0 < a_j < 1$ for $j = 1, 2, \ldots, n - 1$ and $a_n = 0$. The procedure is to select candidate 1 if $x_1 > a_1$; otherwise, reject candidate 1 and move on to candidate 2. In general, candidate j is selected if candidates $1, 2, \ldots, j - 1$ have been rejected and $x_j > a_j$. Note that forcing $a_n = 0$ ensures that candidate n is selected if candidates $1, 2, \ldots, n - 1$ are rejected, guaranteeing that the threshold procedure defined by $\mathbf{a}$ always selects one of the candidates.

Example 1: Suppose that $n = 3$ and consider the threshold vector (0.8, 0.6, 0). Candidate 1 is selected if $x_1 > 0.8$; otherwise, $x_1 \leq 0.8$, candidate 1 is rejected, and candidate 2 is considered. Candidate 2 is selected if $x_2 > 0.6$; if $x_2 \leq 0.6$, candidate 2 is rejected, and candidate 3 is automatically selected (i.e., she is the default candidate). Because the probability that a candidate's score exceeds a is $1 - a$, it follows that candidate 1 is selected with probability 0.2, candidate 2 is selected with probability $0.8 \times 0.4 = 0.32$, and candidate 3 is selected with probability $1 - 0.2 - 0.32 = 0.48$.

For a general threshold procedure, $\mathbf{a}$, denote the event that candidate j is selected by A_j. Clearly, exactly one of the events $A_1, A_2, \ldots, A_n$ must occur. For $j = 1, 2, \ldots, n$, the probability of A_j is

$$\Pr\{A_j\} = p_j^{\ S} = (1 - a_j) \prod_{i<j} a_i$$

where the product is 1 if $j = 1$. In particular, $p_1^S = 1 - a_1$, $p_2^S = a_1 (1 - a_2)$, $p_3^S = a_1 a_2 (1 - a_3)$, …, and, because $a_n = 0$, $p_n^S = a_1 a_2 \ldots a_{n-1}$. For Example 1 with $n = 3$ and $\mathbf{a} = (0.8, 0.6, 0)$, these formulas agree with our direct calculation of $p_1^S = 0.2$, $p_2^S = 0.32$, and $p_3^S = 0.48$.

A constant threshold procedure is an easy-to-understand case. For any t such that $0 < t < 1$, define $\mathbf{a}(t) = (t, t, \ldots, t, 0)$; in other words, $a_j = t$ for $j = 1, 2, \ldots, n - 1$, and $a_n = 0$. For $\mathbf{a}(t)$, $p_1^S = 1 - t$, $p_2^S = t(1 - t)$, $p_3^S = t^2(1 - t)$, …, $p_{n-1}^S = t^{n-2}(1 - t)$, and $p_n^S = t^{n-1}$.

We wish to design a threshold procedure so that the score of the candidate selected is as high as possible. Given that candidate 1 is selected, candidate 1's score, X_1, is uniformly distributed on $(a_1, 1)$, so the expected score of candidate 1, given that she is selected, is $E[X_1 \mid A_1] = ½ (1 + a_1)$. Similarly, for all $j = 1, 2, \ldots, n$, $E[X_j \mid A_j] = ½ (1 + a_j)$. It follows that the expected score of the candidate selected is

$$E[X] = \sum_{j=1}^{n} E\{X_j | A_j] \Pr\{A_j\} = \frac{1}{2} \sum_{j=1}^{n} (1 - a_j^{\ 2}) \prod_{i<j} a_i$$

where, because $a_n = 0$, the last term of the summation is

$$(1 - a_n^{\ 2}) a_1 a_2 \cdots a_{n-1} = a_1 a_2 \cdots a_{n-1}$$

For Example 1, with $n = 3$ and $\mathbf{a} = (0.8, 0.6, 0)$, it is easy to see that the expected score of candidate 1, given that candidate 1 is selected (probability 0.2), is 0.9. Similarly, the expected score of candidate 2, given that candidate 2 is selected (probability 0.32), is 0.8. Otherwise (probability 0.48), candidate 3 is selected with expected score 0.5. Thus, $E[X] = (0.9 \times 0.2) + (0.8 \times 0.32) + (0.5 \times 0.48) = 0.676$, which agrees with the formula above.

For the constant threshold procedure $\mathbf{a}(t)$, it is easy to verify that $E[X] = 1/2\ [1 + t - t^n]$. By calculus, $E[X]$ is maximized when $t = n^{-1/(n-1)}$. For instance, when $n = 9$, the maximum occurs at $t = 0.760$, where the value of $E[X]$ is 0.838. Thus, when there are $n = 9$ candidates, DM can ensure that the expected score of the selected candidate is .838 by simply selecting the first candidate whose score exceeds 0.76, and taking candidate 9 if none of the first eight candidates meets this condition.

To determine the optimal threshold vector, $\mathbf{s} = (s_1, s_2, \ldots, s_{n-2}, s_{n-1}, 0)$ — defined as the one that maximizes $E[X]$ — we begin by studying the function $f(x) = 1 - x^2 + xw$, where w is fixed. The maximum of $f(x)$ occurs at $x = w/2$ and is equal to $F(w) = 1 + w^2/4$. Now consider the dependence of $E[X]$ on the threshold vector $(a_1, a_2, \ldots, a_{n-2}, x, 0)$, where $a_1, a_2, \ldots, a_{n-2}$ are fixed and positive. It is easy to see that

$$E[X] = K_{n-1} + \frac{1}{2} a_1 a_2 \cdots a_{n-2} [(1 - x^2) + x]$$

where K_{n-1} does not depend on x. Taking $w = w_n = 1$, we see that $E[X]$ is maximized at $x = s_{n-1} = w_n/2 = ½$, and that the maximum value of $E[X]$ can be obtained by substituting $F(w_n) = F(1) = 5/4$ for the expression in square brackets.

Now assume that $a_{n-1} = s_{n-1}$ and consider the dependence of $E[X]$ on $x = a_{n-2}$. If $a_1, a_2, \ldots, a_{n-3}$ are fixed and positive, then we have that

$$E[X] = K_{n-2} + a_1 a_2 \cdots a_{n-3} [(1 - x^2) + x w_{n-1}]$$

where $w_{n-1} = F(w_n)$ and, as previously, K_{n-2} is constant. It follows that $E[X]$ is maximized by $x = s_{n-2} = w_{n-1}/2$, and that the maximum of the expression in square brackets is $w_{n-2} = F(w_{n-1}) = 1 + w_{n-1}^2/4$. Substituting, $s_{n-2} = 5/8$ and $w_{n-2} = 89/64 = 1.3906$.

This procedure can be continued, producing $s_j = w_{j+1}/2$ and $w_j = F(w_{j+1})$, which is equivalent to

$$s_j = \frac{1}{2}\left(1 + {s_{j+1}}^2\right)$$

for $j = n - 1, n - 2, \ldots, 2, 1$, starting with $s_n = 0$. For $n = 9$, the optimal threshold vector $\mathbf{s}$ is given by

s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9
0.836	0.820	0.800	0.775	0.742	0.695	0.625	0.5	0

The value of $E[X]$ achieved by **s** – the maximum expected score that can be attained by any threshold procedure – is $F(s_1) = ½ (1 + s_1^2) = 0.8498$.

Denote by $C^S(\mathbf{a})$ the number of candidates assessed by the Score Method with threshold vector **a**. If the event A_j occurs (i.e., candidate j is selected), then j candidates were assessed so $C^S = j$. The expected number of candidates assessed, $E[C^S(\mathbf{a})]$, is therefore

$$E[C^S(\boldsymbol{a})] = \sum_{j=1}^{j} j \cdot p_j{}^S = 1 \cdot (1 - a_1) + 2 \cdot a_1(1 - a_2) + \cdots + n \cdot a_1 a_2 \cdots a_{n-1}$$

This series simplifies to

$$E[C^S(\boldsymbol{a})] = 1 + a_1 + a_1 a_2 + \cdots + a_1 a_2 \cdots a_{n-1}$$

In Example 1, it is easy to see that $E[C^S(\mathbf{a})] = 2.28$.

For the constant threshold procedure $\mathbf{a}(t)$, it is easy to verify that $E[C^S] = 1 + t + t^2 + \ldots + t^{n-1} = (1 - t^n)/(1 - t)$. As is to be expected, smaller values of t correspond to smaller values of C. For the quality-maximizing constant threshold $t = 0.760$, $E[C^S(\mathbf{a}(t))] = 3.812$. The optimal threshold vector **s** produces $E[C^S(\mathbf{s})] = 4.2384$.

For any $j = 1, 2, \ldots, n$, let B_j be the event that candidate j's score is higher than the score of any other candidate. Recall that A_j is the event that candidate j is selected. Our third measure, $\Pr\{Q^S\}$, satisfies

$$\Pr\{Q^S\} = \Pr\{\cup_{j=1}^{n}(A_j \cap B_j)\} = \sum_{j=1}^{n} \Pr\{B_j|A_j\}\Pr\{A_j\} \quad (1)$$

because the events $A_1 \cap B_1$, $A_2 \cap B_2$, etc. are all disjoint.

We illustrate $\Pr\{Q^S\}$ using examples with $n = 3$ and threshold vector $(a_1, a_2, 0)$. It is easy to see that, conditional on A_1, candidate 1's score, X_1, is uniformly distributed on $(a_1, 1)$. Because all candidates' scores are independent, it follows that candidate 1's score, $X_1 = x_1$, is greater than the score of any other candidate with probability x_1^2. It also follows that

$$\Pr\{B_1|A_1\} = \frac{1}{1 - a_1}\int_{a_1}^{1} x_1^2 \, dx_1 = \frac{1}{1 - a_1}\frac{1 - a_1^3}{3}$$

For example, if $a_1 = 0.8$, then the conditional probability that candidate 1 is best, given that candidate 1 is chosen, is 0.8133. Of course, candidate 1 is chosen with probability

0.2, so the (unconditional) probability that candidate 1 is selected and is best is 0.2 × 0.8133 = 0.1627.

Now suppose that event A_2 occurs. Then the conditional distribution of candidate 2's score is uniform on (a_2, 1). Moreover, because candidate 1 was not chosen, the conditional distribution of candidate 1's score must be uniform on (0, a_1). Assuming that $a_1 \geq a_2$, it follows that

$$\Pr\{B_2|A_2\} = \frac{1}{1-a_2}\left[\int_{a_2}^{a_1} \frac{x_2^2}{a_1} dx_2 + \int_{a_1}^{1} x_2\, dx_2\right]$$

because, if $X_2 = x_2$ where $a_2 < x_2 < a_1$, the conditional probability that $X_1 < X_2$ is x_2/a_1. Hence, the probability that $X_2 = x_2$ is the highest score is x_2^2/a_1. It follows that

$$\Pr\{B_2|A_2\} = \frac{1}{1-a_2}\left[\frac{a_1^3 - a_2^3}{3a_1} + \frac{1-a_1^2}{2}\right]$$

(Note that, if $a_2 = a_1$, as in a constant-threshold procedure, the first term on the right side disappears.) Using Example 1, where $a_1 = 0.8$, and $a_2 = 0.6$, the conditional probability that candidate 2 is best, given that candidate 2 is selected, is 0.7583. Therefore, the probability that candidate 2 is selected and is best overall is 0.32 × 0.7583 = 0.2427.

To find the value of $Pr\{Q^S\}$ associated with the Example 1 threshold vector (0.8, 0.6, 0), all that remains is to calculate the probability that candidate 3 is best given that candidate 3 is selected, which occurs with probability 0.48. By the same reasoning as earlier, that conditional probability is

$$\Pr\{B_3|A_3\} = \frac{1}{a_1 a_2}\int_0^{a_2} {x_3}^2 dx_3 + \frac{1}{a_1}\int_{a_2}^{a_1} x_3 dx_3 + \int_{a_1}^{1} 1\, dx_3$$

because, as usual, if A_3 occurs, then it must be the case that $X_1 < a_1$ and $X_2 < a_2$. It follows that

$$\Pr\{B_3|A_3\} = \frac{{a_2}^3}{3a_1a_2} + \frac{{a_1}^2 - {a_2}^2}{2a_1} + 1 - a_1$$

For the threshold vector (a_1, a_2, 0) = (0.8, 0.6, 0), we have $Pr\{B_3|A_3\} = 0.525$, so the unconditional probability that candidate 3 is selected and is best is 0.48 × 0.525 = 0.252. It follows that, for Example 1,

$$Q^S = (0.2 \times 0.8533) + (0.32 \times 0.7583) + (0.48 \times 0.525) = 0.6573$$

In summary, the threshold procedure based on (0.8, 0.6, 0) selects the best candidate with probability almost 2/3.

To evaluate Q^S for a general threshold vector with n candidates, a similar set of integrals must be evaluated. Recall that $a_n = 0$ and set $a_0 = 1$. Define $b_1 = 1$ and $b_j = b_{j-1}a_{j-1}$ for $j = 2, 3, \ldots n$. (Thus, for $j > 1$, $b_j = a_1 a_2 \ldots a_{j-1}$.) It can be shown that

$$\Pr\{B_j|A_j\} = \frac{1}{1-a_j}\sum_{k=1}^{j}\left[\frac{1}{b_{j-k+1}}\int_{a_{j-k+1}}^{a_{j-k}} x^{n-k}dx\right] = \frac{1}{1-a_j}\sum_{k=1}^{j}\frac{{a_{j-k}}^{n-k+1} - {a_{j-k+1}}^{n-k+1}}{(n-k+1)b_{j-k+1}}$$

As in (1), the sum over j of the values of $\Pr\{B_j \mid A_j\}p_j^S$ equals Q^S. In particular, for $n = 9$ and the optimal threshold vector **s**, it can be verified that $Q^S = 0.5474$.

We summarize our comparison of our three methods in Table 3. We compare the methods for the selection of one of $n = 9$ candidates on three criteria:

- Probability that the best candidate is selected, Q;
- Average number of candidates assessed, $E[C]$; and
- Expected rank, $E[X]$, or expected score, E(X), of the selected candidate.

Criterion	Standard Method	Reserve Method A	Reserve Method B	Score Method (Optimal Thresholds)
Q	0.4060	0.6976	0.7393	0.5474
$E[C]$	6.2786	6.2786	6.6536	4.2384
E(R) or E(X)	$E[R] = 2.9167$	$E[R] = 1.4583$	$E[R] = 1.4167$	$E[X] = 0.8498$

Table 3: Comparison of the selection methods on three criteria

To compare the overall quality of the Score Method with the others, note that the expected score of the best of 9 candidates is 0.9, and the expected score of the second-best is 0.8. Thus, the expected score of a candidate selected by the Score Method is about midway between the expected positions of the top candidate and the second-best candidate, which in fact is quite close to the expected rank of about 1.5 for a candidate selected by one of the Reserve Methods.

4. Conclusions

Finding the best candidate would be easy if the choice could be postponed until all candidates have been assessed. Making the reserve candidate the best candidate seen so far, and selecting the reserve after all candidates are assessed, guarantees that DM selects the best candidate. But this scheme involves luxuries that DM may not be able to afford, both in the time needed to see all the candidates and the effort needed to assess their qualities.

What makes the Secretary Problem intriguing, and its standard solution far from obvious, are two restrictions that are built into the Standard Method: (1) once a candidate is selected, there is no further consideration of other candidates; and (2) the successful candidate's selection is based only on ordinal information—whether she is better or worse than the other candidates who have been assessed so far. We compared the Standard Method with other methods that relax either condition (1) or (2) in terms of

(i) the probability of selecting the best candidate;
(ii) the average number of candidates assessed by DM before one is selected; and
(iii) the quality of the candidate selected.

We illustrated our analysis with the example of 9 candidates. Applying the Standard Method, DM assesses the first 3 candidates but does not select any of them; instead, comparing them and later candidates with each other until one candidate is preferable to all the earlier candidates (including the first 3). If this never happens, the 9th candidate is selected by default.

The two Reserve Methods hold in reserve the best of the first 3 candidates and may choose her instead of the 9th candidate. Reserve Method A does so automatically, whereas Reserve Method B does so only if she is not preferable to the 9th candidate; if the 9th candidate is preferable, she is the choice on the last round. Both Reserve Methods raise the probability of selecting the best candidate from 41% for the Standard Method to 71% for Method A and 74% for Method B. Method A does not change the expected number of candidates that are assessed (6.3), and Method B increases it slightly (6.7). In sum, the Reserve Methods' big advantage over the Standard Method is to substantially increase the probability of selecting the best candidate.

The Score Method is comparable to the Reserve Methods in the overall quality of the selected candidate — our third criterion. The expected score of the selected candidate is equivalent to an expected rank of about 1.5 for the ordinal methods, which is to say that the candidate selected will tend to be about midway between the best and next-best candidates.[6] This puts her considerably above the expected rank of 2.9 given by the Standard Method.

The Secretary Problem is clearly relevant to actual hiring decisions made in the real world, in which DM must strike a compromise between choosing a reasonably good candidate and seeing many before deciding on one who may well be the best overall. In fact, DM may well try to keep the best candidate seen so far in reserve, or set a high initial threshold that can be gradually lowered as the supply of candidates dwindles. Our Reserve and Score Methods show how these factors come into play when the restrictive conditions of the original problem are loosened.

[6] The comparison is approximate, reflecting that the (cardinal) expected score is about midway between the expected scores of the ordinally best and next-best candidates.

References

Bruss, F. Thomas (1984) "A Unified Approach to a Class of Best Choice Problems with an Unknown Number of Options," *Annals of Probability*, 12, 3: 882-889.

Ferguson, Thomas S. (1989) "Who Solved the Secretary Problem?" *Statistical Science* 4, 3: 282-296.

Freeman, P.R. (1983) "The Secretary Problem and Its Extensions: A Review," *International Statistical Review*, 51, 2: 189-206.

Gardner, Martin (1960) "Mathematical Games," *Scientific American*, 202, 2: 150-153.

Hahn, Niklas, Martin Hoefer, and Rann Smorodinsky (2022), "The Secretary Recommendation Problem," *Games and Economic Behavior*, 134, 199-228.

Li, Longjian, and Alexis Akira Toda (2022), "Incentivizing Hidden Types in Secretary Problem," Preprint. https://arxiv.org/pdf/2208.05897.pdf

Sakaguchi, Minoru (1961), "Dynamic Programming of Some Sequential Sampling Design," *Journal of Mathematical Analysis and Applications*, 2, 3: 446-466.

Smith, M. H., and J. J. Deely (1975). "A Secretary Problem with Finite Memory," *Journal of the American Statistical Association*, 70, 350: 357-361.

Wikipedia (2023), "Secretary Problem," https://en.wikipedia.org/wiki/Secretary_problem, accessed 2023-02-17.