University of Pisa
Scuola Superiore Sant'Anna

Department of Economics

Master of Science of Economics

# Dynamic threshold models of optimal stopping and the role of intelligence

Supervisors:

Mauro Sodini

Luca Polonio

Candidate:

Silvio Ravaioli

Academic Year 2013/2014

*"A good decision is based on knowledge*
*and not on numbers."*
Platone

# Abstract

The aim of this thesis is to investigate the secretary problem using a twofold approach: compare the performance of models adopting different degrees of complexity and validate the results testing which of them fits behavioural data the best.
From a theoretical point of view we modify two of the main assumptions of the traditional secretary problem. We use absolute values randomly drawn from a normal distribution instead of rankings for the applicants and remove the constraining 0-1 utility function. In this way the goal is to maximize the expected gain instead of the probability of picking the highest value within a sequence of alternatives.
This problem is less rigid than the traditional one and allows picking non-candidate options. We believe this variation can model some real choice situations such as limited time offers and decisions involving economic variables characterized by stochastic fluctuations.

We initially consider previously studied heuristics and modify them in order to study the performance in case they accept second-best options. After that we solve the problem using a computational demanding *dynamic threshold model* that updates information using observed alternative.
Additionally, we design two new heuristics, called *cutoff-threshold* and *maximum-threshold rule*, that include features of both previous heuristics and threshold model.

120 subjects participated in an optimal stopping experiment with 30 different sequences and various tests measuring intelligence, working memory and risk aversion. Our cutoff-threshold rule fits behavioural data better than all the other models, showing the preliminary role of exploration and the trade-off between the optimal policy and the already identified heuristic. Additionally we study the learning process during the task and the role of intelligence of participants. Results confirm previous studies on the role of intelligence as major predictor of performance, both as value earned and precision of the model adopted.


KEYWORDS: Optimal stopping, Behavioural economics, Heuristics, Threshold, Intelligence, Learning

# Contents

# 1 General introduction

Economics is the discipline that studies choices.
Consumption, production and investment choices. Political, fiscal and social choices. In each of these cases there are different options, and the decision maker selects one among them in order to pursuit her goal.
Decision problems are usually described using mathematical language in order to characterize optimal solution. Real problems are simplified assuming perfect knowledge about all the available alternatives. When stochasticity is introduced we use laws that describe the degree of uncertainty, avoiding ambiguity.

In order to show a simplified case we propose a textbook example.
Shopping situation is an archetypical case of decision problem: if you go to the supermarket to buy a chocolate bar you will reach the appropriate shelf and check all the available varieties. They can differ by cocoa percentage, flavour, size, price. Given your taste, willingness to pay (and gluttony) you will chose one (or more!) among the chocolate bars. Economic theory tells us that you are trying to maximize a fictitious utility function under perfect knowledge of the alternatives and unlimited memory and computational ability. Actually the choice scenario described above is less obvious than we could expect at a first sight.
To take a perfectly informed decision we need to check all the different options, compare them on more dimensions, and remember all these information.
To do this we need, first of all, plenty of time.
In the real world, given the same shopping situation, we could answer in different ways. For example we could buy a certain product because we use to do so or a friend suggested us, because there is only one option close to our desire, or just pick a random one. We cannot ex-ante be sure these strategies correspond to maximizing choices.

Virtual and real auctions represent a clear example of choice with non-traditional features. In these situations bidders need to take fast decisions and there is a competition with other potential buyers, so the bids take into account many factors including expected replies by the other bidders.
Information about the product sold may be partial, or at least not so detailed as in

the case of a chocolate bar or a bottle of shampoo at the supermarket.

Finally we can consider the case of booking a flight. The decision maker knows the destination and the dates she desires to fly, so it is easy to find the solution with the best price among different companies. The only information that is not available is how much prices will change tomorrow, or the day after, and so on until the day of the departure. The price can both rise or fall, but once you bought your ticket it is not possible to take advantage of a more convenient offer.

All these typologies of decision framework and many others show characteristics that differ from the traditional choice problem, as described in the chocolate example. It is not always possible to get the best option, or forecast the best moment to place an offer, given available information.

Researchers put a lot of effort in the design of more complex problems and solutions. On one side it is possible to include limited information, or partial information with the possibility to improve the knowledge, for example putting some effort in a costly search activity.
On the other side, time plays a crucial role. So-called recall option, i.e. the possibility to change your mind and accept something you refused before, can be absent or limited. For example you can consider interest rates on different home mortgages and try to find the lowest rate. Interest rates, like many other economic variables, are characterized by stochastic fluctuations. If rate increases you cannot return to older, lower one.

One of the goals of search theory is to explain how individuals behave when they have imperfect or incomplete market information.
In particular, we decide to focus on the so-called "secretary problem", an optimal stopping problem that represents an hiring situation: the decision maker chooses when she wants to terminate the search by hiring an applicant, and gives up the opportunity of hiring another applicant potentially better, who has yet to be interviewed.

The first statement of the problem appeared in the February 1960 column of Martin Gardner in *Scientific American*. Lindley (1961) solved the problem first in a scientific journal, and three years later Chow et al. (1964) generalized the solution.
Since then researchers studied many variations on the traditional problem. Results change according to the assumptions of the problem, for example full, partial or no information, research costs and structure of the utility function. Seale and Rapoport (1997) designed three main heuristics to describe human behaviour in a task corresponding to the problem. These rules, called cutoff rule, candidate count rule and successive non-candidate rule, are used in all the situations with no information, where relative ranking is the only crucial element used to choose.
Different assumptions about the distribution of applicants and utility functions (as in Muller 2000, or Costa and Averbeck 2015) bring to a distinct solution model based on acceptance regions.

In our problem we operate two main changes on the assumptions of the traditional secretary problem. First of all we substitute the relative ranking information with cardinal values. The values are randomly picked from a normal distribution whose parameters are not known. The decision maker can gain information and estimate mean and variance by observing the applicants. Hence we describe a partial information problem.
The second element is given by the utility function (or, to be more precise, the gain function) assigned to the decision maker. In the traditional problem there is a 0-1 utility function: picking the best is the only goal and choosing the second best is considered as bad as picking the worst one. Here we reward the decision maker with exactly the same cardinal value she chose.

We believe that applying these two modifications[1] we obtain a more appropriate framework to describe simplified economic choice situations.
First of all we allow learning during the single sequence: at the beginning of a sequence we start with no information, and our confidence over the estimates improves

[1]We are actually modifying three of the assumptions, as the first change affects two separate aspects (data generation and information given to the decision maker).

during the task because of the higher number of applicants met.[2]
Second, absolute values play a role. When we compare two results we are interested not only in the ranking but also in the distance between them.
Finally we do not limit our pursuit to the choice of the best one. Relaxing this assumption means allowing a higher heterogeneity in the choices. In the traditional version of the problem the selection of a non-candidate applicant, i.e. an applicant that is not the best one met so far, is considered as a mistake. We want to include the possibility to choose lower values without a tight classification criterion.[3]

We study the normal-distribution variation of the secretary problem in three steps. We initially study the performance of old heuristics described in the literature. In particular we redesign them and allow non-best selection.[4] Our analysis shows that under some conditions the performance of the new rules is no worse than the original.

We also analyze the corresponding optimal strategy. The *dynamic threshold model* we describe operates according to the principle of acceptance regions presented by Muller (2000). If we include uncertainty about the parameters of the distribution, the optimal policy simultaneously manages the exploration and exploitation aspects. Each new observation is in the process of estimate of the parameters, and is evaluated as possible choice.

Finally we introduce two new rules. Since the threshold model is very demanding from a computational point of view, we embed some of its features in the new rules. *Cutoff-threshold* and *maximum-threshold rule* are much simpler but their performance are significantly higher than traditional heuristics.

In order to test empiric relevance of the newly introduced rules, 120 subjects partici-

[2]It is important to stress that the design of our task seeks to avoid a priori beliefs about the unknown parameters. As we discuss later, at the beginning of each sequence the decision maker has no information about the parameters of current sequence.

[3]Adopting absolute values we also include heterogeneity driven by risk aversion and attitude to manage high and low values.

[4]We compare the models when choice is possible only in correspondence of a best candidate or in the case second-best or third-best candidates can be selected.

pated to a paid experiment. Participants solved 30 problems with different sequences and completed different questionnaires to measure intelligence (Raven Advance Progressive Matrices and Cognitive Reflection test), risk aversion (Holt & Laury risk aversion test) and working memory (free recall Working Memory Test and Wechsler Digit Span test).
Participants were rewarded between 10 and 20 euros according to the accuracy of their choices in the optimal stopping task under normal-distribution variation.
We analyze performance, waiting time, average gain and number of choices made according to the described rules.
After that we study the role of intelligence, risk aversion and memory as predictors of behaviour.
In particular we focus on the role of intelligence and the learning process during the experiment.

The dissertation is made up of two distinct parts.
The first part is dedicated to a review of the literature. We describe the main features of an optimal stopping problem and in particular the assumptions of the traditional secretary problem. We show the main heuristics in the literature and compare their performance. Previous experiments play an important role in the review of the literature. We recap the main behavioural studies on traditional problem and its variations. Finally we summarize the tests used during the experiment in order to measure intelligence, memory and risk aversion, and we briefly discuss the role of learning.
In the second part of the dissertation we describe our variation of the problem, the optimal policy and the performance of both old and new heuristics. We discuss the role of heterogeneity in modelling accuracy and learning.
Finally we describe the experimental paradigm and show behavioural results.
We conclude with a brief discussion about the areas of applications of the problem, a summary of main results and some possible implementations.
Appendices contain the instructions that subjects received during the experiment and the pseudo-codes used in the design of models.

# Part I

# Optimal stopping: a review of the literature

## 2 Optimal stopping problems

Optimal stopping theory is concerned with problems of choice of the time to take an action (Ferguson, 1967), and in general with situations of choice among a sequence of alternatives without recall, in order to maximize an expected reward or minimize an expected cost in a task with some degree of uncertainty.
Problems of this kind are studied in statistics and operational research, but simplified cases can be found in everyday life: hire an employee (Stein et al., 2003), replace a machine (Monahan, 1982), buy an airplane ticket (Shu, 2008), open a champagne bottle, virtual auctions, mating and stock trading (Rothschild, 1974) are only some examples.

Stopping problems are defined by two objects:
- a sequence of random variables $X_1, X_2, \dots$
- a sequence of real-valued reward functions $Y_1(X_1), Y_2(X_1, X_2), \dots$
The problem consists of the choice of the time to stop to maximize the expected reward. For each $n = 1, 2, \dots$ after observing $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$ elements of the sequence it is possible to stop and receive the known reward $y_n(x_1, ..., x_n)$ or continue and observe $X_{n+1}$.

A stopping rule consists of a sequence of functions of the observed elements, representing the conditions for stopping:[5] $\phi = (\phi_1(x_1), \phi_2(x_1, x_2), ...)$.

This class of problems usually does not allow free recall option, hence it requires an algorithm able to sequentially process data. In the machine learning terminology

[5]In case of a random stopping rule $\phi$ represents the probability of stopping.

this type of algorithms is called on-line algorithms (Rethmann and Wanke, 2001). This term indicates that the problem is to be solved in real-time without the ability to have all the data and then determine the solution, which in the best choice problem case would be trivial.

The two main approaches to solve optimal stropping problems are the martingale approach[6] and the Markovian approach.[7] Under certain conditions optimal stopping problems can be described by Markov processes by a measurable function. For further information see Ferguson (1967) or Tsitsiklis and Van Roy (1999).
Monahan (1982) identified models and algorithms dealing with partially observable Markov decision processes, a generalization of a Markov decision process which permits uncertainty regarding the state of a Markov process and allows for state information acquisition. These models have wide applications in areas such as quality control, machine maintenance, internal auditing, teaching strategy, search algorithms and even analysis of health-care systems.
An important tool in a vast number of stochastic problems is the Bellman equation (Bellman, 1957). This equation can be used in a maximization (or minimization) problem with decisions to be taken at different times, optimizing the path with the recursion of its periods.

[6] A martingale is a stochastic process for which the expectation of the next value in the sequence is equal to the present observed value even given knowledge of all past events. The martingale approach is based on backward induction, for cases in discrete and finite time, and essential supremum, for discrete and continuous time and finite or infinite horizon.

[7] A Markov process is a stochastic process that satisfies the Markov property, i.e. the conditional probability distribution of future states of the process (conditional on both past and present states) depends only upon the present state, not on the sequence of events that preceded it.

# 3 Traditional secretary problem

The secretary problem is a classic example of optimal stopping problem (Lindley 1961, Gilbert and Mosteller 1966). The first version of the problem was written by Martin Gardner, who published it in a 1960 column of *Scientific American*.
The name comes from a possible real-world scenario of the problem: trying to hire the best secretary among a certain number of candidates, with imperfect knowledge about their quality, adopting sequential interviews, and without recall option.
Given a finite number of alternatives shown sequentially, the decision maker can choose whether accept current option or refuse it. The only information available for each element is its relative quality with respect to previous alternatives met, assuming the possibility of sorting ex post all the alternatives.
If current alternative is accepted the sequence is immediately stopped and none of remaining options can be chosen. Otherwise, the sequence continues to the following alternative and previous ones cannot be selected. The alternatives are finite, hence if the decision maker reaches the last one, she is forced to select it.

The assumptions of the secretary problem, as stated by Stein et al. (2003) are:
1. The number of applicants for employment, $n$, is finite and known;
2. A single position is available;
3. The $n$ applicants are interviewed (evaluated) sequentially in random order with each of the $n!$ orderings equally likely;
4. The decision maker can rank order all applicants from best = 1 to worst = $n$ without ties;
5. Only the relative ranks of the applicants are made known to the decision maker, rather than a numeric measure of utility;
6. Once rejected, an applicant cannot be recalled;
7. The decision maker's goal is to maximize expected payoff: 1 if the best applicant is selected, 0 otherwise.

Starting from this basic framework, many variations were implemented: utility functions different from classical 0-1, explicit costs for acquiring new information and absolute values instead of relative rankings are only some of the examples that will be presented in the following section. Freeman (1983) and Chun (2000) realized two

reviews of articles that modify one or more of the assumptions listed above.
In all these cases there is a common element: the trade-off between refusing current offer, to seek higher values, and accepting it, in the case current option is the best one or it maximizes expected value. This trade-off corresponds to the two distinct phases we can recognize as exploration and exploitation, whose combination represents a widespread dilemma (for example see March, 1991).

The traditional secretary problem is not only an *optimal stopping problem*, but also a *best choice problem*, because the goal is to select the best one of the $n$ alternatives (assumption 7). In this cases the maximization of the expected value corresponds to the maximization of the probability of selecting the best option. This aspect is firmly linked with the absence of recall option (assumption 6): without it the problem could be trivially solved gathering all the data available and choosing the best option at the end of each sequence.

Among the possible stopping rules, three classes of heuristics (cutoff rule, candidate count rule and successive non-candidate rule) were investigated by Seale and Rapoport (1997, 2000) as descriptive models of the behaviour of subjects who took part to experiments. All these rules are determined by the value of a single parameter, and by the length $n$ of the sequence.
For sake of clarity, we name "applicant" any element of the sequence, and "candidate" any element greater than all the previous ones.

The cutoff rule consists of ignoring the first $r-1$ applicants and choosing the first candidate thereafter.
The candidate count rule considers only candidate alternatives; it counts the number of candidates identified and selects the $c$-th candidate.
Successive non-candidate rule (or non-candidate count rule) operates in the opposite way; it counts the number of non-candidate applicants and selects the first candidate met after an uninterrupted subsequence of $k-1$ or more non-candidates.
In all the three cases the last element of the sequence is selected if no applicant satisfies the rule. This aspect does not improve the performance of the cutoff model, whereas it slightly increases the success rate of remaining heuristics. In fact it allows

the selection of the best option if it occupies the last position of the sequence, even if it does not satisfy the selection criteria.

The cutoff rule includes the optimal rule, i.e. the rule able to maximize the probability of picking the best.
We analyze separately each of these three heuristics as descripted in Stein et al. (2003) with a detailed description of the three rules and the corresponding performances.

## 3.1 Cutoff rule

The cutoff rule applied to the secretary problem consists of ignoring the first $r-1$ applicants ($r \in N, r \in [1, n]$ ) and choosing the first candidate thereafter.[8]
Given the value of $n$, there exist a value of $r^*$ able to maximize the probability of choosing the best applicant, and this value corresponds to the best policy. We study the performance of the rule for all the values $r$ can take, and find $r^*$.

We define $W_n(r) = P(success)$ as the probability of selecting the best out of $n$ applicant adopting the cutoff rule with the parameter $r$.
We know the terminal condition $W_n(n+1) = 0$, i.e. it is impossible that the best is not included in the sequence of applicants.
For $r = 1, 2, ..., n$ we have

$$W_n(r) = \frac{1}{r} \cdot \frac{r}{n} + \left(1 - \frac{1}{r}\right) \cdot W_n(r+1)$$

The probability that the applicant $r$ is a candidate is $1/r$, whereas the probability that the candidate in position $r$ is the best is $r/n$. If the element at position $r$ is not a candidate, we move to applicant $r+1$, adopting the cutoff rule we would have using $r+1$ as value of the parameter.
We can explicitly define the probability of success as

$$W_n(r) = \frac{r-1}{n} \sum_{k=r}^{n} \frac{1}{k-1}$$

[8]We remind that any element of the sequence is an applicant, and we name candidate only those element greater than all the previous ones.

To find optimal value for $r$ we proceed as in Gilbert and Mosteller (1966) and approximate $W_n(r)$ for large $n$ in such a way that $r/n$ tends to a given number $p$ that represents the fraction of applicants to ignore before choosing the first candidate ($p \in R, p \in (0,1]$ ).
As $n$ tends to infinite we define $V(p)$ the probability of success, $\Delta p = \frac{1}{n}$, $\frac{r}{n} = p$ so $\frac{1}{r} = \frac{1}{n \cdot p} = \frac{\Delta p}{p}$.
We can now rewrite the probability of success as
$V(p) = \Delta p + (1 - \frac{\Delta p}{p}) \cdot V(p + \Delta p)$ with terminal condition $V(1) = 0$.

We can more easily rewrite this expression as $\frac{V(p+\Delta p)}{p} = 1 + \frac{V(p+\Delta p) - V(p)}{\Delta p}$
and for $\Delta p \to 0$ we get

$$V'(p) = 1 - \frac{V(p)}{p}$$

Finally we solve the differential equation: $V(p) = -p \cdot \ln(p)$.
To find the optimal strategy we derive with respect to $p$ and set the first order condition $V'(p) = 0$. The solution is $p^* = e^{-1}$, corresponding to $r^* = N/e$.
This value corresponds to a global maximum because $V''(p) = -1/p < 0$, so the function is concave.

We can recognise this result intuitively in the graph below (from Stein et al., 2003).

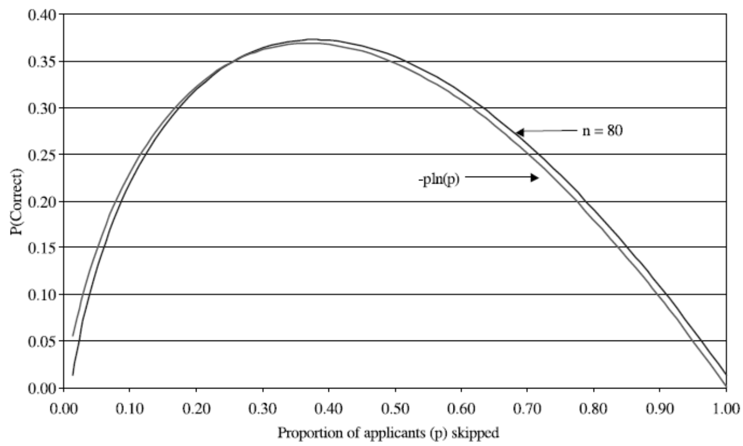


Figure 1: Performance of cutoff rule for different values of p

In the figure we can see the plot of the success probability $V(p)$ with respect to the proportion $p$ of applicants skipped in the cases of $n = 80$ and for large $n$ (generally

$n > 200$). The plot is almost flat in correspondence of the peak, representing just slightly lower success rates in the neighborhood of the optimal value.

## 3.2 Candidate count rule

The candidate count rule considers only candidate alternatives; it counts the number of candidates identified from the beginning of the sequence and selects the $c$-th candidate ($c \in N, r \in [1, n]$ ). No applicant is ignored using this rule.

We define $g(i, j), 1 < i \leq j \leq n$ the probability that there are exactly $i$ candidates among the first $j$ applicants (considering a given number $n$ of elements in the sequence).
$g(i, j)$ is computed for all the feasible combinations of $i$ and $j$.[9]
The condition that defines whether applicant $j$ is a candidate is:

$$P \ (i \text{ candidates in the first } j) =$$
$$= P \ (i-1 \text{ cand in } j-1, \ j \text{ cand}) + P \ (i \text{ cand in } j-1, \ j \text{ not cand})$$

We can hence write this expression as

$$g(i, j) = \frac{1}{j} \cdot g(i-1, j-1) + \left(1 - \frac{1}{j}\right) \cdot g(i, j-1) \quad for \ 1 < i \leq j \leq n$$

and since $g(j, j-1) = 0$ (it is impossible to have more candidates than applicants) we obtain

$$g(j, j) = \frac{1}{j} \cdot g(j-1, j-1) \quad for \ 1 < j \leq n$$

We know $g(0, j-1) = 0$ so we get

$$g(1, 1) = \left(1 - \frac{1}{j}\right) \cdot g(1, j-1) \quad for \ 1 < j \leq n$$

Applicant 1 is always a candidate so we can impose an initial condition $g(1, 1) = 1$. In this way we can start the recursion of the previous equations

$$g(j, j) = \frac{1}{j!} \ , \ \ g(1, j) = \frac{1}{j}$$

We have a special case if the final applicant happens to be a candidate. In this case we assume that the candidate is selected even if the decision conditions are not met.

[9] We introduce the value of the parameter $c$ after all the terms are computed.

If the last applicant is a candidate it will certainly be the overall best, so it is always rational to select it.

With the recursive method it is possible to compute $g(i, j)$. Hence we fix a value of the parameter $c$. The decision rule will bring to a successful choice if either:
- candidate $c$ occurs in the first $n-1$ applicants and the candidate is the overall best one or
- up to $c-1$ candidates appear in the first $n-1$ applicants and the last applicant is the overall best one.

The probability of success in each of these two cases can be computed as follows:
- there are $c-1$ candidates in the first $j-1$ applicants and candidate $c$ corresponding to applicant $j$ is the overall best; we sum the probability over all possible positions of $j$

$$\sum_{j=c}^{n-1}\Big[g(c-1, j-1)\cdot\frac{1}{j}\cdot\frac{j}{n}\Big]$$

- there are $c-1$ or less candidates in the first $n-1$ applicants and the last applicant $n$ is a candidate which turns out to be the overall best one

$$\sum_{i=1}^{c-1} g(i, n-1)\cdot\frac{1}{n}$$

Summing the two separate components we get the probability that the chosen candidate is the overall best one

$$P(success) = \frac{1}{n}\Big[\sum_{j=c}^{n-1} g(c-1, j-1) + \sum_{i=1}^{c-1} g(i, n-1)\Big]$$

## 3.3 Successive non-candidate rule

The successive non-candidate rule counts the number of non-candidate applicants and selects the first candidate met after observing a subsequence of $k-1$ or more successive non-candidates ($k \in N, k \in [1, n]$).

The applicant selected in this way is better than at least $k-1$ previous items plus any candidate before. Furthermore, being an heuristic, it captures the desire of choosing a relatively high value after a sufficiently long block of non-candidates.

If we consider fixed values for $k$ and $n$, we can describe the sequence of numbers as a Markov chain. At any stage $r$ (total number of applicants interviewed) $i$ is the current number of successive non-candidate.
If applicant $r+1$ is not a candidate, then $i+1$ is the next state of the system. This happens with probability $1 - \frac{1}{r+1} = \frac{r}{r+1}$.
Instead, if the applicant is a candidate (which occurs with probability $\frac{1}{r+1}$ we can face two cases:
- if $i < k-1$ this candidate is not selected, the new state of the system is $i = 0$ and the sequence continues;
- if i $\geq k-1$ then the candidate is selected and the process stops.[10]

Let $\gamma(i, r)$ be the probability of success given the state of the system $i$ at stage $r$, with $n$ fixed and known. This function represents the situation where $r$ out of $n$ applicants were interviewed and the $i$ immediately preceding applicants observed were non-candidates.
We want to compute the value $\gamma(0, 1)$ which is the probability of success starting at the beginning of the problem. The first applicant is always a candidate, so the process always starts in state $i = 0$ on stage 1. We calculate the values of $\gamma(i, r)$ in correspondence to the following stages as intermediate computations:

$$\gamma(i,r) = \begin{cases} \frac{1}{r+1} \cdot \gamma(0, r+1) + \frac{r}{r+1} \cdot \gamma(i+1, r+1) & if \quad 0 \leq i < k-1 \\ \frac{1}{r+1} \cdot \gamma(s, r+1) + \frac{r}{r+1} \cdot \gamma(i+1, r+1) & if \quad k-1 \leq i < r-1 \end{cases}$$

Additionally we add here terminal conditions $\gamma(i, n) = 0$ for all $i = 1, ..., n-1$ (if we continue the process until the end of the sequence with a run of non-candidates, we have selected no candidate so we have no chance of success).
Furthermore, $\gamma(0, n) = 1$: the process ends in state 0 at time $n$ if the last applicant is a candidate but we had not seen enough successive non-candidates (e.g. state $i$ was less than $k-1$ on the previous stage). In this case we are forcing the selection of a candidate that may appear in correspondence of the last stage $n$.
As we already noted the last applicant is certainly the best value if it turns out to

[10] When the candidate is selected the process enters an absorbing state $s$. A state of a Markov process is called absorbing if once in that state there is no chance of leaving that state. The existence of an absorbing state (e.g. choose a candidate and stop the sequence) satisfies assumption 6.

be a candidate.

If we make a selection on stage $r+1$ and enter the absorbing state $s$, the probability that this candidate is the best one is

$$\gamma(s, r+1) = \frac{r+1}{n}$$

Thus the second equation above can be simplified to

$$\gamma(i, r) = \frac{1}{n} + \frac{r}{r+1} \cdot \gamma(i+1, r+1) \quad if\ k-1 \leq i \leq r-1$$

We can treat $k-1$ as an absorbing state, since we do not need to differentiate $k-1$ from higher values. Hence

$$\gamma(k-1, r) = \frac{1}{n} + \left(\frac{r}{r} + 1\right) \cdot \gamma(k-1, r+1)$$

In this way we can recognize $\gamma(k-1, r-1) = W_n(r)$ because $\gamma(k-1, r-1)$ is the probability of success assuming we start in state $k-1 > 0$ after $r-1$ applicants. Applicant $r-1$ is not a candidate but the rule will select the next candidate, if any, from applicants $r, r+1, ..., n$. The cutoff rule with parameter $r$ skips the first $r-1$ applicants and then selects the next candidate after that, thus the two decision rules will make the same selection from this stage forward. Taking into account these simplifications we can rewrite the system as

$$\gamma(i, r) = \begin{cases} \frac{1}{r+1} \cdot \gamma(0, r+1) + \frac{r}{r+1} \cdot \gamma(i+1, r+1) & if \quad 0 \leq i < k-1 \\ W_n(r+1) & if \quad i = k-1, k \leq r \end{cases}$$

for $r = 1, ..., n-1$, with $\gamma(i, n) = 0$ for all $i = 1...k-1$ and $\gamma(0, n) = 1$.
Computing the elements in correspondence to all the stages the value for the initial position $\gamma(0, 1) = P(success)$ is straightforward.

## 3.4 Performances of the rules

As we noticed by the description of the rules, the performance of each models depends on the value of the unique parameter.
Figure 2 shows the percentage of success choices in a traditional secretary problem with $n = 100$ in correspondence to the different level of parameters ($r$, $c$ or $k$ according to the heuristic, within the range [1,100]).

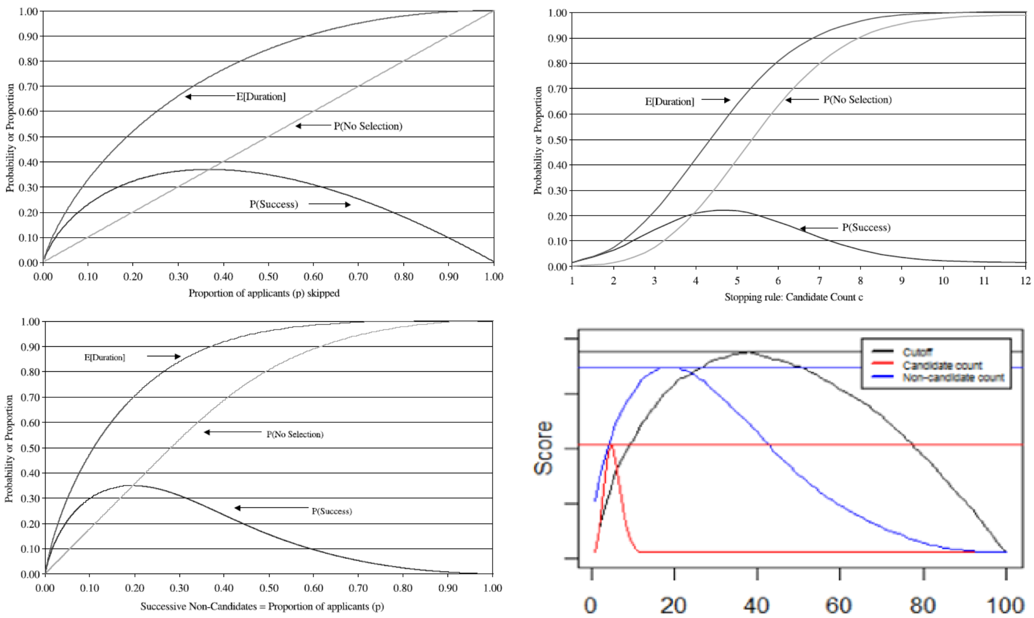


Figure 2: Performance of cutoff, candidate count and successive non-candidate count

The rules display slightly different success values in correspondence of optimal values of the parameters. In particular, cutoff rule and successive non-candidate rule differ only by a 7%, while candidate count rule performs 18% worse than the optimal policy. Observing this result we remark the performance of cutoff is robust to the total number of observations, while the performance of the candidate count rule declines with increasing values of $n$ (0.217 at $c = 5$, $n = 80$, 0.174 at $c = 6$ and $n = 500$, 0.146 at $c = 9$ and $n = 5000$). Hence we conclude this rule produces only moderately good results for small $n$, and the performance rapidly declines with higher number of applicants.

Successive non-candidate rule shows the opposite characteristic. The probability of success significantly increases moving from a low value of observations: with $n = 6$ and $k = 3$ $P(success) = \gamma(0, 1) = 0.275$, for $n = 5000$ and $k = 977$ $\gamma(0, 1) = 0.3476$, a probability of success of the best application at about 94% of the optimal value of cutoff with the fraction $e^{-1}$.

If we adopt the fraction $p = \frac{k}{n}$ for successive non-candidates,[11] we can define the

[11] We define fraction $p$ with $n$ sufficiently high.

result in terms of $p$: wait for a run of non-candidates equal to the proportion $p$ of all applicants available, then pick the first candidate. If $p^* = \frac{977}{5000} = 0.1954$ is used, then the probability of success is maximized within this class of heuristics (there is little variation in the optimal $p^*$ for larger values of $n$).

Successive non-candidate rule is an excellent heuristic in correspondence to the optimal value, with properties similar to the optimal policy, however the heuristic is considerably more sensitive to departures from its optimal parameter $p^* = 0.1954$ than is the cutoff rule, corresponding to lower tolerance for deviations.

Generally we can see the shape of the curves are dramatically different. The graph corresponding to the cutoff rule is smooth on the top. This means the result is very similar to the optimal policy if the cutoff parameter is in the neighborhood of the best.

Finally we can compare the performance of each heuristic in correspondence of its optimal parameter. Table 1 shows the optimal parameter for each rule in the case of $n = 100$ observations, success percentage, no choice percentage (i.e. the probability of reaching the last element of the sequence without finding any candidate satisficing the requirements) and average position chosen.

| | Cutoff | Candidate | Non-candidate |
|---|---|---|---|
| Optimal parameter | r = 28 | c = 6 | k = 5 |
| Success % | 31 | 21 | 28 |
| No choice % | 30 | 39 | 32 |
| Average position | 51 | 63 | 56 |

Table 1: Performance of heuristics, values for n=100 and 100,000 simulations

On average we can see not only the success rate is different, but also the other characteristics. In particular the average position is higher in correspondence of the candidate count rule with respect to competing models.

# 4 Experiments and variations of the problem

Since the Seventies, researchers run many experiments in order to test human behaviour and ability to act accordingly to optimal policy. Most of the studies about optimal stopping show a general tendency to early stopping. Decision makers choose a value before finishing the exploration share of the sequence according to cutoff model. This behaviour causes a suboptimal result due to the lack of analysis of the distribution of alternatives.

Late stopping, the opposite result, is found in the experiments where subjects know the distribution of options and in particular are informed about an hypothetical ideal alternative.

In Section 3 we described the seven assumptions of the traditional secretary problem. Each of them can be relaxed in different ways. Variations on the traditional task have an impact on the optimal policy and in some cases subjects act differently.

Different assumptions offer more realistic formulation of the standards problem. Rather than assuming a single position Gilbert and Mosteller (1966) and Sakaguchi (1979) consider a secretary problem in which more positions are available.

Pressman and Sonin (1972), Rasmussen and Robbins (1975) and Gianini-Pettit (1979) consider the case where the decision maker only knows the distribution of the value of $n$.

The assumption of random arrival of applicants can be replaced by other assumptions; Cowan and Zabezyk (1978) assume applicants are interviewed at time points of a Poisson process. Relative rank can be replaced by other assumptions, probability distribution can be known in advance or it can be learned during the interview (Sakaguchi 1979, Gilbert and Mosteller 1966, De Groot 1970, Stewart 1981, Petruccelli 1982).

The absence of recall option can be generalized in several different ways. Smith (1975) introduced the possibility that an applicant, if selected, has some probability of not being available. Another generalization allows the recall of one of the last $m$ rejected applicants, and it can be selected only if it is still available.

Assumption about the 0-1 utility function can be generalized in several ways. The decision maker can receive utility $u_i$ if the applicant selects the $i$-th best, or directly receive a value proportional to the cardinal number she chose.

To show the results of main experiments we divide this section into three parts. In the first one we describe the first experiments and the main evidences from tasks with the assumptions described in the traditional version of the secretary problem. After that we show some variations on the assumptions and how they can impact behavioural results. Finally we show some of the insights about more radical variations on the problem, such as group version of the secretary problem, multi-attribute implementation and neural studies.

## 4.1 First experiments and main evidences

Rapoport and Tversky (1970) run the first experiment about search costs using the secretary problem framework. In their task they used sequences with offers randomly drawn from normal distributions whose parameter values were known. On each trial the decision maker could either pay a fixed cost and sample another offer or stop the search and receive a payoff equal to value of last offer minus the total cost of search. Seven subjects took part to the experiment within a study about distribution detection. Throughout the first period they participated in a signal detection experiment and they were asked to distinguish two normal distribution with mean 1630 or 1797 and common standard deviation equal to 167. During the last period of experiment they took part to the optimal stopping study.

The two different distributions were known, $n$ was bounded and known, decision makers saw all the values even after choice.

The experiment was conducted under different conditions: with and without recall option, using different length of the sequence (8 or 24), and different constant cost conditions (including no cost).

Behavioural data show that decision makers solved most of the problems optimally, but in general subjects tended to stop earlier than predicted by optimal policy. Rapoport and Tversky compared the choices with a cutoff point rule, as the subjects should stop the search whenever an offer exceeds some fixed cutoff (threshold value).

Results are aligned with the model studied by Kahan et al. (1967). According to their study the point is selected at the beginning of each problem and is allowed to vary from one problem to another. In this task a different cutoff point can be selected for every length $n$ and every cost condition $c$, and the decision maker stops the search on a given trial with the first offer whose value exceeds the predetermined cutoff point. If no value is observed the decision maker should continue sampling until the last stage. A weaker version of the model allows the decision maker to change the point at the beginning of every trial. In this way it is possible to explain about 90% of the choices. Some common behaviours cannot explained by this model. In case of recall option subjects sometimes refuse a value and then change their mind and take it back. Even when recall option is not allowed decision makers can refuse a value, then choose a lower value later in the sequence, paying extra search costs.

This experiment is one of the first behavioural studies on this topic. The concept of cutoff point is similar to the threshold value we describe in case of perfect knowledge about the distribution. The task is not a best choice problem, and does not contain an exploration part. The design of the experiment wants to test Kahan decision model, but the weak version does not specify if changes in the cutoff point are explained as computational ability, learning process or endogenous risk aversion.

Kogut (1990) presented a problem of search for low prices. Given different prices of a product randomly drawn from a known distribution, the problem includes a constant cost of search and perfect recall. Assuming risk neutrality, results show again systematic early stopping.

As in previous experiment the search should continue until the point where marginal costs are greater than marginal benefits. The task asked to search for the lower price or the highest wage from a distribution of offers. Values were randomly generated from a known uniform distribution of ten prices. Subjects had to buy the product and resell it at a constant price (1.50 dollars).

To test whether searching behaviour is influenced by sunk cost, i.e. the resources already spent on an activity, different conditions were adopted. The baseline cost for a new observation was 0.08 dollars. Adding a further cost for the first search (equal to 0.40 dollars) should have no impact on the choices. Results show previously born cost are not ignored, and subjects tend to continue searching as profits

fall and recall later during the task. Decisions seem to be based on a total rather than a marginal value, in contrast to the theoretical prediction.
Related literature explain sunk costs influence decisions in two distinct ways: the evidence show that with sunk costs individuals are more likely to abandon their search (Kahneman and Tversky, 1979) rather than more strongly committed to it (Thaler, 1980). Subjects stopped early 38% of times with respect to optimal choices, risk aversion is a possible explanation.

The exploration-exploitation scenario was firstly proposed by Stigler (1961). In its model the decision maker visits $n$ stores, obtains a price quotation from each one and then buys from the lowest-price store. Given a distribution of prices and the cost for acquiring new information, the search should stop when the cost is greater than the expected gain for searching a new shop. This value depends from the cost of search $c$ and the distribution of prices (that is assumed known). The key concept of "reservation price" is used to define a sequential optimal rule that describes the ex-ante number of shops the decision maker should decide to visit.
The expected gain for searching after $n$ observations is $G_n$. The decision makers should choose $n$ so that $G_n \geq c > G_{n+1}$. Decreasing the cost of search or increasing price dispersion rises the intensity of search, i.e. it increases the number of shops to visit.

Kahan et al. (1987) tested the effect of different distributions in the optimal stopping problem and the role of individual or group task. In this case the task is very close to the problem described in the traditional version of the secretary problem. Subjects had to find the largest of a set of 200 different numbers sampled from one of three distributions: a triangular distribution with positive skew (beta distribution with parameters a=1, b=2), one with negative skew (beta distribution with a=2, b=1) and a rectangular distribution (beta distribution with a=1, b=1). All the distributions were adjusted to have mean 1250, with standard deviations 884, 442 and 722 respectively. The task was performed individually or in small groups.
Numbers were written on cards and participants won 2 dollars every time they chose the highest card. The distribution adopted was not known in this problem. On average participants observed over 150 on 200 before choice in group conditions and

120 in individual condition. The difference between group and individual condition is significant. Values of average waiting were not not significantly different under the three conditions of distribution.

Results show late stopping: only 12 out of the 88 participants chose a value previous to position 74 (i.e. the value $\frac{200}{e}$ corresponding to the cutoff position to move from exploration to exploitation parts of the problem) and almost half of the participants did not choose early enough and did not pick any candidate, reaching the end of the sequence.

In order to explain results the authors propose a very broad subjective-cutoff model tested with an arbitrarily selected cutoff between 20 and 100. This general model is able to explain 85% of choices. Main violations to the model include selecting an early candidate in a position before 20, select an applicant that is not a candidate, or reject a candidate observed after position 100.

Seale and Rapoport (1997) compared the ability of the main heuristics (cutoff, candidate count and successive non-candidate rules) to fit behavioural data in a task with all the main features of the traditional problem: maximize the probability of selecting the best observation from a finite set of alternatives inspected sequentially. 50 subjects took part to the experiment under two conditions about the length of the distribution ($n = 40$ and $n = 80$) and run 100 trials. In case of best choice they received immediate confirm they chose the top ranked option and won 0.30 dollars in case $n = 40$ or 0.50 dollars for $n = 80$.

Despite many answers can be done according to more than one heuristic, result support the cutoff decision rule as the most adopted model. Authors use the term "cutoff threshold value" in the description of the model. The term threshold used here has a different meaning with respect to what we use later in our model.

On average subjects performed well and selected the best candidate 31.9% of the times. This compares to a success rate of 37.2% if they had known and used the optimal strategy. Subjects tended to select a candidate too early in the process, waiting on average 21% of time instead of 37% as prescribed by optimal policy.

The early stopping effect is confirmed by the overall cutoff value adopted: 13 instead of 16 for $n = 40$ and 21 instead of 30 for $n = 80$.

Data show learning effect only under condition $n = 40$: there is a significant block effect, and 15 out of 25 subjects moved closer to optimal solution comparing two halves of the task.

## 4.2 Experiments with variations on the assumptions

Seale and Rapoport (2000) introduced an experiment where the number of applicants is finite but unknown.

In this secretary problem with uncertainty the decision maker knows the a priori distribution of the number $n$.

50 subjects run 100 trials under two condition (maximum number of applicants $n = 40$ or 120). The actual number of applicants $m$ was randomly and independently drawn on each trial and could assume any value between 1 and $n$ with equal probability.

The optimal policy for this problem suggests to observe a fraction of applicants equal to $\frac{r^*}{n} = \frac{1}{e^2} = 0.135$ with a probability of success that approaches $\frac{2}{e^2} = 0.2707$ (Presman and Sonin, 1972).

If the applicant accepted by the subject was the top ranked one, the subject was paid 0.30 dollars for the trial, otherwise she was not paid at all. Results show no learning effect under both the conditions. Furthermore there is a strong early stopping bias. Authors explain it as an endogenous fixed cost per observation.

In the experiment designed by Rothschild (1974) decision makers learned about the probability distribution while they searched for it. At the beginning of each task the probability distribution from which they are searching is unknown. Observing values from a distribution it is possible to estimate the parameters. The author showed the qualitative properties of optimal search strategies are in many instances the same as in the simpler case when the distribution is assumed known.

After every price quotation the decision maker can accept current value or pay an amount $c$ to receive another price quotation, with no privilege of recall. Rothschild assumes the decision maker has a prior distribution and as the search continues she gathers more information about the distribution of prices, which she assimilates by using his prior according to Bayes's rule.

If the knowledge of prices is represented by the parameters $\mu, \rho$ which are updated after each new information, the optimal strategy is calculated by induction. Since

the correct distribution of prices is unknown, acceptable regions of price change as information increase, so that optimal Bayesian search procedure cannot be characterized by a single reservation price, as it is with the known distribution.
In our problem we operate in a slightly different way. We do not assume any prior about the parameters of the distribution and do not include search costs.

Bearden (2006) adopted rank-based selection and cardinal payoffs as an extension of the secretary problem.
The decision maker can observe up to $n$ applicants whose values are random variables drawn from an uniform distribution on the interval $[0, 1]$. She receives a payoff $x_t$, the realization of random variable $X_t$, for selecting the $t$-th applicant. For each encountered applicant she learns whether it is the best so far. It is important to stress that this information is different from the relative position. We can say that the decision maker receives the information $I_t = 1$ if the value is a candidate, $I_t = 0$ otherwise.
The objective is the maximization of the expected payoff. Differently from traditional problem this is not a pick-the-best task. The decision maker gets the true value she chose. Since the values are drawn from a uniform distribution on $[0, 1]$ the expected value of a candidate in position $t$ is $E_t = E(X_t|I_t = 1) = \frac{t}{t+1}$.
Since $\frac{dE_t}{dt} > 0$, if it is optimal to select an applicant with $I_t = 1$ then it is optimal to select an applicant with $I_{t+k} = 1$, for every $k \geq 0$. Hence it is never optimal to select an applicant with $I_t = 0$.
The optimal solution for this problem is given by a certain cutoff position $c_n^*$ corresponding to the optimal value $V_n^* = V_n(c_n^*)$.
The optimal policy for this problem is slightly different from the traditional problem. The decision maker should skip the first $\sqrt{n-1}$ applicants, then select the next encountered applicant whose value is a maximum.
The author states that the payoff scheme he presented is more natural compared to the classical secretary problem. We claim that the assumptions adopted have tight limits due to the limited information. As the decision maker does not get the information about the rank she cannot discriminate if current value is a second or third best. According to the position within the sequence it could be optimal to accept a second best observation. This scheme does not admit this case and the

author does not justifies this strong assumption about limited information.

## 4.3 Further experiments

Shu (2008) designed a problem of future-biased search with a secretary problem with a quest for the ideal.

In the experiment participants received an endowment of 9,000 dollars and were asked to buy 15 tickets with the possibility to use one reward coupon and two 30% discount coupons. Tickets had different prices randomly drawn from a known distribution of prices. The optimal selection strategy prescribes certain rules in order to determine the right timing to use the discounts.

According to results subject spent on average more per trial and used the free trip ticket to save lower value. This bias (use later the tickets and spend more) is a violation of the usage rules and can be explained as participants are overly focused with finding the "perfect match".

Two different conditions on the task (higher second value and less extreme results) led to improved performances. This evidences stress the role of focalism.

The optimal policy describes a series of thresholds for the ticket price at which the free ship should be used in early period. Late stopping phenomenon corresponds to an extended search beyond optimal endpoints when looking for the best option. The myopic behaviour can be explained as searchers overestimate the probability of a desired outcome and underestimate the value of a second-best outcome.

Bearden, Murphy and Rapoport (2005) proposed a multi-attribute extension of the secretary problem.

In their experiment the decision maker observed the relative ranks in the $k$ attributes and received a payoff for each attribute, according to the absolute rank.

They run experiments with symmetric and asymmetric payoffs with two attributes. Subjects took part to 100 trials with $n = 30$ and $k = 2$ and were paid for a single randomly selected trial.

The optimal policy for this problem describes a set of cutoffs for each feasible pair of relative ranks. The algorithm accepts an applicant if the expected gain exceeds a certain threshold depending on payoffs and position.

Results show a strong evidence of early stopping, as subjects attribute dispropor-

tionate weight to select an applicant who is acceptable on both attributes instead of being high on just one of them. This choice criterion fits a modified satisfying rule, that can be considered important in many daily situations.
The authors found a strong learning effect through the task. Regressing the mean earning for each trial onto the trial numbers indicate earnings increased with experience. Subjects also tended to search longer with experience, and these evidence can be linked as subjects seem to have learned that searching longer improves payoffs.

Lee and Paradowski (2007) arranged a group decision making experiment.
Groups of five participants had to choose a value from a sequence of five random numbers between 0 and 100 (uniform known distribution). The goal is to choose the maximum of the sequence without recall.
Group interaction was mediated only by networked computers, with no other free interaction. After each observation each participant was asked to give an initial individual decision. Everyone received the recommendations of the other group members, then the subjects provided a potentially revised decision.
The authors run the experiment under three different conditions about the group interaction and decision: consensus to accept from all group members, majority, or acceptance of an appointed group leader.
Under some conditions groups significantly outperformed even their best members. Results show people do not often revise their decisions, but in the consensus and leadership conditions participants were more conservative in their initial decisions. This conservatism removes the individual bias toward choosing values too early in the sequence.
Only in the majority condition people continued behaving as they did individually and the group showed the same bias in decision making.

Costa and Averbeck (2015) conducted a neural study in an optimal stopping task.
They collected behavioural and fMRI data during the task designed as a threshold crossing problems (as a special case of optimal stopping problems, see DeGroot, 1970). The values of the threshold defining stopping versus continuing were calculate dynamically after each new piece of information, and the number of remaining samples was taken into account. As in previous cases the trade-off between declining

sufficient options to make an informed choice and not missing a good option can be modelled formally using a Markov decision process.
The participants had to select an high-ranking item from sequential sampling task. The problem was presented under two conditions about the length of the sequence (lists of 8 or 12 element).
Participants were rewarded 5 dollars for the best option, 3 dollars for the second best, 1 dollar for the third best.
Observations were sampled form a Gaussian distribution with a normal-inverse-$\chi^2$ prior (Gelman et al., 2004).
The threshold model adopted to explain data was implemented with a cost-to-sample estimated for each individual subject in order to better fit data. The cost-to-sample significantly improved the model's prediction with respect to the ideal policy in 16 of 32 subjects.
Comparing "take current option" versus "choice to decline it" the neural activation pattern is similar to a threshold crossing problem. It is possible to distinguish activations in the anterior insula, dorsal anterior cingulate, ventral striatum, and parietal-frontal areas. These brain regions are implicated in reward representations and encode the threshold crossing that triggers decisions to commit to a choice.
This pattern of activation is highly similar to the activation in Furl and Averbeck (2011) for a related information sampling task.
Given a set an internal thresholds the decision maker samples until she get an option that crosses current threshold. Each sample serves both to update one's estimate of the underlying distribution as well as being a choice candidate.
The utility $u$ of a state $s$ at time $n$ can be described by the model

$$u_n(s_n) = Max_{a \in A_{sn}} \left\{ r_n(s_n, a) + \int p_n(j|s_n, a) \cdot u_{n+1}(j) \; dj \right\}$$

where $A$ is the set of available actions and $r$ is the reward obtained if action $a$ is taken.
The problem described here is similar to our task. The main difference is due to the utility function the authors adopted, with a reward linked to the relative ranking within the trial.

# Part II

# Dynamic threshold models

## 5 Introduction

In the previous parts we focused on the features of the traditional secretary problem. Experiments on the basic task enquire the role of waiting time, and many studies underline that early stopping is the main bias in behavioural data (Rapoport and Tversky 1970, Seale and Rapoport 1997).
Furthermore, it is not possible to discriminate good from optimal policies, as the number of best choices is the only index of the performance. In the traditional framework the decision maker gains no reward for a near miss. In some experiments researchers introduced significantly smaller rewards for second best alternatives (Costa and Averbeck 2015).
Finally, in the traditional problem there is no space for learning. Relative ranking information about the applicants do not require any assumption about the shape of the originating distribution. Successive observations can be used only to calculate the probability of picking a candidate and that current candidate is the overall best option.

The potentiality of an optimal stopping problem allows us to include both these aspects and enquire their impact on final performance.
After a brief recap of the traditional secretary problem we define the changes in the assumptions that lead to our normal-distribution variation.
We consider first previously studied heuristics (Stein et al., 2003) and compare their performances adopting the traditional version and in the case we allow non-best choices.
After that we solve the problem of maximization of the expected gain for each sequence, i.e. each trial of the task. To do so we define a dynamic threshold model that updates the estimates of the unknown parameters after each observation.
We also design two new heuristics that include features from both threshold model and previous rules. These heuristics are less demanding from a computational point

of view but their performance are better than old heuristics because of learning through observations.

Results collected with the experiment inform us about the aspects we discussed about in this part.
First, we can test which of the rules we described (old and new ones) fits the best behavioural data.
Second, we compare behaviour and common biases in this variation with the evidences from the traditional version of the secretary problem.
Third, we use data collected with other tests measuring intelligence, working memory and risk attitude to study the drivers of the performance, if any.
Finally we enquire the role of learning though the task. If there is an improvement in the performance, we consider if it is generated by a reduction in the early stopping bias or by better estimates of the unknown parameters.

We conclude describing some of the areas of application of the results, with particular reference to financial and consumption topics, and a few possible implementations of the task in order to enquire more detailed aspects of the problem.

# 6 Normal-distribution variation

## 6.1 A brief recap of the traditional secretary problem

Before discussing our variation on the optimal stopping problem we want to briefly recap the features of the benchmark problem (Stein et al., 2003).

The assumptions of the traditional secretary problem are:

1. The number of applicants for employment, $n$, is finite and known;

2. A single position is available;

3. The $n$ applicants are interviewed (evaluated) sequentially in random order with each of the $n!$ orderings equally likely;

4. The decision maker can rank order all applicants from best $= 1$ to worst $= n$ without ties;

5. Only the relative ranks of the applicants are made known to the decision maker, rather than a numeric measure of utility;

6. Once rejected, an applicant cannot be recalled;

7. The decision maker's goal is to maximize expected payoff: 1 if the best applicant is selected, 0 otherwise.

Furthermore, we want to remind that unless otherwise indicated there is no explicit cost in moving from an element to the following one of the sequence.

Among the possible stopping rules, Seale and Rapoport (1997, 2000) investigated three classes of heuristics as descriptive model of the behaviour of subjects who took part to an experiment. All these rules are determined by the value of a single parameter, and by the length $n$ of the sequence.

From now on we use the name "applicant" to indicate any element of the sequence, and "candidate" to define an element greater than all the previous ones.

The cutoff rule consists of ignoring the first $r - 1$ applicants and choosing the first candidate thereafter.

The candidate count rule considers only candidate alternatives; it counts the number of candidates identified and selects the $c$-th candidate.

Successive non-candidate rule (or non-candidate count rule) operates in the opposite way; it counts the number of non-candidate applicants and selects the first candidate met after an uninterrupted subsequence of $k - 1$ or more non-candidates.

In all the three cases the last element of the sequence is selected if no applicant satisfies the rule.
The cutoff rule includes the optimal rule, i.e. the rule able to maximize the probability of picking the best under the assumptions listed above.

## 6.2 An easy variation: maximize the expected value

Before studying the normal-distribution variation, we can start analyzing a simplified problem, shown in Muller (2000).
The author considers a problem with a fixed number $n$ of alternatives (i.i.d. real random variables $X_1, ..., X_N$) that can be observed sequentially and assumes the decision maker knows the common distribution of the observations. It is possible to add a fixed cost $c$ per observation, and this assumption (with $c \gg 0$) is strictly necessary if there is no limit to the number of observations available.
The reward for choosing observation $i$ is $R_i = X_i - c \cdot i$.

In order to solve the problem we have to find the optimal sequence of stopping regions, that is the set of values the optimal policy should accept in each possible step of the sequence.
It is crucial to stress that we are assuming initial perfect knowledge of the distribution, and no learning effect occurs during the sequence.

Defining $T$ the random variable corresponding to the position of the value chosen by the stopping rule, and $R_T$ the random gain obtained by this stopping rule, the problem is $MaxR_T = MaxE(X_T - c \cdot T)$.

We define $J_k(s)$ the optimal expected reward if there are still $k$ possible observations and $s$ is the momentarily available offer.
We can see $J_0(s) = s$ and $J_k(s) = Max\{s, EJ_{k-1}(X) - c\}$.
It is optimal to accept $s$ if and only if $s \geq r_k^* \equiv EJ_{k-1}(X) - c$.

We can calculate reservation level $r_k^*$ by recursion:
$r_1^* = EX - c$ and $r_k^* = EMax\{X_1, r_{k-1}^*\} - c$.
Reservation level and optimal expected reward are increasing in $k$ and for $k \to \infty$

they converge to $r^*$ and $J(s) = Max\{s, r^*\}$.

For simplicity Muller makes some examples in his paper using offers that are uniformly distributed on a certain interval.
For example given the interval $[0, 100]$ and a fixed cost $c = 2$ the reservation levels, calculated by iteration, are $r_1^* = 48, r_2^* = 59.52, r_3^* = 65.71, ..., r^* = 80$.

This structure based on reservation levels is similar to our threshold model, but requires initial perfect knowledge about the distribution. Muller's models fit experimental data collected by Sonnemans (1998) using a known discrete uniform distribution in the interval $[1, 100]$.

## 6.3 Normal-distribution variation and performance of previous rules

We consider now a more difficult framework. With respect to traditional problem we change three assumptions.[12] First, we use cardinal values instead of rankings. These values are randomly generated by a normal distribution. Second, the decision maker knows how values are generated, but does not know the parameters of the distribution. Mean and variance differ through the sequences, and the decision maker does not know the range. Third, we substitute the 0-1 utility function with a linear correspondence between gain and chosen value.

Like in previous variation on the problem we observe absolute values and want to maximize the expected value choosing whether accept or refuse current alternative, with a finite number of alternatives (we assume now no cost for gathering new observations). The difference we introduce here is that we do not know a priori the distribution of the values.
We assume the only information available at the beginning of a sequence is that all the values are extracted from a common normal distribution, but the two parameters that describe the distribution (mean and variance) are not known. We also assume

[12] First and second change are strongly connected and represent the implementation of a unique change that effects both assumptions 4 and 5.

the decision maker knows the values are independently and identically distributed.

Similarly to traditional problem solved by cutoff rule we can clearly recognize the trade-off between early and late stopping. In particular in this case the information about the distribution are "updated" because at each step the decision maker gets a new variable randomly drawn from the unknown distribution.
This means that, in correspondence to each step, the beliefs about the distribution change (i.e. the estimates of the values of mean and variance are updated) and are more reliable (it is possible to attribute an increasing degree of confidence).

To sum up, the assumptions[13] of this normal-distribution variation on the secretary problem are:
1. The number of applicants for employment, $n$, is finite and known;
2. A single position is available;
3. The $n$ applicants are interviewed (evaluated) sequentially in random order with each of the $n!$ orderings equally likely;
4*. The values of the applicants are randomly generated by a normal distribution;
5*. The decision maker knows the cardinal value of the applicants and knows they are generated by a normal distribution, but she does not know the parameters (mean and variance) of the distribution;
6. Once rejected, an applicant cannot be recalled;
7*. The decision maker's goal is to maximize expected payoff: her gain is equal to the cardinal value of the selected applicant.
As in the traditional problem we include no explicit cost in moving from an element to the following one of the sequence.

In this new framework we do not consider a 0-1 utility function. The three main heuristics in the literature are focused on selecting the best applicant, so a priori we do not know their performance in this new problem.
In the traditional problem picking the second best was as bad as choosing the worst applicant. In our case this is not true. We can modify cutoff rule, candidate count

[13] Asterisks indicate the assumption we changed from the traditional problem.

rule and successive non-candidate rule in order to allow second-best choices.

In general we can characterize any of the three heuristics by two parameters:

- the parameter $p$ that indicates how many observations wait before starting the selection process. This corresponds to the single parameters adopted in traditional formulation of the heuristics ($r$, $c$ or $k$);

- a new parameter $k$ that defines the relative rank of the highest value to compare to current option. If we consider $k = 1$ the rules coincide with original definitions. With $k = 2$ current option is compared with the second best alternative met so far.

  We could say otherwise that $k$ changes the definition of candidates: a candidate is an applicant with a value higher than the $k$-th best one met until that point of the sequence.

Figure 3 shows how performance of each heuristic vary according to the combination $(p, k)$ of parameters. Performance is now defined as average win instead of probability of success.
To allow comparison we obtained results by repeated simulation (100.000 simulations for the results shown in the figure) using standardized normal distributions, i.e. normal distribution with $\mu = 0$ and $\sigma = 1$, and sequences with $n = 20$.

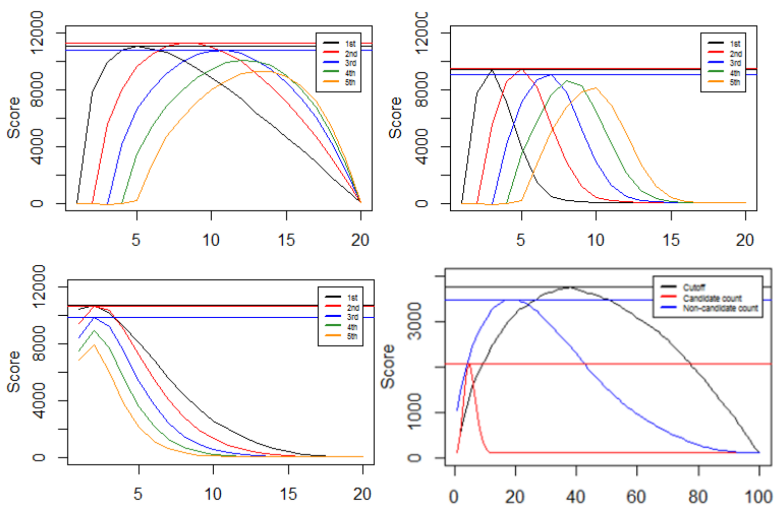


Figure 3: Performance of the models under different parameters

Looking at the curves we can see that the rules have similar performances with $k = 1$ and $k = 2$ using different parameters of $p$. If we increase the ranking of the benchmark candidate we can notice the optimal value of $p$ is higher as well. For example the cutoff rule with $n = 20$ shows the best performances with the combinations $k = 1, p = 5$ (score 1.18), $k = 2, p = 11$ (score 1.19) and $k = 3, p = 13$ (score 1.17).
Similar results for candidate count rule ($k = 1, p = 3; k = 2, p = 5; k = 3, p = 7$).
Successive non-candidate rule displays a different structure. The shape of the curves are similar, with the score for $k = 2$ slightly below the original rule, and with evidently worse performance in correspondence to other levels. The optimal value of $p$ is not different in these cases ($p = 3$).

# 7 Dynamic threshold models of optimal stopping

## 7.1 Solving the problem: the dynamic threshold model

In a full information problem like the one discussed in Muller (2000) options are random variables drawn i.i.d. from a distribution assumed to be known to the decision maker before the search starts.
In partial information problems this assumption is relaxed. In our case we assume the decision maker knows that the distribution is normal, but she must learn its mean and variance during the search process.
Finally, in a no-information problem the distribution is unknown and cannot be learned during the search process, as in the traditional formulation of the secretary problem. In that case the unique information is the relative ordinal value of each alternative.

To solve our variation on the problem we have to find the region of acceptance for each combination of values observed and number of items left. We solve first the problem in the case of known parameters of the normal distribution, then we introduce uncertainty about one or both the parameters and solve the problem with uncertainty.

In all the cases we use a threshold model, similar to those described in Rapoport and Tversky (1970), Bearden, Murphy and Rapoport (2005) and Costa and Averbeck (2015). For each stage we accept an applicant if its value exceeds the corresponding threshold, and refuse it otherwise. In case of partial knowledge every new applicant is used as an information to improve the estimate of the parameters.

In the case of perfect information we use the classic framework of initial expected value maximization problem. As in Muller's paper the stopping rule is defined by the stopping regions, i.e. the regions of values such that accepting is a better option than refusing.
Let $s$ be current available offer, and $k$ the number of possible observations left before the end of the sequence.
The problem consists in the maximization of the expected value of the reward ob-

tained at each stage according to the optimal policy. By recursive method we want to maximize the expected value of the reward at the beginning at the sequence.
We define $J_k(s)$ the optimal expected reward if there are still $k$ possible observations and $s$ is the available offer.

By definition $J_0(s) = s$, i.e. when the decision maker is facing the last element (so there are 0 observations left) she can do nothing but accept current option.
If there are one or more observations left, the decision maker should take the action (accept or refuse current option) that maximized the expected value. According to this $J_k(s) = Max\{s, EJ_{k-1}(X)\}$ where X represents the random variable.
The reservation level, corresponding to the threshold value, is equal to the expected value of the gain at the following step.

$$r_k^* \equiv EJ_{k-1}(X).$$

Hence it is optimal to accept option $s$ if $s \geq r_k^*$.

The reservation level at the last-but-one step is equal to the expected value of the random variable. If the decision maker declines current option, she will reach the last observation and is forced to accept it. In case the distribution is known

$$r_1^* \equiv EX = \mu$$

If the mean is unknown, the decision maker can estimate its value using previous observations since we are facing a partial information problem. In this case the reservation level is modified into

$$r_1^* \equiv EX = \hat{\mu}.$$

For $k > 1$ the expected value at the following step can be calculated by the distribution of $X$.

$$E(X_k) = pr(X \leq E(X_{k-1})) \cdot E(X_{k-1}) + pr(X > E(X_{k-1})) \cdot E(X|X > E(X_{k-1}))$$

Given the distribution of the random variable, we consider separately the two sides divided by the reservation level at the following step. The part on the left represents the values that would not be accepted. For those values the decision maker will refuse the offer and go to the next option, and her expected value is equal to the corresponding step, so it is given regardless the current option $s$. On the other side, the values on the right would be accepted as they are above the threshold value. In this case the distribution of $X$ determines both the probability and the expected value of the right tail.

We define $E(X_k)$ the expected value of the gain when there are $k$ observations left using a certain rule. We also define $E(X_1) \equiv E(X)$.
We can now write the problem

$$MaxE(X_n) \text{ with } X \sim N(0,1).$$

The reservation level at each stage is evaluated iterating the computation

$$J_k(s) = Max\{s, EJ_{k-1}(X)\}$$

After each observation the decision maker decided whether to accept or refuse current observation according to the threshold

$$J_k(s) = \begin{cases} s & (accept) \quad if \quad s \geq r_k^* \equiv EJ_{k-1}(X) \\ r_k^* & (refuse) \quad if \quad s < r_k^* \end{cases}$$

$$EJ_k(X) = Max\{X, EJ_{k-1}(X)\} = \begin{cases} X & if \quad X \geq EJ_{k-1}(X) \\ EJ_{k-1}(X) & if \quad X < EJ_{k-1}(X) \end{cases}$$

We can solve recursively. In case of perfect knowledge and standardized normal distribution $r_1^* = EJ_0(X) = 0, r_2^* = 0.4637, r^3 = 0.6752$ and so on.
Given a sequence made up of 20 elements we get a series of reservation levels in correspondence of each stage $r_{19}^*$ , ... , $r_1^*$ and the expected gain at the beginning of the trial $r_{20}^* = EJ_{19}(X) = 1.43$.

In case of imperfect knowledge about the distribution the decision maker knows that the distribution used to generate observation is a normal distribution. According to the assumptions, she may know only the mean or the variance, or none of them.
We maintain the maximization problem and the solution by recursive method, but the expected values are unknown at the beginning of the sequence as it is not possible to define the original distribution.

Observations are used to infer these values. In correspondence to each information set we can estimate the originating distribution's parameter (or parameters). The more values we know, the higher our confidence for the estimated parameters is. Because of this, the stopping thresholds are updated after every new observation. Acceptance choice also depends from the number of trials left. The closer we are to the end of the sequence, the lowest the threshold value is.

The optimal policy requires a dynamic threshold model that operates at each round with the following three steps:

1. Update the estimate of the parameters $\hat{\mu}$ and $\hat{\sigma}^2$ of the distribution using the new observation;

2. Create a series of thresholds for $k = 1, 2, ...$ depending on number of observed values $i$ and estimated values of mean and variance $\hat{\mu}$ and $\hat{\sigma}^2$: $J_k(\hat{\mu}, \hat{\sigma})$;

3. Accept the current alternative $s$ only if it is above the threshold corresponding to current number $k$ of remaining observations.

If both the parameters are known, this model collapses to the previously described one.
If only mean is unknown, the estimation of the variance is substituted with the true value and vice versa.
In the optimal policy the estimate of the mean is calculated as average of the observed values, whereas the variance is calculated as unbiased sample variance:

$$\hat{\mu} = \sum_{1}^{i} s_i \,, \qquad \hat{\sigma}^2 = \frac{\sum_1^i (s_i - \hat{\mu})^2}{i-1} \quad (i > 1)$$

$$J_k(\hat{\mu}, \hat{\sigma}) = Max\{X, EJ_{k-1}(\hat{\mu}, \hat{\sigma})\}$$

Unlike the threshold model with perfect knowledge, adopting the dynamic threshold model it is not possible to calculate the expected reward at the beginning of the sequence.

## 7.2 New heuristics for the problem

Candidate count and successive non-candidate are the main heuristics that Seale and Rapoport (1997) studied to fit behavioural data. They can be easily compared to the cutoff rule and in particular to the optimal policy for the traditional problem on dimensions including performance, average position and percentage of no choice. Heuristics can be useful to explain suboptimal behaviour. Gigerenzer and Todd (1999) describes the role of homo oeuristicus and stresses the importance of simplified rules for learning and performing well, sometimes even better than in the case of optimal policy. In case of uncertainty or ambiguity heuristics can be more flexible and reliable than a fixed rule.

Hommes (2013) describes non-optimizing heuristics in the theory of financial markets. Simple rules can be computationally less demanding than structured models, so they are easier and faster to adopt. According to Hommes observation, financial markets are characterized by high volatility, and time plays an important role in decision making. Furthermore, results depend from a large number of interacting actors, hence it is not possible to find a univocal solution for the problem.

Finally, in case of maximizing or minimizing problems there is a unique solution or a set of solutions. Heuristics allow a much higher degree of heterogeneity. In our case heterogeneity can be related to the type of rule the decision maker adopts or the parameters she considers for the choice.[14]

For our problem we have already described a generalization on previously studied heuristics. We introduced the possibility to choose non-best candidates and compared the robustness of the performance.

We introduce now two new rules specifically designed for this task[15] called *cutoff-threshold rule* and *maximum-threshold rule*.

[14] For an example of heterogeneous heuristics adopted in an optimal stopping task, see Seale and Rapoport (1997). They check which of the combinations rule-parameter can fit the best behavioural results for a group of 32 participants to an optimal stopping task.

[15] During the pilot and the experiment the participants were asked to explain the rules they adopted to take decisions. Answers from the pilot gave us some insights about possible simplifications subjects used.

As the name suggests, the cutoff-threshold rule includes features both from cutoff rule and threshold model. As in the cutoff rule, the decision maker initially observes a certain number of values as exploration. These observations do not lead to any choice and allow a first estimate of the unknown parameters.
After the cutoff position, thresholds are calculated using estimates for mean and variance. In particular we consider a cutoff position $i = 5$ and generate thresholds linearly with respect to the number of observations left.
To express the linear relation we multiply the number of observations left[16] by a fixed parameter (in this case $l = 0.1$). In correspondence to observation $i = 6$, i.e. with 14 remaining choices, the threshold value is calculated as

$$J^c(k = 14, l = 0.1) = \hat{\mu} + 1.3 \cdot \hat{\sigma}$$

In general the threshold value for the cutoff threshold rule is

$$J^c(k, l) = \hat{\mu} + l \cdot (k - 1) \cdot \hat{\sigma} \ , \quad i > 5$$

The maximum-threshold rule is a less demanding heuristic as it does not require to estimate the variance. It adopts the maximum value observed during the sequence as a benchmark for the following observations.
Also in this rule the first observations are used as an exploratory part of the task; after that different thresholds are computed.
After the cutoff position the decision maker considers sample average and maximum value observed. The reservation level is the sum of the estimate of the mean and the difference between maximum and mean, multiplied by a decreasing coefficient (a fixed parameter $l = 0.1$ that multiplies the number $k-1$ of observations left).

In our case we consider again $i = 5$ as cutoff position for the exploratory task and define $M(i) = Max(s_1, s_2, ..., s_i)$ the highest observed value. Threshold are calculated linearly with respect to the number of observations left. in correspondence to observation $i = 6$ we calculate the threshold as

$$J^m(i = 6, k = 14, l = 0.1) = \hat{\mu} + 1.3 \cdot (M(i) - \hat{\mu})$$

[16] We consider $k-1$ instead of $k$ to preserve the property of reservation level equal to the estimate of the mean in correspondence to the last but one observation.

In general the threshold value for the maximum threshold rule is

$$J^m(i,k,l) = \hat{\mu} + l \cdot (k-1) \cdot (M(i) - \hat{\mu}) \ , \quad i > 5$$

Both of these heuristics share three essential features with the optimal policy described by the dynamic threshold model.
First, they update the threshold values after each new observation considering both the number of steps left and the confidence in the estimates on the parameters.
Second, they explicitly take into account the mean, and at the last-but-one step they accept any value above average. This characteristic is based on a risk-neutral assumption and excludes any effect due to the absolute cardinal value of the observations.
Finally, the values of the thresholds are decreasing with respect to the number of choices left. This is calculated differently from the dynamic threshold model as the relation in the heuristics is linear.[17]

We want to compare the performance of the dynamic threshold model with previous and new heuristics considering full or partial information about the distribution.
Old heuristics do not use extra information if available, and because of this their performances do not change, and maximum threshold does not include the correct value of the variance if available.

Table 2 shows the results. The table contains the average performance for traditional and second-best cutoff model, candidate count, successive non-candidate, dynamic threshold, cutoff-threshold and maximum threshold.[18] Performance is considered as average gain with standardized normal distributions. Results are calculated using 100,000 simulations.
A detailed description of the choice criteria according to each rule can be also found in Appendix B - Pseudo-codes.

[17]The design of cutoff threshold and maximum threshold rules include some of the features explicitly described by participants in their comments to the choice task.

[18]Best performing parameters were adopted for the heuristics. Cutoff uses $k = 1, p = 5$, cutoff 2 includes the second best $k = 2, p = 10$. Cutoff threshold and maximum threshold use the parameters described above: cutoff in position $i = 5$ and $l = 0.1$

| **Model** | $\mu$ known<br>$\sigma$ known | $\mu$ known<br>$\sigma$ unknown | $\mu$ unknown<br>$\sigma$ known | $\mu$ unknown<br>$\sigma$ unknown |
|---|---|---|---|---|
| Cutoff | 1.1732 | 1.1732 | 1.1732 | 1.1732 |
| Cutoff 2 | 1.1864 | 1.1864 | 1.1864 | 1.1864 |
| Candidate | 1.0459 | 1.0459 | 1.0459 | 1.0459 |
| Non-candidate | 1.0948 | 1.0948 | 1.0948 | 1.0948 |
| Dynamic threshold | 1.3465 | 1.3238 | 1.3124 | 1.2864 |
| Cutoff threshold | 1.2951 | 1.2737 | 1.2662 | 1.2434 |
| Maximum threshold | 1.2326 | 1.2326 | 1.2184 | 1.2184 |

Table 2: Performance of different models under full, partial or no initial information.

As we can see the expected performance of the family of threshold rules is higher than all the previous heuristics, because they consider the distribution of values and can update information with new observations, learning during the sequence.

# 8 Modelling heterogeneity

As a first step in the description of the problem we solved the case of perfect knowledge about the parameters of the distribution and we found the reservation levels

$$J_k(s) = Max\{s, EJ_{k-1}(x)\}$$

After that we introduced uncertainty about one or both the parameters

$$J_k(s, \hat{\mu}, \hat{\sigma}) = Max\{s, EJ_{k-1}(x, \hat{\mu}, \hat{\sigma})\}$$

We want now to test the best response model with imperfect knowledge and introduce heterogeneity in the way of computing the threshold values. We consider two main elements that can change between agents and drive different choices.
The first one is the ability of estimating the parameters. We can define this ability as the result of interaction between intelligence and memory skills, that determine the accuracy in the creation of beliefs about the distribution.
Risk attitude is the second element we introduce.
Both these elements can have an impact in the evaluation of the reservation level. In particular we can assume an higher level of risk aversion decreases the threshold value, whereas the ability to estimate parameters can separate the value from the optimal one in both the directions.
We assume both these traits have little impact when the decision maker is close to the end of the sequence, whereas their importance gets bigger as the number of observation left increases.[19]

We define the impact of these factors using a coefficient $m(k, \alpha, \beta)$ that represents the gap between optimal and personal reservation level. The coefficient depends on

[19] We can intuitively explain the reason for this assumption for both the elements. Low ability in the estimate of the parameters plays a minor role when a big number of observations have been collected, but can be crucial after a few values because of the lack of a priori beliefs about the distribution and the difficulty of discriminating outliers from average values. Similarly, risk aversion cannot have a major role when only a few observations are left, while the decision maker facing a relatively high value after a few observations will be tempted to accept it and stop the sequence.

$k$ (number of observations left), a parameter $\alpha$ that describes the risk attitude[20] and $\beta$ that captures the accuracy in the estimate of the parameters using the information available.[21]

We do not distinguish the impact of the two components on the coefficient $m$ and use a unique parameter $g$ to define the gap. In this way we obtain a coefficient $m(k, g)$ that only depends on the number of observations left and on the value of the gap.[22]

$$m(k, \alpha, \beta) = m(k, g) = m(1, g)^k = (1 + g)^k$$

As we already explained this approach is able to diminish the impact of intelligence, memory and risk aversion when the number of observations left declines. The threshold in correspondence to each stage are computed by iteration as usual

$$J_k(s, \hat{\mu}, \hat{\sigma}, g) = Max\{s, EJ_{k-1}(x, \hat{\mu}, \hat{\sigma}) \cdot m(k, g)\}$$

This approach is useful to test the robustness of the performance of the threshold model in case of small deviations from the optimal policy.[23]

Again we multiply the threshold value in correspondence of each stage by the coefficient $m(k, g)$ close to 1.

We consider the interval $g \in [-0.01, 0.01]$ with 0.001 precision and compare the performance of the dynamic threshold model using the different values of the thresholds in 100.000 sequences.

The graph shows no significant difference comparing the performance of the basic model and the values close to one assigned to the coefficient. This result can be explained by the limited number of times where this slight change in the threshold causes a different choice.

[20]The value of $\alpha$ represents the risk attitude. It can assume positive values (risk seeking) or negative values (risk aversion). In case of risk neutrality, as we have for the optimal policy, $\alpha = 0$.

[21]The parameter $\beta$ describes the accuracy in the estimate. It is defined as the gap from the correct threshold value (so it actually describes the imprecision) and assumes positive values. In case of perfect accuracy, as we have in the optimal policy, its value is $\beta = 0$.

[22]This approach can be used to implement a probabilistic threshold model with fuzzy reservation levels. Given the optimal value of the threshold according to a model and a gap, it is possible to define three regions: one of acceptance and one of reject, and an intermediate region where the probability of acceptance increases as the value is closer to the actual threshold value.

[23]The detailed procedure adopted to test the robustness is described in Appendix B.

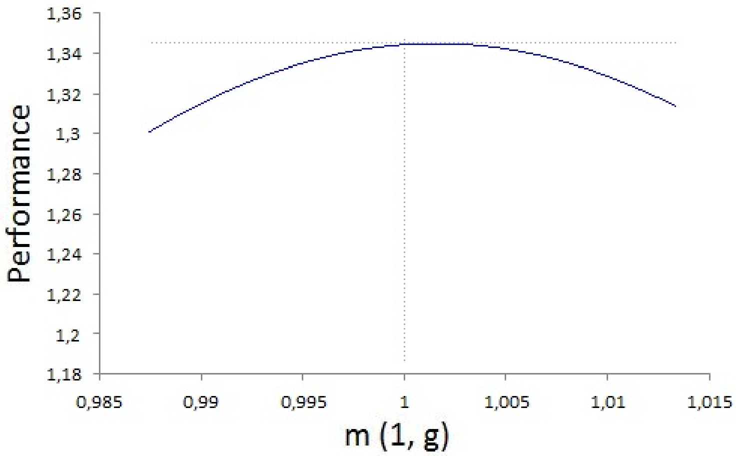


Figure 4: Dynamic threshold model with different values of g

If we consider values in a broader interval we can recognize a significant decrease in the score obtained, and plotting the performance as a function of the value of $m(1, g)$ we can clearly recognize a smooth curve with the maximum in correspondence of $m(1, g = 0) = 1$.

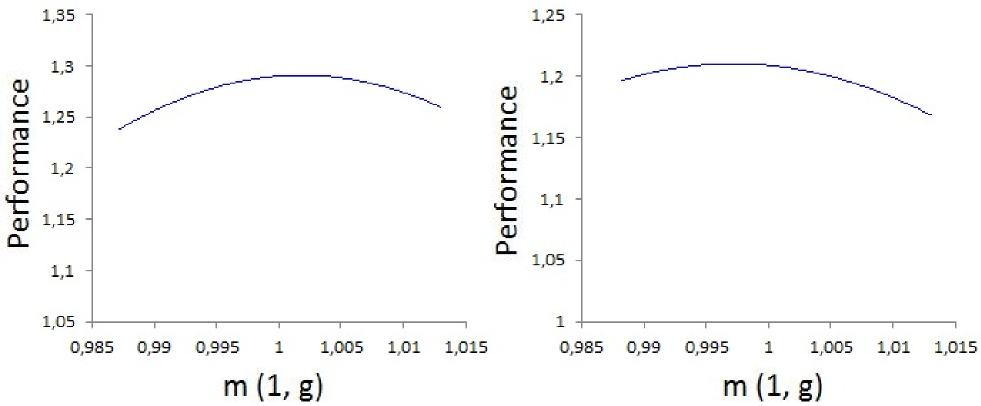


Figure 5: Robustness of cutoff threshold and maximum threshold models

Cutoff threshold and maximum threshold rules show similar results, but the best performance does not correspond to the case without deviance. In correspondence to the cutoff threshold rule we find the best results with $m(1, g) = 1.004$.
This rule is naturally more risk averse than the dynamic threshold model and a slightly higher value of the coefficient partially corrects this aspect.

It is possible to introduce and discuss other elements of heterogeneity. The first one we implicitly introduced is in the different rules that can be adopted to face the problem. Old rules and heuristics in the class of threshold models are some examples. It is possible to design a plurality of heuristics that share some of the features of the previous ones. Within each model it is possible to change the parameters, like the benchmark position and the position of the cutoff observation in the cutoff rule.

A second aspect of heterogeneity is the one we described above with reference to risk attitude. We included a characteristic of the agent that is constant through all the task, facing different trials of the same problem.
Risk aversion becomes endogenous if we shift the value of gap $g$ after every trial according to previous gains. For example we can increase risk aversion after the selection of a low value and vice versa. This change can be temporary (only due to previous choice) or cumulative (the decision maker considers all previous gains).

A similar aspect can be found with a discrimination between different cardinal values. Some decision maker can act differently when she faces high or low (absolute) values. It is easy to include this feature, and in particular recognize if different rules or parameters are adopted in correspondence to sequences with high mean (when decision makers tend to be more conservative). The waiting time before choice is the main element that can discriminate this behaviour.

Finally we can define heterogeneity in the distribution of not-systemic biases. The decision maker can learn from previous trials and improve her performance by updating the rule adopted or the computational ability (estimate the parameters, improve the precision). This aspect is strongly connected to the type of feedback she receives and the goal of the task.

# 9 Experimental results

## 9.1 Subjects, data collected and experimental design

In order to validate the results of the previous sections we decided to test which of the models fits behavioural data the best. We conducted an experiment made by two parts. In the first one subjects performed an optimal stopping task under the assumptions defined by the normal-distribution variation of the secretary problem. In the second part we collected data about possible drivers of the performance: intelligence, working memory and risk aversion.

Two sessions of the experiments were conducted at University of Trento (Department of Psychology) and University of Pisa (Department of Economics). In all the sessions subjects were undergraduate and graduate students at their respective universities (average age 21 years).[24]

120 subjects (40 males and 80 females) participated in the experiment. They were recruited by advertisements asking for volunteers to take part to a paid computer-controlled experiment concerning decisions under uncertainty with payoff contingent on performance. The experiment, including explanations and questionnaires, required one hour and a half and subjects were paid between 10 and 20 euros according to their performance during the main task.

Subjects arrived at the laboratory in small groups (usually two to five) and were seated in a room equipped with PCs. The experimenter explained the secretary problem task with particular emphasis on how the distribution change as new sequences is observed.[25] After the instructions subjects performed a sample exercise

[24]As a preliminary step we run a pilot experiment with 10 students (5 males and 5 females) at University of Pisa. Participants completed the questionnaires and a version of the optimal stopping task shorter than the final one (15 sequences instead of 30). Pilot results has been used to define details of the main task in order to focus on the main conjectures. Additionally we collected feedback about the single questionnaires and in particular about rules adopted during the optimal stopping task.

[25]The detailed description of the instructions can be found in Appendix A.1.

to have a chance to become familiar with the procedure. Participants did not receive any written copy of the instructions and were invited to ask questions either during the description of the task or after the sample exercise.

Each subject completed 30 independent trials after the sample one. Each trial consisted of a number of values equal to 20. The absolute value of current applicant was displayed on the screen after a short message that indicated the number of values already observed. As we consider crucial for different rules to know the number of observations left, we make this information available at each step, as a difference between the total number of applicants and the current number of refused offers.
In correspondence to each observation only current absolute value was displayed. Previous observations were not available on the screen and subjects did not have the possibility to write them down.
Values lasted on the screen for 2.5 seconds. In order to skip the options subjects had to wait, as numbers automatically flowed on the screen. To select current value it was sufficient to press the space bar on the keyboard. The selection process did not allow any recall option and the sequence was stopped immediately after choice. Subjects could not verify the goodness of their choices as remaining options were not shown.
Subject were informed that they would be facing values from normal distributions with unknown parameters of mean and variance. During the instructions a brief explanation of main characteristics of a normal distribution was given to participants.

The sequences were divided into three groups according to the value of the mean of the generating distribution. Participants knew distributions had different values but were not informed about the group of the sequence of the current trial. At the end of the experiment subjects drew three numbers corresponding to three sequences, one for each group (sequences with low, average and high mean). Final payment was equal to the average of the values selected by the participant in the extracted sequences, plus 3 euros as show-up fee. Because of the values of each group of sequences subjects received as a payment sums between 10 and 20 euros.

To maintain comparability and interpret experimental results, all the participants

viewed the same 30 sequences in different orders. The 20 elements of each sequence always maintain the positions.
We generated the sequences randomly from different normal distributions with various means and variances and checked the randomization splitting the 30 sequences into three block. The order subjects faced sequences is random within each block, but blocks are introduced in the same order. In this way we can compare the performance of each block and compare choices and learning effect.
Each block contains sequences with the position of choice according to cutoff rule antecedent, equal or subsequent the threshold model. Furthermore, a second check is on the position of the "best option" (randomly assigned). Finally each block includes sequences with different parameters. Values of mean are contained in the interval 4-16, standard deviation in the interval 1-3.[26]

In correspondence of each trial we save position and absolute value selected by the participant, responding time, sequence and block. We also have information about possible correspondence between participant's choice and models and a between chosen value and best element of the sequence.
Using all these information it is possible to recognize pattern of behaviour according to one or more of the models we described.

To better understand behavioural results we include in the experiment some questionnaires. In this way we want to capture the role of intelligence, working memory and risk aversion. These elements could be drivers of the performance, as they are simpler elements of the optimal stopping task.
In order to measure intelligence we adopt RAPM (Raven Advance Progressive Matrices), a test that captures the ability to recognize patterns, and CRT (Cognitive Reflection test), a series of questions that indicate the tradeoff between a rapid, intuitive answer and a more pondered result.
Holt & Laury test is adopted to measure risk aversion, whereas free recall Working Memory Test and Wechsler Digit Span test are used to collect working memory

[26]Further details about generation and randomization of sequence can be found in Appendix B.9.

parameters. The optimal stopping task is complex and contains elements from all these three categories. Intelligence and pattern recognition are necessary to estimate unknown parameters and evaluate current options, working memory plays a big role in the adoption of previous information, finally risk aversion can effect choice causing early stopping.[27]

## 9.2 Research questions and conjectures

Lee (2006) argues optimal stopping problems naturally distinguish between performance based on achieving optimal outcomes, i.e. choosing the final value, and performance based on following optimal decision processes, i.e. choosing the correct value. Simon (1976) termed these different measures procedural and substantive rationality, respectively, and noted that procedural measures are less noisy.

We enquire both the aspects analyzing the performance measured by average gain and the adherence to one of the models we defined in the previous sections.

Unless the subject is directly asked about the decision rule she used, there is no obvious and nonambiguous way to infer the rule from her responses on any particular trial. Even if we assume that the subject adheres to one rule, there is no way to determine the exact value of the model's single parameter from her response. We focus on the decision rules introduced earlier and proceeded to test them.

We focus on four main aspects using available data.

First, we can test which of the rules we described (old and new ones) fits the best behavioural data. As the problem we use is different from the traditional secretary problem we want to study the effects of the changes on the search behaviour.

Second, we compare behaviour and common biases in this variation with the evidences from the traditional version of the problem.

Third, we use data collected with other tests measuring intelligence, working memory and risk attitude to study the drivers of the performance, if any.

Finally we enquire the role of learning though the task. If there is an improvement in the performance, we consider if it is generated by a reduction in the early stopping bias or by better estimates of the unknown parameters.

[27] Appendix B contains instructions for all the questionnaires.

Our main research questions and conjectures cover this four areas.
Our research focuses on two aspects neglected by previous experiments: the role of intelligence and the class of model adopted in case absolute values of the applicants are shown to the decision maker.
Can we identify drivers of performance? Which components of intelligence have an impact on the choice in the optimal stopping task? Intelligence and working memory are candidates as drivers of performance.
Does risk attitude have an impact on the ability of selecting best alternatives and on the model adopted? We want to test if risk averse subject suffer early stopping or risk seeking subjects choose optimally or suffer late stopping.
One of our hypothesis is that risk aversion does not change during the experiment, whereas the ability of estimating the parameters improves through the experiment for all the subjects. We want to verify if high-intelligence subjects converge faster to the unbiased estimates, corresponding to higher gains.

We want to verify if learning occurs, comparing different blocks. If yes, it can be measured as average gain or precision in the adoption of one of the models.
About the role of intelligence, if any, we want to discriminate if its effect is important for all the performance (since the first sequence) or impacts learning during the experiment?[28]
Finally, with respect to old heuristics, we want to compare traditional cutoff rule with its second-best variation. In this way we want to define which of the possible definitions fits behavioural data the best.

## 9.3 Results

Behavioural data can be studied using different perspectives.
Table 3 contains correlation values and offers a complete description of the relationships between questionnaires and performance in the optimal stopping task.
In particular we are interested in two different classes of parameters that describe

[28] This question corresponds to the distinction between faster *learning by thinking* or *learning by doing* driven by trials and errors.

performance. On one side we consider average position chosen, indicating the waiting time, average gained value and number of choices corresponding to best options. Additionally we consider the adherence with different behavioural models.

| | Position | Value | Best | Cutoff | Cutoff 2 | Candidate | Non-cand | Dynamic th. | Cutoff-th. | Max-th. |
|---|---|---|---|---|---|---|---|---|---|---|
| **RAPM** | 0,1453 | 0,3179 | 0,2912 | 0,2713 | 0,3013 | 0,1750 | 0,3252 | 0,2272 | 0,3222 | 0,2106 |
| **CRT** | 0,0762 | 0,2127 | 0,2862 | 0,2125 | 0,2100 | 0,1699 | 0,2119 | 0,2210 | 0,2526 | 0,1651 |
| **Holt&Laury** | -0,1135 | -0,0179 | -0,0191 | -0,0831 | -0,1279 | 0,0155 | 0,0013 | -0,1587 | -0,0422 | -0,1315 |
| **Digit span** | 0,0873 | 0,1128 | 0,0544 | 0,0811 | 0,0628 | 0,1964 | -0,0067 | 0,0646 | 0,0812 | 0,0211 |
| **Working memory** | 0,1986 | 0,0057 | 0,1244 | 0,1089 | 0,1652 | 0,0744 | 0,0668 | 0,1627 | 0,0825 | 0,1450 |

Table 3: Correlation between score in the questionnaires and performance in the optimal stopping task and adoption of the main decision rules.

Results show a strong positive correlation between RAPM - CRT scores and the adoption of models. Working memory tests show little correlation, and HL scores (corresponding to risk aversion) are weakly negatively correlated to the performance. Intelligence appears as the strongest predictor of performance. Performing regression of the performance expressed in various terms we can see that intelligence tests (Raven test and Cognitive reflection test) are the only variables with a significant positive coefficient (1% significance).

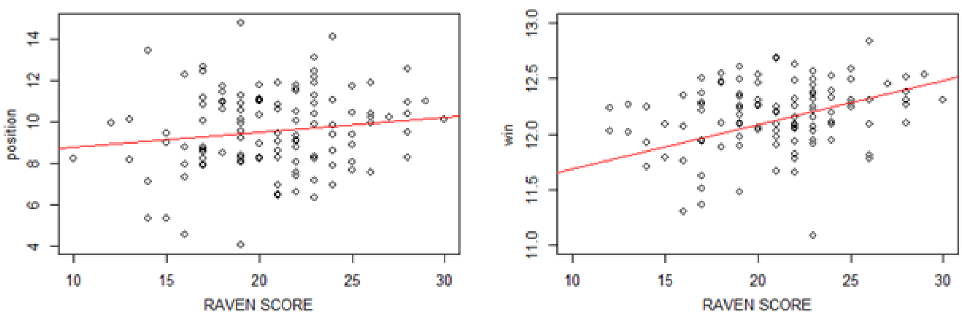


Figure 6: Distribution of performance with respect to Raven score

Figure 6 shows the performance (win) and average position selected.
In order to study the impact of intelligence we consider two separate clusters corresponding to quintiles according to the performance in the Raven Advanced Matrices Test.

Table 4 shows strong differences in all the variables. Intelligence measured by the

| CLUSTER | RAPM | CRT | H&L | Digit span | Working m. | Position | Value | Cutoff | Cutoff 2 | Dyn. th. | Cut-th. |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 15,3 | 0,1 | 4,6 | 4,2 | 6,2 | 9,1 | 11,9 | 6,4 | 5,8 | 6,8 | 7,7 |
| 2 | 19,1 | 0,3 | 5,0 | 4,6 | 6,3 | 9,9 | 12,1 | 8,2 | 7,4 | 8,0 | 9,8 |
| 3 | 22,1 | 0,5 | 5,0 | 4,4 | 6,3 | 9,4 | 12,1 | 8,1 | 8,0 | 8,0 | 9,7 |
| 4 | 25,9 | 0,7 | 4,8 | 4,7 | 6,4 | 9,9 | 12,3 | 9,3 | 8,9 | 8,7 | 11,3 |
| Average | 20,7 | 0,4 | 4,9 | 4,5 | 6,3 | 9,5 | 12,1 | 8,0 | 7,6 | 7,9 | 9,6 |

Table 4: Average values of tests and performance for groups created by RAPM score

two tests is strongly correlated (RAPM and CRT), but higher clusters are associated with higher score in the working memory tests , average position and value. In particular we recognize little difference between clusters 2 and 3, whereas groups 1 and 4 are significantly different.

Adoption of behavioural models is strongly correlated with the RAPM score, as the number of choices coherent with behavioural models always increases in the highest clusters, and this effect is particularly higher if we consider the cutoff-threshold model.

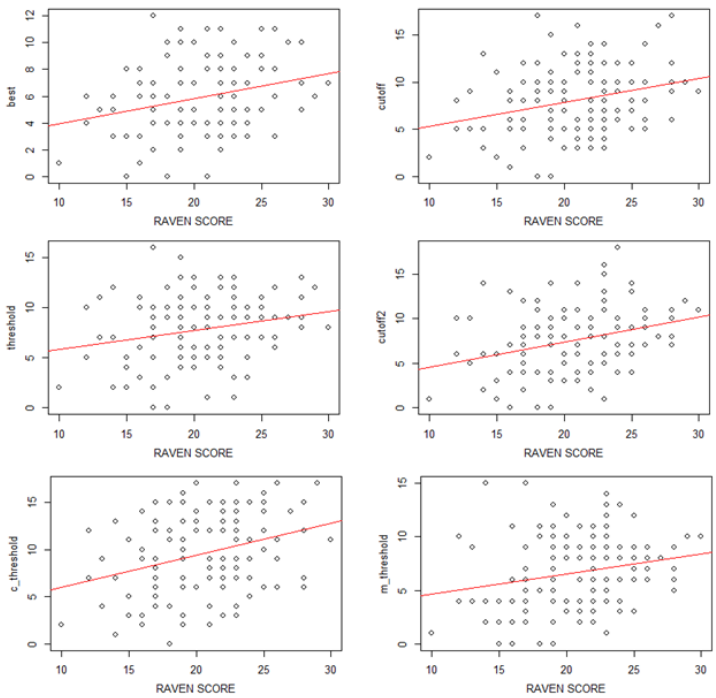


Figure 7: Number of choices according to different models

Figure 7 displays the distribution of the models adopted. Cutoff-threshold models is widely adopted, in particular for high RAPM scores (bandwidth 20-30 correct answers).

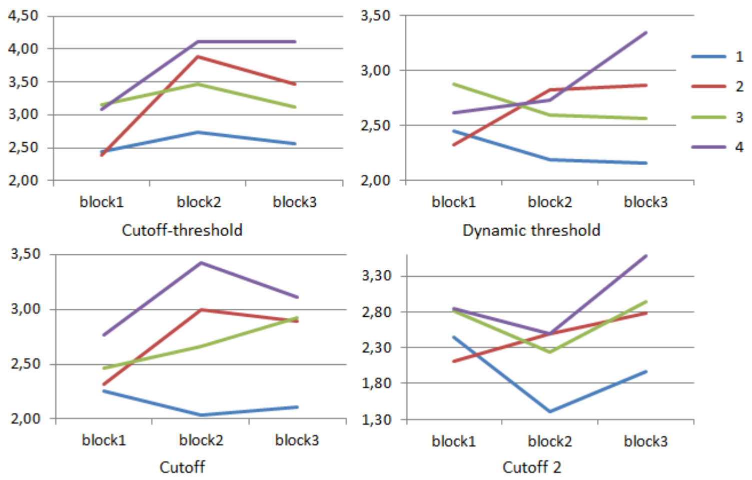


Figure 8: Learning - adoption of behavioural models during the experiment

The distribution of models adopted is also connected with the learning phenomena. Comparing average number of choices according to main models we recognize a general increase between block 1 (sequences 1-10) and block 2 (sequences 11-20). This effect is parallel to performance of the different clusters and can be studied also under different criteria to build clusters.

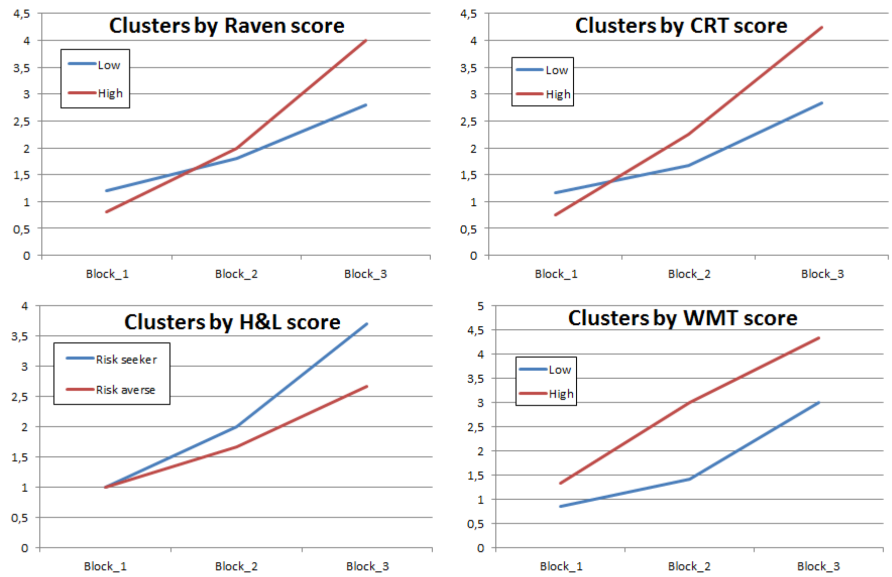


Figure 9: Learning - number of best choices according to different clusters

Division of the subjects according to RAPM, CRT, HL and WMT score always display learning effect in the number of choices corresponding to the best value within the sequence (figure 9).

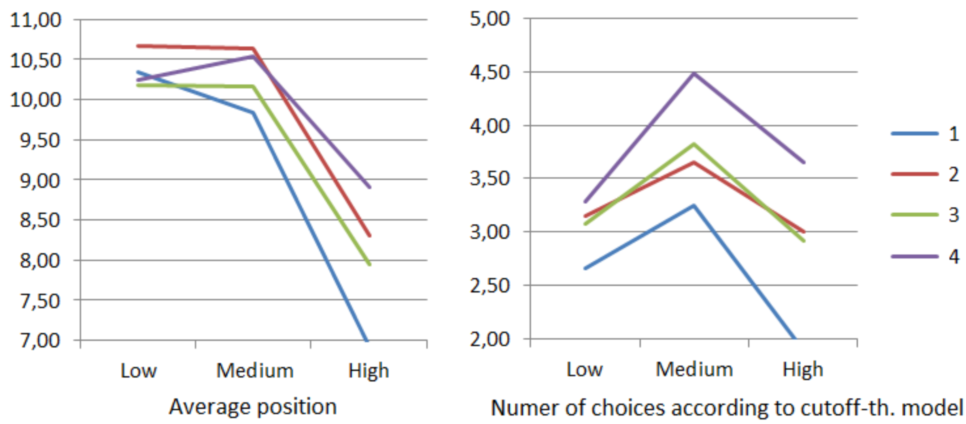


Figure 10: Different behaviour under sequences with low and high mean values

Finally we consider the difference between sequences with higher or lower values of the mean. Results shows participants generally behave differently, with worse performance under high sequences. This result is led by a dramatic reduction in the number of observations before choice.

Figure 10 shows the relation between this phenomenon and the role of intelligence. Average position is not different under low and medium sequences, but it is significantly lower if we consider high sequences. What is more interesting is how different clusters behave: the first cluster (low score) shows a reduction of over 3 positions, the fourth cluster (high score) only 1.5. Similar evidences appear from the analysis of the number of choices according to main models, and in particular cutoff-threshold model.

Participants of all the clusters perform better under medium sequences, with a significant decline under high values.

# 10 Areas of application

Optimal stopping problems have a wide range of applications.[29] Our variation on the traditional secretary problem introduces two aspects that make the choice situation more plausible. First, a choice close to the best one generates only a small reduction in the gain. Second, we can introduce learning effect in the precision of estimate of the unknown parameters of the distribution within each trial.

The normal distribution assumption is necessary to allow comparison with the optimal policy. Introducing heterogeneity about the distribution is a possible implementation of the problem.[30] Kahan et al. (1967) tested the effect of different distributions within the traditional framework and found no significant difference in the choice pattern used by subjects. We already tested the performance of variations on old heuristics. Results are comparable with the normal distribution case with slightly different optimal values for the position parameter according to the skewedness of the distribution.

As general examples of areas of applications we propose here three main areas: finance, staff selection and consumption.

About finance we do not think this model and scenario can be applied to professional broker choices. The number of observations available, the specialization in the recognition of trends and the role of sophisticated computational tools cover most of the elements of uncertainty of the decision maker. Furthermore, if we consider financial markets we know that final results come from the interaction between investors and expectations play a major role.

If we think about naïve investors the scenario could be more credible. Households looking for a loan have time constraints, uncertainty about economic dynamics (as interest rates can suddenly change) and perform a costly learning activity comparing

[29]More examples of applications in statistics and operational research can be found in Rothschild (1974) and Monahan (1982).

[30]We can design a task with no prior knowledge about the shape of the distribution. In this case we can consider normal, beta, power law, uniform and other distributions. The role of memory and intelligence in this case would not be related to the ability of estimating the unknown parameters. The most important aspect in this case is in fact the ability to recognize the shape of the distribution, and in particular the dispersion of values in the right tail of the distribution.

more alternatives. If they refuse an offer they could have no perfect recall option, as the opportunity passed or its features changed.
Similarly, small firms comparing alternative projects to implement or asking credit to different banks display most of the characteristics of the scenario we described.

The fictitious choice situation described by the secretary problem refers to staff selection. Is it possible to apply this framework in a "real" hiring problem? It is interesting to test whether human resources do apply some model for their selection processes, and at the moment the literature does not offer any study focused on this aspect.
Small and big firms face different aspects of the problem. In particular we think for small companies getting more information represents a bigger cost if there is no specific division with best practices and continuous activity, whereas big companies can face major issues related to time constraints.
To study the process we can also implement our model with new features, in particular using larger time windows and partial recall option.[31]

The literature about optimal stopping and the secretary problem presents many examples of consumption decision: buy plane tickets, visit stores, "look for partner", etc.
We think it is possible to apply this model to decisions involving uncertainty about product quality. The insights about the role of intelligence have clear marketing implications. Absent-minded consumers can be easily distracted by high discounts, choose the first product on the shelf at the supermarket and buy the first plane ticket they find.
Can this "exploration vs exploitation" framework get the major intuitions about the role of intelligence? According to the effects of learning, repeating the task with different values we could even propose a mild policy implication based on the implementation of different fictitious choice situations in order to improve the confidence for real decision and reduce potentially harmful distortive effects.

[31] If applicants look for other job opportunities, human resources have decreasing probability of successful recall if they submit late offers. Smith (1975) first considered this feature in an optimal stopping problem.

# 11 Conclusions

Researchers adopted the secretary framework to represent decision problems including hiring, replace a machine, buy a ticket, mating and stock trading.
In all these situations typical economic assumptions are not satisfied. Limited knowledge of alternatives or uncertainty about the effects, time constraints and computational limits are some of the causes of failure in the design of a pure maximization problem.

Exploration vs exploitation tradeoff is another versatile element and can be applied in many situations, from the search of a flat to rent to the research of a cloth to buy during the discount period.
March (1991) described this dilemma between self-excluding exploration of new possibilities (representing variation, search, experimentation, discovery, risk taking, creativity) and exploitation of old certainties (corresponding to choice, production, effort, refinement, efficiency, selection, implementation).

Variations on the secretary problem allow a wide range of applications. The decision problem we propose is able to describe a simplified version of a large range of common real-life situations that require role of learning through repeated observations and time constraints. We propose applications and suggest further research in the areas of domestic finance and consumption.
The models we analyze represent different degrees of complexity in the ability to use previous information to create correct expectations.
Modifications on traditional heuristics allow different levels of risk aversion. The inclusion of non-optimal candidates in the range of acceptable alternatives does not reduce the performance, and under some conditions the expected win increases.
Newly introduced threshold models (dynamic threshold, cutoff-th. and max-th.) include learning mechanisms and increase their precision in the estimates considering the observed applicants.
The dynamic threshold model represents the optimal policy but is very demanding from a computational point of view. The other rules obtain good results and require less effort in the estimation of the correct values of the unknown parameters.
In the matter of exploration and exploitation the class of threshold rules combines

them allowing coexistence. Each observation is adopted to refine estimates and evaluated as an applicant.
As we discussed, time plays a crucial role and is tightly connected to learning. In all the models we identified a basic and necessary level of exploration before decision.

We used behavioural data from an experiment with 120 participants to validate the models and analyze which of them fits human choices the best.
Versatile cutoff-threshold model accounts for systematic pattern of behaviour observed in human decision making to capture real choices. Participants often explicitly declared in the questionnaire that they waited a certain number of observations before evaluating the alternatives. Subjects with early stopping choices had no opportunity to learn from mistakes (no learning by doing, nor "regret" driving reinforcement learning).
Working memory and risk aversion display no impact on the performance or in the model adopted by participants. Intelligence is the major driver of performance and accuracy in the systematic adoption of models and ability to pick highest values.
This result is interesting because it captures the roles of intelligence and learning in a task characterized by a crucial role of timing.

We did not include explicit costs for information neither in the theoretical problem nor in the experiment. Information are costly because of time (endogenous cost of research) and other resources necessary to get them. In a more complex problem this feature need to be taken into account.
Behavioural experiments showed how costs can affect decision creating biases due to explicit sunk costs. Little effort was put in the design of simple tasks able to pair learning and sunk costs and remove this bias.

The secretary problem contains other limits in terms of assumptions.
For example low and high absolute values are considered in the same way even in the normal-distribution variation. Behavioural results show a significant different comparing sequences with different mean values. It is hence necessary to understand if a unique model is able to describe both the cases or specific characteristics of each background determine separate behavioural models.

This aspect has real implications, since buying a ticket for a plane or signing a mortgage have different absolute levels and temporal impacts.
Perfect memory and no recall option are strong assumptions as well. Both of them can relaxed introducing partial or declining features.
Changes in choice preferences of the decision maker or the applicants represent another possible issue. An agent whose offer had been previously declined could retire her proposal in a second moment, and this aspect is coherent with partial or no recall option. It is quite obvious if we consider the secretary problem in the formulation called *fiancée problem*. Koestler (1960) wrote that Johannes Kepler struggled with this problem when interviewing “applicants” for his second wife. After interviewing the fifth, he decided to propose to the fourth one, who turned him down because he had waited too long, or perhaps because he had other interviews.

Finally we want to stress the central role of threshold value. This concept corresponds to the acceptance regions defined by Muller (2000) but can also be applied in case of imperfect knowledge.
The concept of threshold is versatile and coherent with the “satisficing level” described by Simon (1955), and can easily modified to include elements of the prospect theory (Kahneman and Tversky, 1979).
Because of this, thresholds are a good tool in case we do not allow a proper maximization process and include uncertainty and learning.
One further implementation of this concept in presence of uncertainty allows the definition of three distinct regions of a probabilistic threshold model: one of acceptance and one of reject, and an intermediate region where the probability of acceptance increases as the value is closer to the actual threshold value.

# Part III

# Appendix

## 12 Appendix A - Subject instructions

### 12.1 Optimal stopping task

Imagine you are the director of a big company and you want to hire a secretary. Twenty applicants are available for an interview.
You meet the first candidate, interview her, and immediately after that you decide whether to hire her.
If you hire a candidate you immediately stop the interviews. You have already decided who you want to hire and do not need to meet the other applicants.
Otherwise, if you refuse current applicant, you do not have the opportunity to change your mind and choose her in a second moment.

The task you are going to perform at the computer does not requires you to hire a secretary, but the main rules are the same of the problem I have just described you. You have to choose one value among a sequence of numbers, instead of hiring a secretary among a group of applicants.
Every sequence is made by 20 numbers, expressed in euros, corresponding to you possible gains. After the welcome frame, the numbers will sequentially appear on the screen: value 1 out of 20 (for example 7,23 euros), value 2 out of 20 (e.g. 5,61 euros), and so on. Before every value you get the information about which position of the sequence you are currently inspecting.
If you do nothing but wait, numbers automatically flow on the screen. If you want to select the current value on the screen, just press the space bar.
In this way the value is saved on the computer and the sequence is interrupted. You go back to the initial frame and start a new sequence.
There are 30 sequences overall. At the end of the task you will draw three numbers from a box. Each number correspond to a sequence. Your final gain is equal to the average of the values you selected in these sequences, plus a flat amount of three euros as a show-up fee. To win as much as you can you have to select every time

the highest values of the sequences, but every sequence is different from the others.

Do you know normal or Gauss distribution? [meanwhile we draw a distribution as an example and explain the plot and the main properties]
The computer generated in this way the values you will see in the sequences. This type of distribution is characterized by two parameters: mean and variance. The mean [point the mean value in the graph] describes the average value you will find within the sequence. The variance represents how much values are spread around the mean. In general values are concentrated around the mean. If you go further and further from the mean the probability to find a value much higher or lower rapidly decades.
For example consider a normal distribution with a mean of 5 and a variance of 1. About 70% of the values are concentrated between 4 and 6, 95% of the values are concentrated between 3 and 7, and there is a small probability to find some value above 7 or below 3.

The values of mean and variance of the sequence are unknown, but you know that they are stable during each sequence. All of the 20 elements are generated by the same probability distribution, so mean and variance are the same.
After each choice you start a new sequence, corresponding to a new probability distribution. The values of the parameters here can be very different from the values of the previous one. For example you can find an higher or lower value of the mean (e.g. 12) and a greater or smaller variance (e.g. 3). In this case the values within the sequence are much higher and spread more around the mean.

Is a value like 10 euros high or low? You do not know if you have no clue about the parameters. In the first example ten is an high value, in the second example it is low.
In order to win as much as possible you have to choose high values in every sequence. All the sequences are different, and at the end you will draw three values. You will get one value among the sequences with the lower values, one among the average ones, and one among the highest, so it is crucial to do your best in all the cases.

## 12.2 Raven test

In this test you will fill every matrix with one of the eight possible tiles you find. In every matrix there are eight elements, the ninth one in the bottom right corner is missing. Elements within each matrix are linked by some rules, and the rules are different for each matrix. If you analyze the available elements it is possible to find which tile is the only one that correctly fills all the rules.

The first one is used as an example. The sequence of elements follows some regularities.
If you study the rows you can see the bold shape are a square, a circle and a diamond. All the elements are repeated in each row with a different order.
If you consider the columns you can count the number of dotted lines. There are one, two or three lines in each column.
Finally we can observe diagonals. In all the elements in the same diagonal there are dotted lines with the same slope.
Element number 5 is the only one that fills all the rules we described.

Every matrix is designed with different rules within its components. If you study the available elements you are able to find the correct missing element.
There are 29 matrices left. You have 20 minutes to solve as many matrices as you can. Wrong answers are not penalized. You have to correctly answer the highest number of cases.

## 12.3 Cognitive reflection test

Solve the three following problems.
1) A bat and a ball cost \$ 1.10 in total. The bat costs \$1.00 more than the ball. How much does the ball cost? ... cents
2) If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? ... minutes
3) In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? ... days

## 12.4 Holt & Laury risk aversion test

In this test you will choose among your favourite one between a couples of lotteries. The grid is made by ten couples of lotteries, one for each row. Let's look the first row. Imagine you won a ticket and you can decide whether use it to take part to lottery A or lottery B.

Every lottery is defined by two possible gains that can occur with given probabilities. For example in this case lottery A will give you 4 euros with a probability of 1/10 or 3.20 euros with a probability of 9/10. Lottery B, instead, will give you 7.70 euros with probability 1/10 or 0.20 euros with probability 9/10.

Choose the lottery you prefer between lottery A and lottery B and write an X on the corresponding cell, then move to the following couple. There are ten couples, in each of them you have to pick one of the two options. The values of the possible gains are the same in all the cases, probabilities will change for every case.

## 12.5 Working memory tests

Free recall working memory test

I will now read to you ten words. After I finish, you will write on the paper as many words as you can remember. The order is not important, so write down the words as they come to your mind.

Wechsler digit span test

I will now read to you some series of numbers. The series have increasing length. Every time I complete a series, you will repeat after me the same numbers in the same order I told you.

# 13 Appendix B - Pseudo-codes

## 13.1 Data generation and preparatory calculation

Sequence generation: creation of sequences of 20 elements (randomly picked from a standardized normal distribution, with mean 0 and variance 1).
Ordinal value: each element of the sequence is associated to its ordinal position within the sequence (absolute) and within the portion of the sequence up to its value (relative).

## 13.2 Cutoff rule

Input: 20-elements sequence, benchmark position $pos \in [1, 5]$, parameter $par \in [1, 20]$
Output: value chosen according to the model

- Observe first $par - 1$ elements without choosing them
- From the element $i = par$ on, choose the first value greater or equal of the highest value encountered so far (if $pos = 1$) or in general
- From the element $i = par$ on, choose the first value greater or equal of the number $pos$ highest value encountered so far
- If the last element of the sequence is reached without choosing, the 20th element is automatically selected The script saves, for each of the 5 variants and 20 parameters, the choice (position in the sequence, absolute ordinal value and value).

## 13.3 Candidate count rule

Input: 20-elements sequence, position $pos \in [1, 5]$, parameter $par \in [1, 20]$
Output: value chosen according to the model

- Consider only the "candidate" elements of the sequence, e.g. the values higher than the $pos$ highest value encountered so far
- Count how many candidates are encountered, and choose the $par$ candidate

- If the last element of the sequence is reached without choosing, the 20th element is automatically selected

The script saves, for each of the 5 variants and 20 parameters, the choice (position in the sequence, absolute ordinal value and value).

## 13.4 Successive non-candidate rule

Input: 20-elements sequence, position $pos \in [1, 5]$, parameter $par \in [1, 20]$
Output: value chosen according to the model

- Consider the "candidate" elements of the sequence, e.g. the values higher than the $pos$ highest value encountered so far
- Count how many successive non-candidates are encountered
- After $par$ successive non-candidates, choose the first following candidate
- If the last element of the sequence is reached without choosing, the 20th element is automatically selected

The script saves, for each of the 5 variants and 20 parameters, the choice (position in the sequence, absolute ordinal value and value).

## 13.5 Threshold model

**Model without uncertainty**
Part I - Identification of the thresholds
Input: characteristics of the distribution (mean and standard deviation of the normal distribution)
Output: sequence of thresholds

- Implicit: at round 20 every value will be accepted
- Round 19: every value above the mean (i.e. the expected value at this point) will be accepted; compute the expected value of the right tail (above the expected value) and the probability of randomly picking one number in the left tail (below the expected value). The threshold is given by (probability of left tail) $\cdot$ (expected value at round 20, i.e. the mean) + (probability of right tail) $\cdot$ (expected value right tail)

- Rounds 18-1: repeat the process considering the expected value of the following round and updating expected value of the right tail and probability of being in the left tail. Obtain remaining thresholds: (probability of left tail, considering the expected value at the following round) · (expected value at the following round) + (probability of right tail) · (expected value of the right tail)

Part II - Choice rule according to thresholds
Input: 20-elements sequence, Sequence of thresholds from previous part, Characteristics of the distribution
Output: value chosen according to the model

- Rounds 1-19: compare the value with the corresponding threshold value. Accept only if the element of the sequence is higher than the threshold value

- If the distribution is not standardized, multiply the threshold value by the sd and add the mean, then compare this value with the element of the sequence

- Round 20: accept any value

**Model with uncertainty**
Input: 20-elements sequence, sequence of thresholds from previous part, mean or standard deviation of the distribution in case of partial information
Output: value chosen according to the model

- Implicit: never select the first element of the sequence (not enough data to estimate variance)

- At each round (from 2 to 20) consider only the share of the sequence seen so far

- Estimate the parameters: mean and corrected standard deviation (or only one if the other is known)

- Combine the parameters with the basic threshold values:
  (threshold value · sd) + mean

- Compare this result with the corresponding element of the sequence. Choose the element if it is higher than the updated threshold value

- If the last element of the sequence is reached without choosing, the 20th element is automatically selected

The script saves the choice (position in the sequence and value).

## 13.6 Tests of robustness

Part I - Preliminary steps

To test the robustness of the values adopted in the "best threshold" model, we insert in the model an updating coefficient $m(k)$ close to 1 representing the uncertainty in the value of the parameter/s as function of the remaining observations. We want to verify if the threshold values we found are the best-performing ones, or if slightly lower/higher values have better results.

Assigning a value to the parameter $m(1)$, we compute $m(k) = m(1)^k$, $k \in [0, 19]$, where $k$ represents the number of observations left. The classic threshold value is multiplied by $m(k)$ and the element of the sequence is compared to this updated threshold value

Input: 20-elements sequence, sequence of thresholds from previous part, value assigned to $m(1)$ [we tested the interval 0.99-1.01 with 0.001 precision], mean or standard deviation of the distribution in case of partial information.

Output: value chosen according to the model

- At each round (from 2 to 20) consider only the share of the sequence seen so far

- Estimate the parameters: mean and corrected standard deviation (or only one if the other is known)

- Update threshold value multiplying by $m(k)$, where $k$ represents the number of elements left

- Combine the parameters with the updated threshold values:
  (threshold value $\cdot$ sd) $+$ mean

- Compare this result with the corresponding element of the sequence. Choose the element if it is higher than the updated threshold value

- If the last element of the sequence is reached without choosing, the 20th element is automatically selected

The script saves the choice (position in the sequence and value). Part II - Testing robustness
Input: the same as before, using n=100.000 sequences; all the values of $m(1)$ in the interval to test (interval 0.99-1.01 with 0.001 precision)
Output: performance of the model in correspondence of each of the values of the parameter

- Consider a value of $m(1)$ and compute all $m(k)$, $k \in [0, 19]$
- Run the $n$ sequences as shown above, updating the threshold values with $m(k)$
- Save the average performance of the model with the assigned value of $m(k)$
- Repeat for all the values of $m(1)$ in the interval

Results: we can find no significant difference comparing the performance of the model in the base version ($m(1) = 1$) and assigning to the parameter a value close to one. This result can be explained by the limited number of times where this slight change in the threshold causes a different choice.
Considering values of $m(1)$ in a broader interval we can recognize a significant decrease in the score obtained, and plotting the performance as a function of the value of $m(1)$ we can clearly recognize a smooth curve with the maximum in correspondence of $m(1) = 1$.

## 13.7 Cutoff-threshold rule

Threshold sequence generated using estimation of mean and variance.

- Observe first $n = 5$ elements without choosing them
- From the element in position $n + 1$ on, estimate mean and standard deviation and compare the element with a threshold value
- The threshold value is obtained as a sum of the estimated mean and the estimated standard deviation multiplied by a coefficient decreasing linearly through the sequence (from 1.5 at $i = 6$ to 0 at $i = 20$)

- Choose the first elements above the corresponding threshold value
- If the last element of the sequence is reached without choosing, the 20th element is automatically selected.

## 13.8 Maximum-threshold rule

Threshold sequence generated using estimation of mean and highest value encountered.

- Observe first $n = 5$ elements without choosing them
- From the element in position $n + 1$ on, estimate mean and consider the highest value encountered so far, and compare the element with a threshold value
- The threshold value is obtained as a sum of the estimated mean and the product between (the difference between the max and the estimated mean) and (a coefficient decreasing linearly through the sequence, from 1.5 at i=6 to at i=20)
- Choose the first elements above the corresponding threshold value
- If the last element of the sequence is reached without choosing, the 20th element is automatically selected

## 13.9 Generation of the sequences for the experiment

We want to check for a good randomization of the sequences to use for the experiment in order to distinguish the models adopted.
Input: Characteristics required for the sequences to select (random position for the best value, choices according to cutoff rule and threshold model)
Output: Groups of sequences to use for the experiment

- Generate $n = 100$ sequences (20 elements randomly picked from a standardized normal distribution)

- Create a matrix with the useful information of the sequences: position in the sequence of the element chosen according to cutoff rule (c) and threshold rule (t), and highest element of the sequence(b), and differences of these positions (c-b, t-b, c-t NOT in absolute value)

- Save in column (b) the position of the highest value of each sequence

- Save in column (c) the position of the value chosen according to cutoff rule ($pos = 1$, $par = 7$)

- Save in column (t) the position of the value chosen according to threshold model

- Compute the differences: c-b, t-b, c-t

- Create a matrix "filter", where each row represents a sequence, each column a condition (distance between c and t, position of the best), and the value in the cell is 0 (false) or 1 (true)

- Fill the matrix with the conditions (c-t: $>2/<-2/=0$, position of best)

- Create a matrix "type" with the combinations of conditions that we want the final sequences to fulfill

- Combine "type" and "filter", and save in "final" the sequences

# 14 Bibliography


Bearden, J.N. (2006). *A new secretary problem with rank-based selection and cardinal payoffs*, Journal of Mathematical Psychology, Vol. 50, pp. 85-59

Bearden, J.N, Murphy, R.O. and Rapoport, A. (2005). *A multi-attribute extension of the secretary problem: Theory and experiments*, Journal of Mathematical Psychology, Vol. 49, N. 5, pp. 410-422

Bellman, R. (1957). *A Markovian decision process*, Rand Corp Santa Monica CA, p 1066

Brezzi, M. and Lai, T.L. (2002). *Optimal learning and experimentation in bandit problems*, Journal of Economic Dynamics and Control, Vol. 27, N.1, pp. 87-108

Chow, Y.S. et al. (1964). *Optimal selection based on relative rank (the "secretary problem")*, Israel Journal of Mathematics, Vol. 2, N. 2, pp. 81-90

Chun, Y.H. (2000). *Sequential search and selection problem under uncertainty*, Decision Sciences, Vol. 31, N. 3, pp. 627-648

Costa, V.D. and Averbeck, B.B. (2015). *Frontal-parietal and limbic-striatal activity underlies information sampling in the best choice problem*, Cerebral Cortex, Vol. 25, N.4, pp. 972-982

Cowan, R. and Zabczyk, J. (1978). *An optimal selection problem associated with the Poisson process*, Theory Probab. Appl., Vol. 23, pp. 584-592

De Groot, M. (1970). *Optimal statistical decisions*, McGraw-Hill

Ding, J. (2007). *A multiple secretary problem with switch costs*, Diss. Massachusetts Institute of Technology

Ferguson, T. (1967). *Mathematical statistics: a decision theoretic approach*, Academic Press, New York

Freeman, P.R. (1983). *The secretary problem and its extensions: a review*, International Statistical Review/Revue Internationale de Statistique, pp. 189-206

Furl, N. and Averbeck, B.B. (2011). *Parietal cortex and insula relate to evidence seeking relevant to reward-related decisions*, The Journal of Neuroscience, Vol. 31, N. 48, pp. 17572-17582

Gelman A. et al. (2004). *Bayesian data analysis*, New York: Chapman and Hall/CRC

Gianini-Pettit, J. (1979). *Optimal selection based on relative ranks with a random number of individuals*, Advances in Applied probability, Vol. 11, pp. 720–736

Gigerenzer, G. and Todd, P.M. (1999). *Simple heuristics that make us smart*, Oxford University Press

Gilbert, J.P. and Mosteller, F. (1966). *Recognizing the maximum of a sequence*, J. Amer. Statist. Ass., Vol. 61, pp. 35-73

Hommes, C. (2013). *Behavioral rationality and heterogeneous expectations in complex economic systems*, Cambridge University Press

Kahan, J.P., Rapoport, A. and Jones, L.V. (1967). *Decision making in a sequential search task*, Perception & Psychophysics, Vol. 2, pp. 374-376

Kahneman, D. and Tversky, A. (1979). *Prospect theory: an analysis of decision under risk*, Econometrica: Journal of the Econometric Society, pp. 263-291

Kogut, C.A. (1990). *Consumer search behavior and sunk costs*, Journal of Economic Behavior & Organization, Vol. 14, N. 3, pp. 381-392

Lee M.D. and Paradowski, M.J. (2003). *Group decision-making on an optimal stopping problem*, The Journal of Problem Solving, Vol. 1, N. 2, pp. 53-73

Lindley, D.V. (1961). *Dynamic-programming and decision-theory*, Applied Statistics, Vol. 10, N.39–51

March, J.G. (1991). *Exploration and exploitation in organizational learning*, Organization science, Vol. 2, N.1, pp. 71-87

Monahan, G.E. (1982). *A survey of partially observable Markov decision processes*, Management Science, Vol. 28, No. 1, pp. 1-16

Muller, A. (2000). *Expected utility maximization of optimal stopping problems*, European Journal of Operational Research, Vol. 122, pp. 101-114

Paulus, M.P. et al. (2001). *Prefrontal, parietal, and temporal cortex networks underlie decision-making in the presence of uncertainty*, Neuroimage, Vol. 13, N. 1, pp. 91-100

Petrucelli, J.D. (1982). *Full-information best-choice problems with recall of observations and uncertainty of selection depending on the observation*, Advances in Applied Probability, pp. 340-358

Pressman, E. and Sonin, I. (1972). *The best choice problem for a random number of objects*, Theory Prob. Appl., Vol. 17, pp. 657-668

Rapoport, A. and Tversky, A. (1970). *Choice behavior in an optimal stopping task*, Organizational Behavior and Human Performance, Vol. 5, N.2, pp. 105-120

Rasmussen, W.T. and Robbins, H. (1975). *The candidate problem with unknown population size*, Journal of Applied probability, pp. 692-701

Rethmann, J. and Wanke, E. (2001). *Stack-up algorithms for palletizing at delivery industry*, European Journal of Operational Research, Vol. 128, N.1, pp. 74-97

Rothschild, M. (1974). *Searching for the lowest price when the distribution of prices is unknown*, The Journal of Political Economy, pp. 689-711

Sakaguchi, M. (1979). *A generalized secretary problem with uncertain employment*, Math. Japon., Vol. 23, pp. 647–653

Seale, D.A. and Rapoport, A. (1997). *Sequential decision making with relative ranks*, Organizational Behavior and Human Decision Processes, Vol. 69, No. 3, pp. 221-236

Seale, D.A. and Rapoport, A. (2000). *Optional stopping behavior with relative ranks: the secretary problem with unknown population size* , Journal of Behavioral Decision Making, Vol. 2, pp. 280-300

Sethi, S. and Sorger, G. (1991). *A theory of rolling horizon decision making*, Annals of Operations Research, Vol. 29, N. 1, pp. 387-415

Shu, S.B. (2008). *Future-biased search: the quest for the ideal*, Journal of Behavioral Decision Making, Vol. 21, N. 4, pp. 352-377

Smith, M.H. (1975). *A secretary problem with uncertain employment*, Journal of applied probability, pp. 620-624

Sonnemans, J. (1998). *Strategies of search*, Journal of economic behavior & organization, Vol. 35, N.3, pp. 309-332

Stein, W.E., Seale, D.A. and Rapoport, A. (2003). *Analysis of heuristic solutions to the best choice problem*, European Journal of Operational Research, Vol. 151, pp. 140-152

Stewart, T.J. (1981). *The secretary problem with an unknown number of options*, Operations Research, Vol. 29, N. 1, pp. 130-145

Stigler, G.J. (1961). *The economics of information*, The journal of political economy, pp. 213-225

Thaler, R. (1980). *Toward a positive theory of consumer choice*, Journal of Economic Behavior & Organization, Vol. 1, N. 1, pp. 39-60

Tsitsiklis, J.N. and Van Roy, B. (1999). *Optimal stopping of Markov processes: Hilbert space theory, approximation algorithms, and an application to pricing high-dimensional financial derivatives*, Automatic Control, IEEE Transactions, Vol. 44, N. 10, pp. 1840-1851

Van Moerbeke, P. (1976). *On optimal stopping and free boundary problems*, Archive for Rational Mechanics and Analysis, Vol. 60, N. 2, pp. 101-148