ECOLE POLYTECHNIQUE

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

# THE INFLUENCE OF INFORMATION AVAILABILITY ON THE CHOICE OF DESTINATION

F. COMBES
A. DE PALMA

August 31, 2010

Cahier n° 2010-22

DEPARTEMENT D'ECONOMIE
Route de Saclay
91128 PALAISEAU CEDEX
(33) 1 69333033
http://www.economie.polytechnique.edu/
mailto:chantal.poujouly@polytechnique.edu

# The influence of information availability on the choice of destination

F. Combes[*]

A. de Palma[†]

August 31, 2010

**Abstract**

We set a framework where an individual has to choose one among a set of spatially distributed activities. The individual knows the price of each activity, as well as the distance to reach it. She has either full or zero information about each activity's quality. Qualities are modeled by i.i.d. random variables. Under the full information regime, the individual knows the realizations of the qualities; while under the no information regime, she only knows the distribution of the qualities. In that case, she can decide either ex ante, or en route, how many activities to patronize.

[*]Université Paris Est, LVMT, UMR T9403 INRETS ENPC UMLV; Marne-La-Vallée, France; corresponding author.

[†]Ecole Normale Supérieure de Cachan, Cachan; CES; and Ecole Polytechnique, Palaiseau, France.

We analyze the impact of information availability on the choice process, on the distance the individual covers, and on the individual's expected utility. In this framework, more information yields longer distance traveled, but also higher utility. We compute the individual's willingness to pay for information. Finally, we show that providing information may decrease the individual's benefit when congestion arises.

**keywords**: *travel demand, search, logit, information regimes, value of information, differentiation*

# 1 Introduction

Information theory, a well known topic in economics, has influenced during the last decades various areas such as industrial organization, contract theory, finance or decision theory (see, e.g. Gollier et al., 2005). In the transport literature, many interesting articles have been written to study the impact of information on drivers' behavior, notably with respect to the choice of route (see e.g. de Palma and Picard, 2006). A less explored issue is that of the influence of information availability on higher level travel-related decisions, such as the choice of destination, particularly when it is possible for the individuals to use information strategically (e.g. when they can decide to acquire information or not before making a decision). We will explore in this paper how individuals can optimally use costly information.

Transportation involves persons and goods which often (but by far not always) know where to go. A satisfactory trip is one trip which involves a good match between the origin and the destination (match between goods, persons or services). Two persons can have a more or less good match; one person can buy a more or less satisfactory good or acquire a good or a bad service, or proceed to a more or less satisfactory activity. This is essentially a matching problem, which was addressed decades ago in the transport literature (it is known as the multi-commodity flow problem; see e.g. Barnhart and Sheffi, 1993, who present a number of applications of this problem in transport).

In this paper, we explore a parallel approach to associate origins and destinations. Rather than considering a centralized approach which optimally twins origins and destinations, we consider a decentralized process in which individuals look for their best match, in a context of imperfect information. Basically, we consider situations where individuals must choose one among many destinations. We also assume that individuals have limited information on the destinations. Information on a destination is acquired through its exploration. Therefore, the individuals must devote some time and resources to identify a satisfactory destination.

Such situations include, for example, the case of an individual willing to perform some task (buy some shoes, go to the restaurant or find a job, for example) at some discrete locations. The individual knows where the shops are, what the charged prices are, but does not know the match of her preferences with the goods available at each shop. She needs to be physically present at a location to know the quality of the match. When matches are idiosyncratic, such information cannot be transmitted and remains private.

The match between the individual and any alternative is represented by a random variable. The match values of the activities are assumed independent, identically distributed, and their common distribution is assumed known to the individual. If the match value is not observed by the individual, then the only information the individual has on it is its distribution. If the match value is observed by the individual, then the individual knows the realization of the corresponding random variable. The analysis is restricted to the case of rational individuals, who, in particular, acquire information optimally, and use optimally the information they have acquired. Besides, it is also assumed the individuals know the distribution of the match values. Finally, the spatial layout of the possible destinations of the individual is chosen deliberately simple: they are assumed regularly spaced along a linear road.

Information on the various alternatives available can be acquired by the individual or not, depending on the information suppliers. When information can only be acquired via physical presence, two cases are distinguished. In the first case, the individual needs not to commit, and after each visit, can decide to continue the search or not, on the basis of the previously acquired information. If she stops, her benefit will be the largest benefit of all locations explored so far (search with recall). In the frame of this paper, it is relatively simple to derive the decision protocol that optimizes the individual's expected net benefit (gross benefit, i.e. match value minus price) minus transportation cost.

In the second case, the individual has to commit herself to a given time period to find a suitable alternative. She decides ex-ante how much time she wishes to devote to her search, and then cannot alter her decision. Given the time period chosen by the individual, and thus the number of shops visited (defined ex ante), the expected benefit of the individual is based on the distribution of the match values. In such case, the ex ante expected net benefit of visiting a given, fixed number of activities is conveniently calculated using the accessibility measure (or log-sum as defined by Ben-Akiva and Lerman, 1985).

As a benchmark, we also consider a situation where the individual knows the realizations of all match values ex ante. In such case, the individual will select the good or service which will maximize her net benefit. On average, this situation yields the individual a higher net benefit than the two previous ones.

Finally, we consider the three regimes discussed above (two with limited information and one with full information), in the presence of congestion. In such case, the discussion is far more involved.

This paper is the first to model the use of optimal search theory to compute the total distance travelled under different information regimes. A search model analogous to the one considered here was studied by Robbins (1970) in an a-spatial context. In our setting, users know the parameter values of the distribution of the match values. If it is not the case, the situation is far more complex. For a discussion of this problem (involving no space), we refer the reader to de Lara, Chancelier and de Palma (2007).

One major result of our analysis is that the expected distance travelled in the no information regime can be lower than in the full information regime. In such a way, the Internet would increase, and not decrease the total distance travelled. This conclusion is in direct contradiction with the relatively widespread intuition according to which more information implies less travel demand (as the share of travel demand which would stem from the need of the individual to acquire information would be, in this case, superfluous). Our analysis also allows for the calculation of the willingness to pay for information.

In the next section, we present the set of hypotheses on which the main model is based. The three regimes (no information with commitment, no information without commitment and full information) are considered in Section 3. Then, they are compared in Section 4. Finally, simple examples involving search and congestion are provided in Section 5. Concluding comments are presented in Section 6.

# 2 Framework

In the frame of this work, we consider an individual facing a set of activities, indexed by $i \in \mathbb{N}$, located on an infinite line. Moving on this line is costly to the individual. The transportation cost is proportional to the distance covered, up to a positive multiplicative constant $\alpha$ (half of the transportation cost per unit of distance travelled). Each activity is assumed to be at distance $\delta$ from its neighbors. Activity 0 is at distance 0 from the residence of the individual, so that activity $i$ is at distance $\delta i$ from the individual (see Figure 7).

[FIGURE 1]

The individual must choose one of these activities. Each activity provides the individual with a certain degree of satisfaction, referred to as the quality of the activity. The quality of activity $i$ is measured as an idiosyncratic monetary fit between the individual and the activity. The value of the fit is denoted by $\varepsilon_i$. All the activities have distinct qualities: by assumption, $\{\varepsilon_i\}$ are i.i.d. Gumbel random variables, with CDF $F$, density $f$ and scale parameter $\mu > 0$: $F(x) = \exp(-\exp(-x/\mu - \gamma))$, and $f(x) = 1/\mu. \exp(-x/\mu - \gamma)\, F(x)$, where $\gamma$ is the Euler constant, ($\gamma \approx 0.577$). As a consequence, the benefit $u_i$ from choosing activity $i$ consists of two terms, up to a constant additive term set to zero without loss of generality. The first one concerns the transport disutility, and the second the activity's quality: $u_i = -\alpha\delta i + \varepsilon_i$.

We examine the influence of information availability concerning the qualities of the activities $\{\varepsilon_i\}$ on the way the individual chooses her destination. We consider two distinct regimes: the "full information" regime and the "no information" regime. Under the full information regime, the individual knows the precise values of the qualities of the activities (or, using a statistical vocabulary, their realizations) before choosing her destination. The full information regime is labeled by $f$. Under the no information regime, she does not know the realizations of the qualities, and must choose her destination on the basis of their distributions.

Two types of behavior are considered under the no information regime. In the first case, the individual decides *ex ante* the number of activities she will visit. While parsing them, she cannot interrupt her search, nor pursue it. For example, this is the type of behavior of an individual deciding the time she is ready to dedicate to searching a satisfying activity, while setting her whole schedule for a given day. This type of behavior is referred to as "with commitment", and labeled by 0. In the second case, the individual parses the activities sequentially, and can decide at any time to opt for any of the already visited ones. This type of behavior is referred to as "no commitment", and labeled by $d$.

# 3 Analysis

In this section, the behavior of the individual is examined sequentially under the no information regime with commitment, the no information regime without commitment, and the full information regime. In each case, the expected value and standard deviation of the distance the individual covers is derived, as well as the individual's expected utility.

## 3.1 No information regime, with commitment

The no information regime with commitment is defined as follows:

**Definition 1** *Under the no information regime with commitment, the individual cannot know the quality of an activity without visiting this activity. Furthermore, the individual must decide* ex ante *how many activities she will visit.*

Under this regime, the individual has to choose the set of activities she will visit before starting her tour. Once she has visited these activities, she selects the one with the best quality $\varepsilon_i$, at no additional cost (on her way back).

Denote by $n$ the number of activities the individual decides to patronize *ex ante*. The distance she covers is then $(n-1)\delta$. Denote by $d$ this distance. Denote by $\mathcal{U}_0$ her expected utility. Consider $\mathcal{U}_0$ as a function of $d$, with $d/\delta \in \mathbb{N}$. Under the no information regime with commitment, the individual cannot observe the qualities of the activities, and must decide *ex ante* the number of activities she will patronize. In other words, she must choose $d$ so as to maximize $\mathcal{U}_0$.

$\mathcal{U}_0$ is the expected value of the maximum of the qualities of the patronized activities, minus the transport cost:

$$\mathcal{U}_0(d) = \mathbb{E}\left(\max\left(\varepsilon_0; \ldots; \varepsilon_{d/\delta}\right)\right) - \alpha d.$$

Given that the $\varepsilon_i$ are i.i.d. Gumbel random variables, the expected value of their maximum is given by the log-sum formula (Ben-Akiva and Lerman, 1985):

$$\mathcal{U}_0(d) = \mu \ln(d/\delta + 1) - \alpha d,$$

which is a concave function of $d$. As a consequence, $d$ is optimal if and only if:

$$\mathcal{U}(d-\delta) \leq \mathcal{U}(d) \quad \text{and} \quad \mathcal{U}(d) \geq \mathcal{U}(d+\delta),$$

which is equivalent to:

$$(\exp(\delta\alpha/\mu) - 1)^{-1} - 1 \leq d/\delta \leq (\exp(\delta\alpha/\mu) - 1)^{-1}. \tag{1}$$

This leads to the following first result:

**Proposition 1** *Under the no information regime with commitment, the expected distance the individual covers is fixed and given by:*

$$d_0 = \delta \left\lfloor (\exp(\delta\alpha/\mu) - 1)^{-1} \right\rfloor, \tag{2}$$

*and her expected utility is:*

$$\mathcal{U}_0 = \mu \ln(d_0/\delta + 1) - \alpha d_0, \tag{3}$$

(where $x \mapsto \lfloor x \rfloor$ is the floor function and denotes the integer part of $x$.)

Given that the individual must choose the number of activities she will patronize on the basis of the distributions of the $\varepsilon_i$, but not on the basis of their realizations, she cannot inflect her decision using the information she acquires during the course of her search. Therefore, the standard deviation of the distance she covers is zero. From that perspective, the two other cases differ.

The asymptotic behavior of these functions when $\delta$ gets close to zero is given by (where $o_{0^+}(f)$ denotes a function $g(x)$ such that $\lim g(x)/f(x) = 0$ when $x$ gets close to 0, $x > 0$):

**Corollary 1** *Under the no information regime with commitment, for small values of $\delta$, the expected distance the individual covers behaves asymptotically as:*

$$d_0 = \mu/\alpha + o_{0^+}(1), \tag{4}$$

*and her expected utility behaves asymptotically as:*

$$\mathcal{U}_0 = -\mu \ln(\delta) + \mu \ln(\mu/\alpha) - \mu + o_{0^+}(1). \tag{5}$$

*Proof:* See Appendix.

Note that $d_0$ is not differentiable with respect to $\delta$ when $\delta$ is close to zero. Indeed, the individual always chooses an integer number of activities to visit, and this number changes when the parameters of the model change. As a consequence, it cannot change continuously. $d_0$, which is equal to this number times $\delta$, is therefore discontinuous every time there is a change in the number of activities the individual decides to visit. Therefore, it is not possible to approximate $d_0$ at the first order in $\delta$ (Equation (4) is an approximation at order zero). Note also that $\mathcal{U}_0$ increases indefinitely as $\delta$ decreases. In the frame of this model, an increase in the density of activities does not result in the expected distance covered by the individual decreasing toward zero, but toward $\mu/\alpha$, so that the transport cost tends toward $\mu$. The number of activities she visits increases indefinitely, and so does her expected utility as $\delta$ tends toward 0.

This extreme result is a mathematical consequence of the fact that the support of the double exponential distribution does not have a finite upper bound. Assume instead that $\forall i \in \mathbb{N}$, $\varepsilon_i \leq a$ with $a$ a constant, then $\forall \delta > 0$, $\mathcal{U}_0$ is finite since, necessarily, $\mathcal{U}_0 \leq a$.

## 3.2 No information regime, without commitment

The no information regime without commitment is defined as follows:

**Definition 2** *Under the no information regime without commitment, the individual cannot know the quality of an activity without visiting this activity. After each visits, she decides whether or not to continue her exploration or to opt for one of the already visited activities.*

Under this regime, the individual must visit an activity in order to observe its quality. However, she does not have to decide *ex ante* the number of activities she will patronize. She visits activities sequentially. We assume that at any time, she can opt for one of the already observed activities at no additional cost (although, as it will appear later, this assumption is not necessary).

First, the optimal strategy of the individual is derived. Second, the expected value and standard deviation of the distance covered are derived, as well as the expected utility of the individual.

### 3.2.1 Optimal strategy

Consider first a simple situation. Assume the individual has visited activity 0. Denote $e_0$ the realization of $\varepsilon_0$, the quality of activity 0. Based on $e_0$, the individual decides whether to visit activity 1 or not. The expected incremental utility of the decision to go to activity 1 is:

$$\Delta u(e_0) = \mathbb{E}\left((\varepsilon_1 - e_0)^+ - \delta\alpha\right), \tag{6}$$

Note that this is the expression used in finance for an option value. Equation (6) may be rewritten as:

$$\Delta u(e_0) = \int_{e_0}^{+\infty} (e - e_0) f(e) \mathrm{d}e - \delta\alpha, \tag{7}$$

which has a unique solution:

**Lemma 1** *There exists a unique $e_c$ verifying:*

$$\Delta u(e_c) = 0. \tag{8}$$

Indeed, provided that both $\int_{\mathbb{R}} f(e)\mathrm{d}e$ and $\int_{\mathbb{R}} ef(e)\mathrm{d}e$ are finite and that $e_0 \mapsto (e - e_0)f(e)$ is continuous, $\Delta u$ exists and is continuous. It is clearly decreasing in $e_0$, from $+\infty$ to $-\delta\alpha$. Note that this result is not specific to the double exponential distribution. As expected intuitively, $e_c$ is a decreasing function of $\alpha$, $\delta$ and an increasing function of $\mu$. There is no close formula of $\Delta u$ for the double exponential distribution.

By definition of $e_c$, and from Equation (7), we have:

$$\int_{e_c}^{+\infty} ef(e)\mathrm{d}e = \delta\alpha + e_c(1 - F(e_c)) \tag{9}$$

Let us now come back to the original framework. The individual may patronize any of the activities on the road, in any order, and she may stop whenever she wants. However, given the symmetry of the problem, whatever the activities the individual has already visited, the remaining ones only differ by their distance to the current location of the individual. It is therefore always optimal for the individual to reach the nearest unvisited one, *i.e.* to visit the activities in the order of their indices. The expected benefit at step $n$ is thus:

$$\mathbb{E}\left(\max(\varepsilon_0; \ldots; \varepsilon_n)\right) - n\delta\alpha. \tag{10}$$

In this situation:

**Lemma 2** *The optimal strategy is to stop at the first $n$ such that $e_n \geq e_c$, where $e_c$ is implicitly given by Equation* (9).

*Proof:*

The formal proof is rather technical, due to the large number of strategies to consider and evaluate. It can be found in Robbins (1970). Its main arguments are outlined in the Appendix. The result holds for any distribution density $f$ such that $\int_{\mathbb{R}} ef(e)\mathrm{d}e$ exists and is finite, and has an equivalent for the case where $\int_{\mathbb{R}} ef(e)\mathrm{d}e = +\infty$.

The following reasoning can help understanding the intuition behind this result. Indeed, given the symmetry of the problem, it is reasonable to assume the optimal strategy for the individual is to stop as soon as she has observed a quality $e_i$ which is larger than a given limit $\bar{e}$. Let us find out the optimal $\bar{e}$. If the individual follows this strategy, her expected utility is:

$$\mathcal{U}_d(\bar{e}) = \int_{\bar{e}}^{+\infty} ef(e)\mathrm{d}e + \int_{-\infty}^{\bar{e}} \int_{\bar{e}}^{+\infty} (e - \alpha\delta) f(e)\mathrm{d}e f(e_0)\mathrm{d}e_0 + \ldots,$$

which is equivalent to:

$$\mathcal{U}_d(\bar{e}) = \int_{\bar{e}}^{+\infty} ef(e)\mathrm{d}e \cdot \sum_{i=0}^{+\infty} F(\bar{e})^i - \delta\alpha(1 - F(\bar{e})) \sum_{i=0}^{+\infty} iF(\bar{e})^i.$$

Given that:

$$\sum_{i=0}^{+\infty} F(\bar{e})^i = \frac{1}{1-F(\bar{e})},$$

and that:

$$(1-F(\bar{e}))\sum_{i=0}^{+\infty} iF(\bar{e})^i = \frac{F(\bar{e})}{1-F(\bar{e})},$$

the expected utility of the individual is given by:

$$\mathcal{U}_d(\bar{e}) = \frac{1}{1-F(\bar{e})}\left(\int_{\bar{e}}^{+\infty} ef(e)\mathrm{d}e - \delta\alpha F(\bar{e})\right). \tag{11}$$

Its derivative with respect to $\bar{e}$ is:

$$\frac{\partial \mathcal{U}_d}{\partial \bar{e}} = \frac{f(\bar{e})}{1-F(\bar{e})}\left[\frac{1}{1-F(\bar{e})}\left(\int_{\bar{e}}^{+\infty} ef(e)\mathrm{d}e - \delta\alpha F(\bar{e})\right) - \bar{e} - \delta\alpha\right].$$

The optimal $\bar{e}$ then necessarily verifies the following first order condition:

$$\int_{\bar{e}}^{+\infty} ef(e)\mathrm{d}e - \delta\alpha = (1-F(\bar{e}))\bar{e}, \tag{12}$$

which is equivalent to Equation (9). As a consequence, it is also equivalent to $\Delta u(\bar{e}) = 0$, where $\Delta u$ is given by Equation (6). Q.E.D.

The strategy of a rational individual proceeding to a sequential search of a satisfying activity is thus to choose the first activity with a quality greater than $e_c$. This critical quality is such that if the individual observes an activity with a quality larger or equal to $e_c$, the expected benefit from visiting the next activity is smaller or equal to the transport cost.

### 3.2.2 Analysis

From Equations (9) and (11), when the individual applies the optimal strategy given by Lemma 2, her expected utility is:

$$\mathcal{U}_d(e_c) = e_c + \delta\alpha.$$

The probability that the distance the individual covers is equal to $\delta n$ is $(1 - F(e_c))F(e_c)^n$. In other words, the distribution of the distance the individual covers is geometric, up to the multiplicative constant $1/\delta$. The calculation of its expected value $d_d$ and of its standard deviation $\sigma_d$ are therefore straightforward. The results of this section sum up as follows:

**Proposition 2** *Under the no information regime, without commitment, if the individual adopts an optimal strategy, the expected distance she covers is:*

$$d_d = \delta F(e_c)/(1 - F(e_c)). \tag{13}$$

*The standard deviation of this distance is:*

$$\sigma_d = \delta\sqrt{F(e_c)}/(1 - F(e_c)). \tag{14}$$

*The expected utility of the individual is:*

$$\mathcal{U}_d = e_c + \delta\alpha. \tag{15}$$

Note that the expected utility is equal to the expected value of the chosen activity, minus the expected transport cost. Indeed, consistently with Equations (9) and (13), it can be written:

$$\mathcal{U}_d = \mathbb{E}(\varepsilon|\varepsilon \geq e_c) - \alpha d_d.$$

Besides, contrary to the previous case, the standard deviation $\sigma_d$ of the distance the individual covers is positive. Indeed, the decision of the individual depends on the realizations of the qualities of the visited activities, which are random.

Finally, as $\delta$ tends towards zero, we have the following results:

**Corollary 2** *Under the no information regime without commitment, the expected distance the individual covers behaves asymptotically as:*

$$d_d = \mu/\alpha - 3\delta/4 + o_{0^+}(\delta). \tag{16}$$

*The standard deviation of this distance behaves asymptotically as follows:*

$$\sigma_d = \mu/\alpha - \delta/4 + o_{0^+}(\delta). \tag{17}$$

*The expected utility of the individual behaves asymptotically as:*

$$\mathcal{U}_d = -\mu \ln(\delta) + \mu \ln(\mu/\alpha) - \mu\gamma + o_{0^+}(1). \tag{18}$$

### 3.3 Full information regime

The full information regime is defined as follows:

**Definition 3** *Under the full information regime, the individual knows the qualities of all the activities.*

Under this regime, the utility for the individual of choosing activity $i$ is the activity's quality minus the transport cost:

$$u_i = \varepsilon_i - \alpha\delta i.$$

The individual chooses the activity endowed with the highest $u_i$. Given that the $\{\varepsilon_i\}$ are i.i.d. and that their distribution is double exponential, the probability that the individual patronizes activity $i$ is given by the multinomial logit model:

$$\mathbb{P}(i) = \frac{\exp\left(-\delta\alpha i/\mu\right)}{\sum_{i=0}^{+\infty} \exp\left(-\delta\alpha i/\mu\right)},$$

and the expected utility of the individual is $\mathcal{U}_f = \mu \ln\left(\sum_{i=0}^{+\infty} \exp(-\delta\alpha i/\mu)\right)$.

The distribution of the distance the individual covers is geometric. Simple calculations then yield the following results:

**Proposition 3** *Under the no information regime with commitment, the expected distance the individual covers is:*

$$d_f = \delta/(\exp\left(\delta\alpha/\mu\right) - 1). \tag{19}$$

*The standard deviation of the distance is:*

$$\sigma_f = \delta \exp\left(-\delta\alpha/2\mu\right) / (1 - \exp\left(-\delta\alpha/\mu\right)).$$

*The expected utility of the individual is:*

$$\mathcal{U}_f = \delta\alpha - \mu \ln(\delta) + \mu \ln(d_f). \tag{20}$$

Here again, the distance the individual covers is random, as the choice of the destination is directly linked to the qualities, which are also random.

The asymptotic behavior of these functions when $\delta$ gets close to 0 is given by:

**Corollary 3** *Under the full information regime, the expected utility of the individual behaves asymptotically as follows:*

$$\mathcal{U}_f = -\mu \ln(\delta) + \mu \ln\left(\mu/\alpha\right) + o_{0^+}(1). \tag{21}$$

*The expected distance she covers behaves asymptotically as follows:*

$$d_f = \mu/\alpha - \delta/2 + 0_{0^+}(\delta). \tag{22}$$

*The standard deviation of this distance behaves asymptotically as follows:*

$$\sigma_f = \mu/\alpha - \delta + 0_{0^+}(\delta).$$

*Proof:* See Appendix.

The relationship between information availability and travel demand, and the benefit for the individual of information are examined in the next section.

# 4 Comparing regimes

As seen in the previous section, the regimes yield distinct utilities and travel demands. They are compared in Subsection 4.1.

Given that distinct regimes yield distinct regimes, the individual may be willing to pay to change regime and acquire information or freedom to choose. This willingness to pay is defined and measured in Subsection 4.2.

## 4.1 Comparing the expected utilities

It is possible to rank the expected utilities of the individual under each regime, given by Equations (3), (15) and (20). The results are collected in:

**Proposition 4** *The expected utility under the full information regime is higher than under the no information regime without commitment, which itself is higher than under the no information regime with commitment.*

$$\mathcal{U}_0 \leq \mathcal{U}_d \leq \mathcal{U}_f.$$

*Proof:*

Let:

$$\mathcal{U}_f^{(n)} = \int_{\mathbb{R}^{n+1}} \max(e_0; e_1 - \delta\alpha; \ldots; e_n - n\delta\alpha) \prod_{i=1}^{n} f(e_i)\mathrm{d}e_i,$$

and:

$$\mathcal{U}_d^{(n)} = \sum_{k=0}^{n-1} \int_{]-\infty;e_c[^k \times [e_c;+\infty[} \left( (\max(e_0;\ldots;e_k) - k\delta\alpha) \prod_{i=1}^{k} f(e_i) \mathrm{d}e_i \right)$$

$$+ \int_{]-\infty;e_c[^n \times \mathbb{R}} \left( (\max(e_0;\ldots;e_n) - n\delta\alpha) \prod_{i=1}^{n} f(e_i) \mathrm{d}e_i \right),$$

Considering that $\bigsqcup_{i=0}^{n} (]-\infty;e_c[^n \times [e_c;+\infty[) = \mathbb{R}^{n+1}$, and that for all $n$, for all $\{e_i\}_{i\in[0;n]} \in \mathbb{R}^{n+1}$:

$$\begin{cases} \max(e_0;\ldots;e_n) - n\alpha & \leq & \max(e_0; e_1 - \alpha; \ldots; e_n - n\alpha) \\ \ldots & & \ldots \\ e_0 & \leq & \max(e_0; e_1 - \alpha; \ldots; e_n - n\alpha), \end{cases}$$

it implies that $\forall n \in \mathbb{N}$, $\mathcal{U}_f^{(n)}$ is higher than $\mathcal{U}_d^{(n)}$. Or, simple calculations easily prove that:

$$\mathcal{U}_f = \lim_{n\to+\infty} \mathcal{U}_f^{(n)},$$

and:

$$\mathcal{U}_d = \lim_{n\to+\infty} \mathcal{U}_d^{(n)},$$

so that $\mathcal{U}_d \leq \mathcal{U}_f$.

To demonstrate that $\mathcal{U}_0 \leq \mathcal{U}_d$, we will refer once again to lemma 2. Under the partial information regime, the individual may patronize the activities in any order, along any strategy. Assume she decides to visit $d_0/\delta + 1$ activities, then to choose the best one. Then her expected utility is $\mathcal{U}_0$. Or, according to Lemma 2, the expected utility of the individual is necessarily lower if she applies this strategy than if she applies the optimal stopping rule. As a consequence, $\mathcal{U}_0 \leq \mathcal{U}_d$. Q.E.D.

Furthermore, the expected utilities can be compared for small values of $\delta$ using Equations (5), (18) and (21); while the expected distances can be compared using Equations (4), (16) and (22).

**Corollary 4** *When $\delta$ gets close to zero, the expected distances the individual covers under the distinct regimes tend towards the same limit:*

$$d_0 = d_d = d_f = \mu/\alpha. \tag{23}$$

*The expected utilities of the individual under each regime behave asymptotically as follows:*

$$\mathcal{U}_f = \mathcal{U}_d + \mu\gamma = \mathcal{U}_0 + \mu. \tag{24}$$

Note that this difference is equal to $\mu$ up to a multiplicative constant. More information is preferred by the individual, all the more as the individual pays more attention to the quality of the activities. We show however in Section 5 that this statement does not necessarily hold anymore when congestion arises.

[FIGURE 2]

The evolution of the expected distances under each regime with respect to the inter-activity distance $\delta$ is illustrated by Figure 7. It appears from the observation of this figure, as well as from Corollary 4, that the expected distance the individual covers under the no information without commitment regime is lower than under the full information regime. This result may seem counter-intuitive. Indeed, a part of travel demand originates from the need of people to gather information about activities before choosing one of them. Giving people information at home should then decrease travel demand, by making these search trips unnecessary.

However, the contrary is observed here: in fact, a second effect, more subtle, happens, and plays a more important role. When the individual does not have full information about the qualities of the activities, she opts for a conservative strategy where she stops as soon as she has found a satisfying solution, as described by Proposition 1 under the no information with commitment regime and by Lemma 2 under the no information without commitment regime. In both cases, she may miss a better opportunity, located a bit further. She does not miss this opportunity when she has full information. This explains why the expected distance the individual covers is higher under the full information regime than under the other regimes.

This result is closely linked to the hypotheses of the model. To demonstrate that there is no simple, global relationship between information availability and choice of destination, some other situations will be presented in Section 5.

## 4.2 Value of information

From Proposition 4, the expected utility of the individual depends on the prevailing information regime. If the individual were given the choice, she would be ready to pay a certain amount of money to change from the no information regime with commitment to the no information regime without commitment, or to the full information regime. This willingness to pay can be referred to as the value of information, and is calculated in this section.

**Definition 4** *Denote by $V_d$ (resp. $V_f$) the value of information,* i.e. *the willingness of the individual to pay to change from the no information without (resp. with) commitment regime to the full information regime.*

$V_d$ and $V_f$ are easily derived: $V_d = \mathcal{U}_f - \mathcal{U}_d$, and $V_f = \mathcal{U}_f - \mathcal{U}_0$. Furthermore, from Proposition 4:

**Proposition 5** *$V_d$ and $V_f$ are always positive and:*

$$0 \leq V_d \leq V_f,$$

and, from Equation (24), we have:

**Corollary 5** *The asymptotic behaviors of $V_d$ and $V_f$ when $\delta$ is close to zero are:*

$$V_d = \mu\gamma \quad < \quad V_f = \mu$$

The individual is willing to pay for more information, because all other things equal, more information improves the utility of the individual. This improvement increases with the sensitivity $\mu$ of the individual to the qualities of the activities. The individual is also willing to pay for less constraint: the willingness of the individual to pay to change from the no information regime with commitment to the no information regime without commitment, which is equal to $V_f - V_d$, is positive.

[FIGURE 3]

The evolution of $V_p$ and $V_f$ with respect to the inter-activity distance $\delta$ is illustrated by Figure 7. At the examination of this figure, it appears that the value of information increases with the density of activities. In particular, when $\delta$ gets large, the value of information decreases towards zero. Indeed, a large $\delta$ means that the only activity easily accessible to the individual is the one located near her home. The second most accessible activity is located far away. In fact, as $\delta$ increases indefinitely, all happens as if only one activity was accessible to the individual. In that case, information is worthless for the individual, who can observe the quality of that activity at no cost.

A word should be said about the irregular behavior of $V_f$. $V_f$ is equal to the difference between $\mathcal{U}_f$ and $\mathcal{U}_0$. From Equation (20), $\mathcal{U}_f$ is clearly continuously differentiable. This is not the case of $\mathcal{U}_0$, as it appears from Equation (3): $\mathcal{U}_0$ is indeed continuous, but not differentiable for each value of $\delta$ around which $d_0$ changes. This explains the kinks in the curve of $V_f$.

Finally, it is possible to propose from Figure 7 a simple numeric example. Assume the distance between activities is $\delta = 5$ km; the individual is then ready to pay \$4.5 in order to be able to shift from the no information, with commitment regime to the no information, without commitment regime, and about \$3 more to shift to the full information regime.

The model analyzed in this section has yielded some slightly counter-intuitive results concerning the linkage between information availability and travel demand. The question remains though to study how much the results derived here depend on the model's hypotheses. Elements of answer are provided in the next section.

# 5 Complementary analyses

In this section, we show that there is no simple relationship between information availability and travel demand, and that information availability may decrease social welfare when there are more than one individual in the transport system and when there is congestion.

To do so, two very simple models are quickly presented. The first one, presented in Subsection 5.1, concerns a unique individual who has to chose one among two activities. Using this model, we show that providing information does not necessarily increase transport demand.

The second model, presented in Subsection 5.2, is almost identical to the first one, except for the fact that there are two individuals, and that the road the individuals pass on to reach activities is prone to congestion.

## 5.1 Impact of information provision on mobility

Consider an individual located in $A$ having the possibility to patronize an activity located in $A$ for no transport cost, or to patronize an activity located in $B$ for a transport cost normalised to 1. The quality of activity $A$ is set to zero without loss of generality. The quality of activity $B$ is represented by a centered random variable $\varepsilon$, equal to 0 and 4 with probabilities 0.5. Recall that the individual is risk neutral.

Without information, the expected utility of the individual of going to $A$ is always zero. Her expected utility if she goes to $B$ is $2 - 1 = 1$. If she does not have more information, she will always go to $B$. The expected distance she covers is then $d_0 = 1$, and her expected utility is $\mathcal{U}_o = 1$.

With information, the individual knows the qualities of the activities. If the quality of activity $B$ is zero, she chooses activity $A$. On the contrary case, she chooses activity $B$. The expected distance she covers is:

$$d_f = 0.5 \times 0 + 0.5 \times 1 = 0.5.$$

Her expected utility is:

$$\mathcal{U}_f = 0.5 \times 0 + 0.5 \times (4 - 1) = 1.5.$$

The individual's expected utility is higher with full information, but the transport demand has decreased in this case (from 1 to 0.5).

## 5.2 Impact of information provision on social welfare with congestion

In all the situations considered up to now, the expected utility of the individual was improved by information provision. This may be different if congestion arises, especially in the case where information provision has a positive effect on transport demand (in the contrary case, it may have a doubly beneficial effect by reducing congestion in addition to the intrinsic value of information for the individuals). This is illustrated here in a simple case.

Consider two individuals located in $A$, who have to choose between an activity located in $A$ and an activity located in $B$. In order to go to $B$, they must use a congestion-prone road. On this road, the travel time is set to zero if the traffic is limited to one individual. It increases to 1 if the two individuals use the road simultaneously.

For both individuals, the quality of activity $A$ is zero. The quality of activity $B$ from the perspective of individual 1 is 1000. The value of time of individual 1 is 500. On the contrary, individual 2 does not know the quality of activity $B$. She knows this quality is a centered random variable $\varepsilon$, taking the values 1 and -1 with probability 0.5. Her value of time is 0.5. Both the individuals are risk neutral.

**No information regime** Under this regime, individual 2 does not observe the quality of activity $B$ *ex ante*. Whatever the decision of individual 2, individual 1 prefers to go to $B$. Conversely, given the information she has, and given the fact that individual 1 will go to $B$ anyway, the expected utility for individual $A$ to go to $B$ is $0 - 1 = -1$. It is preferable for her to choose activity $A$. The total travel demand is $d_0 = 1$. Given that individual 2 does not go to $B$, the travel cost for individual 1 is zero. individual 1's expected utility is $\mathcal{U}_0^{(1)} = 1000$. individual 2's expected utility is $\mathcal{U}_0^{(2)} = 0$. The social welfare under the no information regime is therefore:

$$W_0 = \mathcal{U}_0^{(1)} + \mathcal{U}_0^{(2)} = 1000.$$

**Full information regime** Under this regime, individual 2 knows the realization of $\varepsilon$. If it is -1, she goes to activity $A$. Conditionally to that case, the sum of the utilities of the two individuals is:

$$W_f(\{\varepsilon = -1\}) = 1000.$$

On the contrary, if the realization of $\varepsilon$ is 1, individual 2 goes to $B$, whatever the decision of individual 1. In this case, individual 1 still prefers to go to activity $B$. The transport demand is $d_f = 2$. The utility of individual 1 is then $\mathcal{U}_f^{(1)}(\{\varepsilon = 1\}) = 500$, and the utility of individual 2 is $\mathcal{U}_f^{(2)}(\{\varepsilon = 1\}) = 0.5$. Therefore, the sum of their utilities is:

$$W_f(\{\varepsilon = 1\}) = 500.5.$$

The social welfare under the full information regime thus depends on the realization of $\varepsilon$. Its expected value is:

$$W_f = 0.5W_f(\{\varepsilon = -1\}) + 0.5W_f(\{\varepsilon = 1\}) = 725.25.$$

Under congestion, providing information to the individuals can decrease the expected social welfare.

These two simple examples are complementary, and both show that the impact of providing information to the users of a transport system can be ambiguous. It can increase distance travelled (see Section 5.1) or even decrease social welfare with unpriced congestion (see Section 5.2).

# 6 Conclusion

This paper presents a simple model, were an individual chooses one activity among a set of alternatives located on a linear road. The behavior of the individual is modeled using discrete choice theory. Results of stopping rule theory are used to derive the results of the paper, i.e. the impact of information on demand for travel.

Three information regimes are examined. Under the no information with commitment regime, the individual has no information ex ante on the match values of the activities, and resorts to a simple heuristic: she determines ex ante the number of activities she will visit. Under the no information without commitment regime, the individual has no information ex ante on the match values of the activities, and she visits the activities along an optimal strategy. This optimal strategy is simple: the first time the match value of a visited activity is higher than an endogenous threshold, the individual stops her search and patronizes this activity. Under the full information regime, the individual knows the matching values of all the activities ex ante. Then, the individual directly chooses the activity which maximizes her net benefit.

In each case, the expected distance covered by the individual is calculated, as well as the variance of this distance, and the individual's expected utility. Closed formulas are not systematically available. When necessary, asymptotic results are derived for large densities of activities.

This model allows for the explicit representation of information availability, and its influence on travel-related decisions. More precisely, it provides an analytical, microeconomic framework for the analysis of the impact of information availability on the choice of destination, and on the utilities of individuals. Some results are derived. First, in the model, the expected utility of the individual is higher under the no information without commitment regime than under the no information with commitment, and even higher under the full information regime. The individual is thus willing to pay for information. The critical assumptions behind the value of information in this model are the heterogeneity of the match values, and the fact that observing these match values entails a cost.

Second, the individual does not choose her destination in the same way with and without full information. With full information, the individual selects directly the best activity, and goes to it. Without information, she has to travel to acquire information. In that case, one could expect that the distance the individual covers is higher, due to the fact that a part of her trip has to be dedicated to the acquisition of information. However, the contrary is observed, which can be interpreted as follows: when the individual has no information, she opts for a conservative protocol of which the objective is to obtain a satisfying result. Better opportunities, located a bit further than the last activity the individual has visited, can then be missed. In the end, the average distance covered is higher under the full information regime than under the two no information regimes.

This result argues against the intuition that more information (e.g. through the expansion of the use of the Internet) mechanically induces less travel demand. Finally, the relationship between information availability and travel demand can be even more intricate when other phenomena, such as congestion, come into the picture. This is illustrated by a small set of simple examples presented at the end of the paper.

The model and results of this paper are based on strong hypotheses, and can thus be improved and extended in several ways. Besides, econometric studies would usefully complete the analytic approach presented here. Some empirical results are already available, although they are usually limited to very specific frameworks. For example, Ferrell (2004) examines the linkage between e-shopping and shopping travel demand; Mokhtarian and Circella (2007) analyses the social factors explaining the intensity of e-shopping.

It is also desirable, for practical application, to improve the realism of the hypotheses concerned with the economic environment and behavior of the individual. In our setting, prices are implicit, and fixed. It some applications, it may be required to consider endogenous prices, as proposed in a linear market by Anderson and Renault (1999). The search strategies change the price pattern in a very simple manner in Anderson and Renault. More realism should be considered for the spacial setting. To study this problem, an explicit network and more elaborated trip search routines should be envisaged (this is also a criticism of our paper). This problem is considerably more complex, since each user should determine her optimal path in the network. We believe that such behavior should be first investigated empirically.

It should also be noted that in the model, the individuals do not have an outside option: they must choose an activity, and cannot decide to opt for say a home delivery. Introducing an outside option in the model makes the analysis more difficult (and results will be implicit). Finally, in this framework, the individual often compares decisions which yield stochastic outcomes: it would be interesting to investigate the influence of risk aversion on the results.

Finally, this work may have implications in the field of activity-based modeling. In activity-based modeling, travel demand derives from the need of individuals to pursue activities located at different places, so that activity schedules and travel pattern decisions are closely intertwined (see e.g. Axhausen and Gärling, 1992; Bhat and Koppelman, 1993). However, it is generally assumed that the individuals have full information about the various available activities. This paper indicates how imperfect information, and its impact on the behavior of individuals, could be introduced in activity-based modeling.

## 7 Aknowledgements

We would like to thank Predit, French Transport Department (ENS Cachan and KUL) for their financial support. The second author would like to thank the Institut Universitaire de France. We would like to thank the editor Hani Mahmassani and anonymous referees for their helpful comments. Both authors benefited from discussion from Stef Proost, Michel de Lara, Nathalie Picard, Jean-Luc Prigent and Nicolas Wagner. Finally, we would like to thank three anonymous referees, for their detailed and helpful comments.

## References

Anderson, S. P. and Renault, R. (1999), 'Pricing, product diversity and search costs: a bertrand-chamberlin-diamond model', *RAND Journal* *of Economics* **30**(4), 719–735.

Axhausen, K. and Gärling, T. (1992), 'Activity-based approaches to travel analysis: conceptual frameworks, models, and research problems', *Transport Reviews* **12**, 324–341.

Barnhart, C. and Sheffi, Y. (1993), 'A network-based primal-dual heuristic for the solution of multicommodity network flow problems', *Transportation Science* **27**(2), 102–117.

Ben-Akiva, M. and Lerman, S. (1985), *Discrete Choice Analysis*, MIT Press, Cambridge.

Bhat, C. and Koppelman, F. S. (1993), 'A conceptual framework of individual activity program generation', *Transportation Research Part A* **27**(6), 433–446.

de Lara, M., Chancelier, J.-P. and de Palma, A. (2007), 'Road-choice and the one-armed bandit problem', *Transportation Science* **41**(1), 1–14.

de Palma, A. and Picard, N. (2006), 'Equilibria and information provision in risky networks with risk-adverse drivers', *Transportation Science* **40**(4), 393–408.

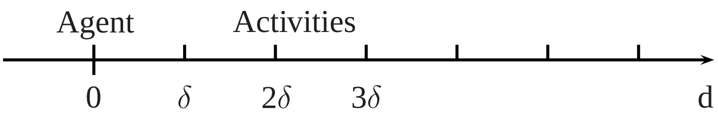
Ferrell, C. E. (2004), 'Home-based teleshoppers and shopping travel', *Transportation Research Record* (1894), 241–248.

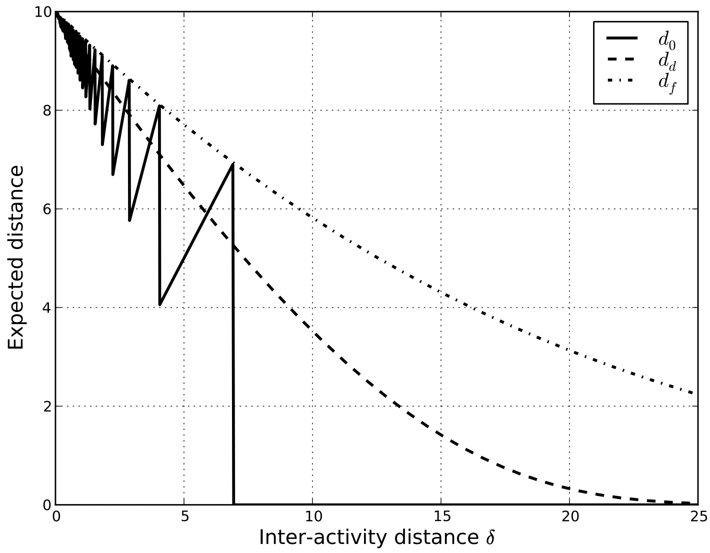
Gollier, C., Eeckhoudt, L. and Schlesinger, H. (2005), *Economic and Financial Decisions under Risk*, Princeton University Press.

Mokhtarian, P. L. and Circella, G. (2007), The role of social factors in store and internet purchase frequencies of clothing/shoes, *in* 'International Workshop on Frontiers Transportation: Social Interactions'.

Robbins, H. (1970), 'Optimal stopping', *American Mathematical Monthly* **77**(4), 333–343.

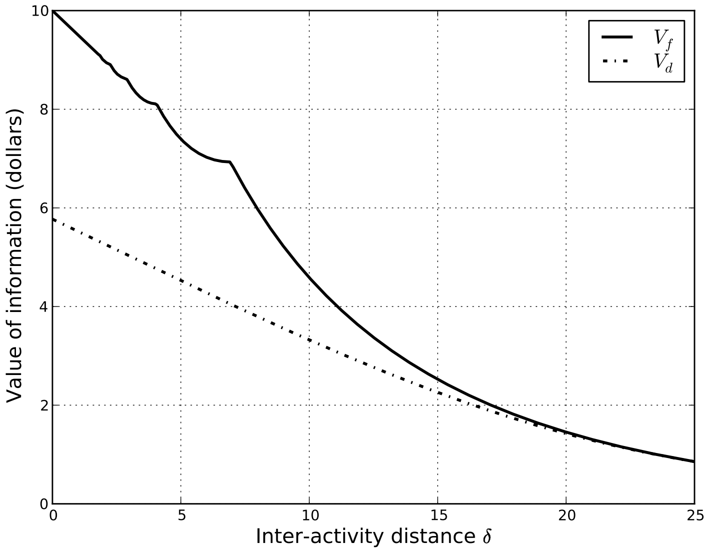
[FIGURE 1]





10
8
6
4
2
0
0
5
10
15
20
25
Expected distance
Inter-activity distance $\delta$
$d_0$
$d_d$
$d_f$


[FIGURE 2]



10
8
6
4
2
0
Value of information (dollars)
0
5
10
15
20
25
Inter-activity distance $\delta$
$V_f$
$V_d$


[FIGURE 3]

[FIGURE TITLES]

[FIGURE 1]

Model framework

[FIGURE 2]

Expected distance covered by the individual.

($\mu = 10$, $\alpha = 1$ \$/mile)

[FIGURE 3]

Value of information

($\mu = 10$, and $\alpha = 1$ \$/mile)

# A Proof of Corollary 1

First and foremost, note that when $\delta$ is close to 0:

$$\exp(\delta\alpha/\mu) - 1 = \delta\alpha/\mu + o_{0^+}(\delta),$$

where $o_{0^+}(f)$ denotes any term $g$ verifying $\lim_{\delta\to 0^+} g(\delta)/f(\delta) = 0$ (in particular, $o_{0^+}(1)$ denotes any term tending toward zero when $\delta$ gets close to zero). Therefore:

$$\frac{1}{\exp\left(\frac{\delta\alpha}{\mu}\right) - 1} = \frac{1}{\frac{\delta\alpha}{\mu} + o_{0^+}(\delta)} = \frac{1}{\delta}\frac{1}{\frac{\alpha}{\mu} + o_{0^+}(1)} = \frac{1}{\delta}\left(\frac{\mu}{\alpha} + o_{0^+}(1)\right)$$

$$= \frac{\mu}{\delta\alpha} + o_{0^+}\left(\frac{1}{\delta}\right).$$

Given the fact that $\mu/\alpha\delta + o_{0^+}(1/\delta)$ increases indefinitely when $\delta$ tends towards zero, and the fact that $\forall x, x - \lfloor x \rfloor < 1$:

$$\left\lfloor \frac{1}{\exp\left(\frac{\alpha\delta}{\mu}\right) - 1} \right\rfloor = \frac{1}{\exp\left(\frac{\alpha\delta}{\mu}\right) - 1} + o_{0^+}\left(\frac{1}{\delta}\right) = \frac{\mu}{\delta\alpha} + o_{0^+}\left(\frac{1}{\delta}\right),$$

and:

$$\delta \left\lfloor \frac{1}{\exp\left(\frac{\alpha\delta}{\mu}\right) - 1} \right\rfloor = \frac{\mu}{\alpha} + o_{0^+}(1).$$

Therefore:

$$d_0 = \frac{\mu}{\alpha} + o_{0^+}(1).$$

Furthermore, since $x \mapsto \ln(x)$ is a concave function whose derivative tends toward 0 when $x$ tends toward $+\infty$:

$$\ln\left(\left\lfloor \frac{1}{\exp\left(\frac{\alpha\delta}{\mu}\right)-1} \right\rfloor + 1\right) = \ln\left(\frac{\mu}{\alpha\delta}\right) + o_{0^+}(1),$$

so that:

$$\mathcal{U}_0 = -\mu\ln(\delta) + \mu\ln\left(\frac{\mu}{\alpha}\right) - \ln(\mu) + o_{0^+}(1)$$

Q.E.D.

# B Proof of Lemma 2

The objective is to find the stopping rule maximising the objective function given by Equation (10). A stopping rule is a random variable $\tau$, function of $\{\varepsilon_i\}$, with values in $\mathbb{N}$ such that the event $\tau = n$ depends solely on $\varepsilon_0, \ldots, \varepsilon_n$. Denote $t$ the stopping rule of the Lemma. It can be rewritten as follows:

$$\{t = n\} \Leftrightarrow \begin{cases} \forall i \in [0; n-1], & \{\varepsilon_i < \varepsilon_c\} \\ \text{and :} & \{\varepsilon_n \geq \varepsilon_c\}, \end{cases} \tag{25}$$

As a consequence:

$$\mathbb{P}(\{t = n\}) = \mathbb{P}((\varepsilon_n \geq e_c) \cap (\varepsilon_{n-1} < e_c) \cap \ldots \cap (\varepsilon_0 < e_c)).$$

It can be rewritten as:

$$\mathbb{P}(\{t = n\}) = \mathbb{P}(\varepsilon_n \geq e_c).\mathbb{P}(\varepsilon_{n-1} < e_c).\cdots.\mathbb{P}(\varepsilon_0 < e_c),$$

or:

$$\mathbb{P}(\{t = n\}) = (1 - F(e_c))(F(e_c))^n.$$

Denote by $x_t$ the random variable corresponding to the utility of the individual applying the stopping rule $t$. Given Equation (25), the expected utility of the individual conditionally to the fact that she has stopped her search at step $n$ when applying the stopping rule $t$ is:

$$\mathbb{E}(x_t | t = n) = \frac{1}{1 - F(e_c)} \int_{e_c}^{+\infty} ef(e)de - n\delta\alpha,$$

which can be rewritten as:

$$\mathbb{E}(x_t | t = n) = e_c - \left(n - \frac{1}{1 - F(e_c)}\right) \delta\alpha.$$

But:

$$\mathbb{E}(x_t) = \sum_{n=0}^{+\infty} \mathbb{P}(\{t = n\}) \mathbb{E}(x_t | t = n),$$

so that:

$$\mathbb{E}(x_t) = e_c + \delta\alpha.$$

To prove that $t$ is an optimal stopping rule, consider another stopping rule denoted $t'$. The individual stops at $t'$, denote $x_{t'}$ her utility. Assume $\mathbb{E}(x_{t'})$ exists and is not $+\infty$. The central argument of the proof in Robbins (1970) is the following inequality, which holds for all $b \in \mathbb{R}$:

$$\max(\varepsilon_0; \ldots; \varepsilon_n) - (n+1)\delta\alpha \leq b + \sum_0^n ((\varepsilon_i - b)^+ - \delta\alpha).$$

Disregarding some methodological precautions, we have:

$$x_{t'} = \max(\varepsilon_0; \ldots; \varepsilon_{t'}) - t'\delta\alpha,$$

so that:

$$x_{t'} < b + \delta\alpha + \sum_0^{t'} ((\varepsilon_i - b)^+ - \delta\alpha).$$

Now assume $b > e_c$. Then, for all $i$:

$$\mathbb{E}((\varepsilon_i - b)^+ - \delta\alpha) < 0$$

Given that we have (refer to (Robbins, 1970) for the proof of this particular equation):

$$\mathbb{E}\left(\sum_0^{t'} ((\varepsilon_i - b)^+ - \delta\alpha)\right) = \mathbb{E}(t')\mathbb{E}((\varepsilon_i - b)^+ - \delta\alpha),$$

necessarily:

$$\mathbb{E}(x_{t'}) < b + \delta\alpha.$$

This inequality is true for all $b > e_c$, so that:

$$\mathbb{E}(x_{t'}) \leq e_c + \delta\alpha.$$

Q.E.D.

# C Proof of Corollary 3

The following lemma will prove useful.

**Lemma 3** *When $\delta$ is close to zero, $e_c$ behaves as follows:*

$$e_c = -\mu \ln(\delta) + \mu \ln(\mu/\alpha) - \mu\gamma + o_{0^+}(1),$$

*and $F(e_c)$ behaves as:*

$$F(e_c) = 1 - \alpha\delta\mu + \alpha^2\delta^2/4\mu^2 + o_{0^+}(\delta^2).$$

*Proof:* See below.

Replacing $F(e_c)$ in Equation (13) we obtain, for small values of $\delta$: $d_d = \mu/\alpha - 3\delta/4 + o_{0^+}(\delta)$ and: $\sigma_d = \mu/\alpha - \delta/4 + o_{0^+}(\delta)$ We will now derive an asymptotic development of $\mathcal{U}_d$ for small values of $\delta$. From Equation (15), $\mathcal{U}_d = e_c + o_{0^+}(1)$. So $\mathcal{U}_d = -\mu \ln(\delta) + \mu \ln(\mu/\alpha) - \mu\gamma + o_{0^+}(1)$. Q.E.D.

# D Proof of Lemma 3

The term $e_c$ is implicitly defined by Equation (8):

$$\int_{e_c}^{+\infty} (e - e_c) f(e) \mathrm{d}e = \delta\alpha.$$

Consider the following function of $t$:

$$G(t) = \int_{t}^{+\infty} (e - t) f(e) \mathrm{d}t.$$

Its first derivative with respect to $t$ is $G'(t) = F(t) - 1$. It is strictly negative. So $G$ is a strictly decreasing function of $t$. As a consequence, $e_c$ is a decreasing function of $\delta$. Therefore, it necessarily has a limit when $\delta$ tends toward zero, and it is straightforward that this limit is $+\infty$. But $F(e_c) = \exp\left(-\exp\left(-e_c/\mu - \gamma\right)\right)$, and :

$$\lim_{\delta \to 0^+} \exp\left(-e_c/\mu - \gamma\right) = 0, \tag{26}$$

so the approximation of $F(e_c)$ for small values of $\delta$ is:

$$\begin{aligned} F(e_c) &= 1 - \exp\left(-e_c/\mu - \gamma\right) + 1/2 \exp\left(-2e_c/\mu - 2\gamma\right) \\ &\quad + o_{0^+}\left(\exp\left(-2e_c/\mu - 2\gamma\right)\right). \end{aligned} \tag{27}$$

Differentiating Equation (8) with respect to $\delta$ gives $-\frac{\mathrm{d}}{\mathrm{d}\delta}(e_c) \int_{e_c}^{+\infty} f(e)\mathrm{d}e = \alpha$. The derivative of $e_c$ with respect to $\delta$ therefore verifies the following equality: $\frac{\mathrm{d}}{\mathrm{d}\delta}(e_c) = -\alpha/(1 - F(e_c))$. Using Equation (27), we obtain:

$$\begin{aligned} 1 - F(e_c) &= \exp\left(-e_c/\mu - \gamma\right) - 1/2 \exp\left(-2e_c/\mu - 2\gamma\right) \\ &\quad + o_{0^+}\left(\exp\left(-2e_c/\mu - 2\gamma\right)\right), \end{aligned}$$

which can be written:

$$\begin{aligned} 1 - F(e_c) &= \exp\left(-e_c/\mu - \gamma\right)\left[1 - 1/2 \exp\left(-e_c/\mu - \gamma\right)\right. \\ &\quad \left. + o_{0^+}\left(\exp\left(-e_c/\mu - \gamma\right)\right)\right]. \end{aligned}$$

Therefore, using Equation (26):

$$1/(1 - F(e_c)) = \exp\left(e_c/\mu + \gamma\right) + 1/2 + o_{0^+}(1),$$

we get:

$$\frac{\mathrm{d}}{\mathrm{d}\delta}(e_c) = -\alpha \exp\left(e_c/\mu + \gamma\right) - \alpha/2 + o_{0^+}(1).$$

Consider now:

$$\frac{\mathrm{d}}{\mathrm{d}\delta}(F(e_c)) = f(e_c)\frac{\mathrm{d}}{\mathrm{d}\delta}(e_c),$$

then:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\delta}(F(e_c)) = 1/\mu \exp\left(-\frac{e_c}{\mu} - \gamma\right) \exp\left(-\exp\left(-\frac{e_c}{\mu} - \gamma\right)\right) \\ \times \left[-\alpha \exp\left(\frac{e_c}{\mu} + \gamma\right) - \frac{\alpha}{2} + o_{0^+}(1)\right], \end{aligned}$$

and, using Equation (27), we get:

$$\frac{\mathrm{d}}{\mathrm{d}\delta}(F(e_c)) = -\frac{p}{\mu} + \frac{p}{2\mu}\exp\left(-\frac{e_c}{\mu} - \gamma\right) + o_{0^+}\left(\exp\left(-\frac{e_c}{\mu} - \gamma\right)\right), \quad (28)$$

which implies:

$$\frac{\mathrm{d}}{\mathrm{d}\delta}(F(e_c)) = -\alpha/\mu + o_{0^+}(1),$$

As a consequence:

$$F(e_c) = 1 - \alpha\delta/\mu + o_{0^+}(\delta).$$

Besides:

$$-\exp\left(-e_c/\mu - \gamma\right) = \ln\left(1 - \alpha\delta/\mu + o_{0^+}(\delta)\right),$$

so that, when developing the right-hand side:

$$\exp\left(-e_c/\mu - \gamma\right) = \alpha\delta/\mu + o_{0^+}(\delta),$$

This equation can be simply be rewritten as:

$$\exp\left(-e_c/\mu - \gamma\right) = \delta\left(\alpha/\mu + o_{0^+}(1)\right),$$

we obtain:

$$-e_c/\mu - \gamma = \ln(\delta) + \ln\left(\alpha/\mu + o_{0^+}(1)\right),$$

and, by developing again the right-hand side, we finally obtain:

$$e_c = -\mu \ln(\delta) - \mu \ln(\alpha/\mu) - \mu\gamma + o_{0^+}(1).$$

We can rewrite Equation (28) as follows:

$$\frac{\mathrm{d}}{\mathrm{d}\delta}(F(e_c)) = -\alpha\mu + \alpha^2\delta/2\mu^2 + o_{0^+}(\delta).$$

But, for any given value of $\delta$, say $\alpha$:

$$\int_0^\alpha \frac{\mathrm{d}}{\mathrm{d}\delta}(F(e_c))\Big|_{\delta=t}\mathrm{d}t = F(e_c)\Big|_{\delta=\alpha} - F(e_c)\Big|_{\delta=0},$$

and $F(e_c)|_{\delta=0} = 1$. As a consequence, $F(e_c) = 1 - \alpha\delta/\mu + \alpha^2\delta^2/4\mu^2 + o(\delta^2)$.
Q.E.D.