TI 2004-082/1
Tinbergen Institute Discussion Paper

Information Overload in Monopsony Markets

Stefano Ficco

Faculty of Economics, Erasmus Universiteit Rotterdam, and Tinbergen Institute.

Tinbergen Institute
The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Please send questions and/or remarks of non-scientific nature to driessen@tinbergen.nl.
Most TI discussion papers can be downloaded at http://www.tinbergen.nl.

Information Overload in Monopsony Markets *

Stefano Ficco[†]
Tinbergen Institute, Erasmus University Rotterdam

July 15, 2004

Abstract

I consider a situation in which heterogenous senders (applicants) compete in order to be selected by one receiver (employer). Productivity is private information to the senders, and the receiver processes imperfect signals (applications) to screen among applicants. The information-processing technology is imperfect: the accuracy of each signal in predicting the unknown productivity decreases with the total number of signals processed. I show that, for a sufficiently large market, information overload occurs as there exist equilibria in which too many people apply and the receiver neglects some applications. For any information-processing technology level, information overload equilibria emerge when the cost of sending applications is low relatively to the existing technology level. The magnitude of information overload is bounded and it is larger if the receiver cannot neglect applications. As a result, an overloaded market in which the receiver has to process all applications is less efficient than an overloaded market where neglecting excessive information is an option.

KEYWORDS: Imperfect information-processing technology, quality and quantity of information, information overload.

JEL CODES: C72, D82, J42.

*I am indebted to Maarten Janssen and Vladimir Karamychev for helpful suggestions and inspiring comments. I wish also to thank colleagues and seminar participants in Rotterdam for useful discussions.

[†]Erasmus University Rotterdam, Tinbergen Institute, Burgemeester Oudlaan 50, 3062 PA Rotterdam, The Netherlands. E-mail: ficco@few.eur.nl. Fax: +31-10-4089147.

1 Introduction.

Typically when economic literature deals with information issues it argues that market inefficiency can be explained in terms of a ”lack” of information (e.g. imperfect or asymmetric information). An indirect implication of such an argument is that the more information is available, the better it is for the efficiency of the market. However in the last years the cost of transmitting information has fallen dramatically and the scarce resource in communication flows is becoming the human attention needed to interpret information. Consequently, problems of information overload may emerge. This consideration naturally leads to question whether the paradigm "the more information the better" holds true in markets where the cost of generating information is low and the resources to process information are limited . This paper addresses this question for the case of a monopsony market.

Consider a situation in which many job candidates compete in order to be hired by one employer. If productivity is private information to employees it is natural to define an information item as the imprecise signal (e.g. one application containing C.V., reference letters and so on) that the employer uses in order to assess the ability of each worker. Therefore, the amount of information generated by the market is endogenously determined by the number of applications actually sent. Provided that information can be processed at no cost, a large monopsony market which generates a large amount of information performs very efficiently. The are two main effects behind this conclusion. The first lies on the supply side and it is the beneficial role of *self selection*: a larger market means more competition which translates in higher workers' productivity. The second lies on the demand side and is a *sample size effect*: a larger amount of information means more people applying for a position, thus a higher probability that a very good worker is available and, eventually, hired. Clearly these two mechanisms may be not be preserved if, on the contrary, the information-processing phase is costly. This naturally leads me to analyze a monopsony model in which the agent in charge of processing information (the monopsonist) bears a cost in doing so. The cost I consider here is not the monetary cost of the screening process, but the opportunity cost of the time and resources needed to process and interpret the available information. Implicit in this assumption is the idea that the resources available to interpret information are limited.

I analyze a model where one receiver (employer) screens among a number of competing senders (applicants) in order to select the best one. Senders' abilities are uniformly distributed. Each sender knows his own ability but does not know the abilities of other applicants. To compete for the position a sender must send a signal (application form) that include some proxies about his ability Sending one signal has a cost, which is equal for all senders. The utility that the sender who obtains the vacancy gets from the position is also assumed to be the same for all senders. The receiver, ex ante, only knows the distribution of abilities. However, by processing signals (reading the applications), the receiver can update the information about the ability of each sender and consequently rank them in order to select the best one. This assessment method is imprecise and imperfect. The imprecision is captured by the noisy nature of signals. The novel assumption that I make here is that of *imperfect information-processing technology*: as the number of applications processed becomes large, the capability of the receiver to rank applicants according to their true abilities vanishes. Therefore I assume that the receiver, after observing the number of applications received, decides first how many of them to read, neglecting, in case, those in excess. Afterwards applications are processed simultaneously and eventually one applicant is selected. Thus, the choice of how many applications to read is a strategic choice of how much attention to allocate to interpret each application. The focus of the paper is to analyze how market mechanisms (*self selection* and *sample size* mechanisms) shape the way in which information is produced and processed, and what are the implications in terms of efficiency

The first finding is that, for a sufficiently large market, there exist equilibria (pooling or partially separating) in which the economic agent in charge of processing applications neglects some of them. These equilibria arise because the positive *sample size* effect is outweighed by the negative *decreasing accuracy* effect: even though the chance of getting really good applicants increases with the number of applications received, it also decreases the actual capability of the receiver to discriminate between applications. Thus, neglecting applications is a consequence of maximizing behavior where the marginal utility of a larger sample size (which is the chance of observing a better application) is compared to its opportunity cost (which is the marginal decrease in the capability of ranking the applicants according to their true abilities). Neglecting applications corresponds to neglecting some potentially valuable information, thus, I refer to this phenomenon as *information overload*. I also notice that, for any information-processing technology level, information overload equilibria emerge when the cost of sending applications is low relatively to the existing technology level.

Second, I find that, provided that sending application is costly, the *self selection* mechanism is preserved: in other words, an increase in the number of potential applicants eventually translates only in higher applicants' ability. This occurs because competition discourages low ability employees from applying and thus, in equilibrium, only the better workers, who have more chances to be hired, are willing to bear the application cost. A direct consequence of the *self selection* mechanism is that, for any positive sending cost, there exists an upper bound on the amount of information that the market can generate and therefore on the magnitude of information overload. The reason is fairly intuitive. As the number of potential senders increases, the competition becomes stronger and, eventually, separation occurs. When this happens self-selection comes to play a role: an additional increase in the market size affects only the ability of the candidates that actually apply, but not their number which, on the contrary, remains constant. Therefore, a positive cost on the supply side identifies a hard bound on the amount of information that is produced, while the technology level identifies a hard bound on the maximum amount of information that the market is willing to absorb. When the cost is low relatively to the technology level, the former bound is larger than the latter and the magnitude of the overload is also bounded.

Third, I compare the efficiency of markets that suffer from information congestion for two different scenarios: one in which applications can be neglected, and one in which this is not possible. I show that an overloaded market where excessive information is neglected is more efficient than one in which the receiver is constrained to process all information. In the first scenario the receiver, by ignoring some information, protects himself from the *decreasing accuracy* effect, thus, the inefficiency lies only on the supply side as there are too high sending costs for the society. In the second scenario, on the contrary, the side effects of information congestion are present also on the demand side: the excessive number of applications processed decreases the utility of the receiver. Moreover, the sending cost are even higher compared to the situation in which applications can be neglected. Two effects play a role in shaping this result. First, if some applications are neglected in equilibrium, senders anticipate that there is a chance that their application might not be taken into account in the first place. Second, when applications are neglected, the signals' accuracy is larger compared to the case in which all excessive applications are processed, and this helps the receiver to sort out relatively bad candidates. Both effects have a discouraging impact on senders and, therefore, a market in which neglecting applicants is an option induces less people to apply.

Finally, I note that, in equilibrium, the total sending costs always decrease in the technology level. Moreover, if information is neglected in equilibrium, the total sending cost does not depend on the cost of sending one single application. The reason for this surprising result is that the elasticity of the number of actual senders with respect to the cost of sending one single application is always unitary: as the cost of sending one application increases the number of applicants in equilibrium decreases in the same proportion and therefore the total cost does not change. If, on the other hand, neglecting applications is not an option, the total sending costs are decreasing in the cost of sending one single application. Thus, in this second scenario, the elasticity of the number of actual senders with respect to the cost of sending one single application is negative and larger than one in absolute value.

The paper relates to two branches of the economics literature: one that analyzes the economic consequences of asymmetric information, and one on limited capacity. Broadly speaking, the former research field was started by Akerlof (1970). More specifically, a large body of the literature has been devoted to investigate the implications of asymmetric information in the labour market[1]. The first contribution in this direction is the pioneering work of Spence (1973). Large part of the subsequent literature is game theoretic in nature[2] and, typically, it focuses on models where two

[1] See Weiss (1995) for a survey of the different ways information asymmetries have been used in studies of the labour market.

[2] See e.g., Cho and Kreps (1987) for the first complete game theoretic analysis of the Spence model.

firms compete for one worker whose ability is unknown. On the contrary, Janssen (2002) considers a screening model in which N workers compete for one job in one firm. I stay within the set up of Janssen (2002), given that I also consider a monopsony situation[3] where the number of potential competitors is exogenous. However, a wage mechanism is not included here. This simplifying assumption is justified as the model is intended to describe the problem of a selection decision at the time in which the decision takes place. Thus, the time horizon within which the model applies is "instantaneous" and the wage is assumed to have already been set at some previous stage [4]. Alternatively, the model can represent economic situations (i.e. beauty contests, prize competitions, grant allocations...) in which the main objective is that of selecting the candidate with the highest expected ability. Yet, the model I propose differs substantially form traditional sorting[5] models. In sorting models a worker has complete control on his signalling activity: by choosing an appropriate investment level (e.g. years of education[6]) he fully determines his signal outcome. In a separating equilibrium signals happen to be fully informative because they allow the employer to sort high ability workers form low ability ones. Separation occurs because, in traditional sorting models, it is assumed that education is more costly the lower is a worker's ability. In my model, on the contrary, a worker does not have any control on his specific signal level, which is simply the outcome of a statistical experiment. Moreover the cost of signalling (that is, the cost of sending one application) does not depend on the specific worker's ability. Yet separation may still occur because better workers produce, on average, better signals and therefore have higher chances of being selected by the employer.

The concept of limited capacity in processing information has also been extensively exploited in the field of economics. The main motivation for this branch of the literature is the widely excepted idea that economic agents are not able to deal, efficiently, with an excessively large amount of information[7]. Early, as well as more recent works in marketing (see Jacoby et al. (1974) and Hahn et al. (1992), among others) provide empirical evidence that supports the adverse effect of excessive information on the quality of decision making. Consequently, models of limited capacity emphasizes the role of cognitive heuristics and simplifying knowledge structures in

[3] See Boal and Ransom (1997), for a survey of the way monopsony models have been used in studies of the labour market.

[4] Alternatively, the model can be restricted to represent certain segments of the labour market in which the wage is exogenously set by other institutions.

[5] I use the term "sort" as meant by Weiss (1995). "[I] will use the term "sorting" to refer to both signalling and screening of workers.[...] Both signalling and screening serve to "sort" workers according to their unobservable abilities.[...] In signalling models the informed (students) move first. In screening models the uninformed (firms) move first. (Weiss, 1995)".

[6] See Nöldeke and Van Damme (1990) for dynamic issues concerning signalling and education.

[7] The idea that the human brain is better equipped for working with relatively small amount of information has been already pointed out by pshychologists in the middle 1950's. See e.g. Miller (1956).

reducing information-processing demands. Lipman (1995), e.g, provides an exhaustive survey of this literature. In most of these models (see, e.g., Van Zandt (2001)) the computational processing limitation is introduced in the form of an exogenous hard bound on the information items that can be processed at a time. Therefore the possibility for the receiver to neglect information is not justified on the ground of an optimizing behavior, but is assumed to be exogenous. On the contrary, the approach that I take here is shared with Ficco and Karamychev (2004) where information overload also emerges as an endogenous behavior of fully rational agents. Yet, the focus of that paper differs as it analyzes information overload in the context of a multi-stage selection of exogenously given alternatives.

The paper is organized as follows. Section 2 describes the model. Section 3 analyzes the problem of the receiver. Section 4 analyzes the problem of the senders. Section 5 provides the equilibrium analysis. In section 6 the implications of information overload in terms of efficiency are discussed. Section 7 concludes the paper. The appendix contains all the proofs.

2 The model.

In this section I first describe the model and discuss the related economic implications. I then provide and justify a more specific assumption on the information-processing technology.

There is a mass N of heterogenous potential senders, where N is a real number $1 \leq N < \infty$ and a mass, normalized to one, of homogenous receivers. Each receiver has a vacancy to offer to a sender. The mass of senders is uniformly distributed on the $[0, 1]$ segment. Each point on the segment corresponds to the ability of a sender and is denoted by $\theta \in [0, 1]$. The decision variable for a sender is whether to send a costly application or not. The decision variables for a receiver are: how many applications to process, and which application to select from the pool of those processed. I denote by $n \geq 0$ the mass of senders who actually apply, and by m the mass of senders actually processed by the receivers. Clearly $0 \leq m \leq n \leq N$. The timing is as follows.

0. Each sender knows where he is located (he knows his own ability), and decides whether or not to send an application. I denote the sender's choice variable by s, where s can take only two values: 0 (do not send) and 1 (send). The cost of sending one application is the same for all senders and is equal to C. I will call *actual* senders those senders who have applied.

1. A random matching function allocates the mass of actual senders n, evenly to the unitary mass of receivers. Thus, each receiver receives a real number, n, of applications.

2. Each receiver observes the number of applications at his disposal, n, and decides how many of them to process, m.

3. Let Θ be a subset of $[0, 1]$ representing the set of possible ability values of actual senders and X a subset of $\mathbb{R}$ representing the set of possible signals about

θ. The receiver simultaneously processes m applications. When an application is processed it produces a signal which is drawn form the distribution $F(\cdot \mid \theta, m)$. It is assumed that the family of distributions $\{F(\cdot \mid \theta, m)\}$ is such that, for any $x_2 > x_1$ the posterior distribution $G(\cdot \mid x_2, m)$ dominates the posterior distribution $G(\cdot \mid x_1, m)$ in the sense of strict first-order stochastic dominance. Moreover it is assumed that for any $\theta_1, \theta_2 \in \Theta$ and $x \in X$

$$\lim_{m\to\infty} F(x \mid \theta_2, m) - F(x \mid \theta_1, m) = 0 \tag{A.1}$$

4. The set of signal outcomes, denoted by $\hat{X}_m \equiv \{\hat{x}_1, .., \hat{x}_m\}$, becomes private information to each receiver who compares signal outcomes and selects the one that maximizes his expected utility.

A strategy for sender θ is denoted by $s(\theta)$. A strategy for a receiver is denoted by $\{m(n), \hat{x}\}$, where $\hat{x} \in \hat{X}_m$ is the signal outcome chosen.

The payoffs are as follows. If sender θ does not apply he gets utility zero. If sender θ applies then the utility he gets equals the value of the vacancy (constant for all senders and denoted by V, where $V > C$) weighted by the probability of being selected minus the cost of signalling. The probability that sender θ assigns to the event of being selected (denoted by Φ) will generally depend on his own ability, the total number of applications sent, and the number of applications actually processed. Therefore, if n senders apply, and m are processed, the payoff of sender θ is

$$u^s(\theta, n, m) = \begin{cases} 0 & \text{if } s(\theta) = 0 \\ \Phi(\theta \mid n, m) V - C & \text{if } s(\theta) = 1 \end{cases} \tag{1}$$

The payoff that a receiver gets from selecting the sender whose signal outcome is $\hat{x}$ equals the expected ability of that sender. The expected ability of sender θ is conditional on the value of the corresponding signal outcome, $\hat{x}$, and the total number of applications processed m. Thus, the ability of a sender θ, as it is perceived by the receiver, is

$$u^r(\hat{x}, m) = E[\theta \mid \hat{x}, m] \tag{2}$$

The assumption of stage 1 is justified by the purpose of this paper which is not that of explaining the mechanism by which senders are assigned to receivers. Since receivers are homogenous, a sender does not perceive a receiver per-se more valuable than another. Therefore I intentionally avoid to model the possibility for a sender to target a particular receiver. The only crucial choice of a sender is that of sending an application or not, where sending an application is indeed a choice of whether to enter the market. When applications are sent, the size of the actual market, (the mass of actual senders n), is endogenously determined and stage 1 provides a stylized picture of the way in which the two sides of the market come to meet each other. The unitary mass of receivers corresponds to a mass of identical information processors, thus each receiver can be thought as the average representative agent of the receivers' population. From now on I will use the term Receiver (with capital R) to mean the representative receiver. Stage 1 also implies that n is not an integer, thus, it does not represent the number of applications received but is a measure of the information produced by the market. This is convenient for the analysis and, with abuse of terminology I will call n simply the "number" of applications received.

An important feature of the model I propose is that information is revealed to the Receiver by observing signals that are drawn simultaneously. This assumption is supported by Moscarini and Smith (2001)[8] and describes here economic situations in which a decision maker decides ex-ante how much to invest in the information-processing phase and, only after such decision is made, the informative outcome is revealed.

The assumptions of stage 3 imply that signals are imperfectly informative and have the monotone likelihood property (MLRP) which implies that higher signal outcomes are *more favorable than*[9] lower signal outcomes. Moreover, assumption $(A.1)$ implies that when infinitely many applications are processed the MLRP vanishes and each signal outcome becomes *equivalent* and *neutral*[10] to the decision maker. This assumption is economically justified as one can easily think of situations in which the agent in charge of processing information has limited resources in order to accomplish this task. A direct consequence of such constraint is that, if the Receiver processes infinitely many applications, he actually allocates zero resources to interpret each one of them and signals turn out to be completely uninformative.

The cost of applying is assumed not to be related to the senders' ability since its interpretation is not that of an opportunity cost, but simply the mere cost of transmitting information (i.e., filling in and sending one application). On the other hand, in this model, the opportunity cost is captured by the fact that more able senders have higher probability of being selected.

It is useful to clarify that both asymmetric information and imperfect information-processing technology are crucial to the model. On one hand, when dealing with the problem of selecting the "best" sender, the Receiver can only make an inference on the senders' ability by using a noisy signal that imprecisely represents the true ability. On the other hand, when dealing with the problem of inferring his chance of being selected, a sender can only use the knowledge about his ability to predict what signal outcome his application will produce. In the model there is a discrepancy between the value of a signal outcome and the true ability value that the signal aims to represent and, moreover, such discrepancy increases with the total number of applications processed. Therefore the larger is the amount of information processed

[8] Moscarini and Smith (2001) show that, in a dynamic continuous time world, one shot non sequential sampling is still optimal given discounting and a constant marginal cost of information.

[9] The concept of "*favorableness*" was first introduced by Milgrom (1981). If Θ is the set representing the possible values of the random parameter θ, and X is the set of possible signals about θ, then, a signal x_2 is *more favorable than* a signal x_1 iff the posterior distribution $G(\theta \mid x_2)$ dominates in the first-order stochastic sense the distribution $G(\theta \mid x_1)$.

[10] Two signals x_1 and x_2 are *equivalent* if $E(\theta \mid x_1) = E(\theta \mid x_2)$. A signal x is *neutral* if, for any prior distribution $G(\cdot)$, $G(\theta) = G(\theta \mid x)$.

by the market, the less is the efficiency with which agents can use their private information in order to make a decision.

Finally, the model can also represent a situation in which neglecting applications is not an option. This can be done by restricting the strategy space of the Receiver to be simply that of selecting one applications form those received. These two scenarios, one in which applications can be neglected, and one in which this is not possible, are compared in section 6.

2.1 The information-processing technology.

In this subsection I propose a more specific information-processing technology which has two desirable properties. Namely I will impose that the two following requirements are met:

(i) The information-processing technology implies a *trade-off between the quality and the quantity of information.*

(ii) The information-processing is *not sensitive to the ability of actual senders.*

Referring to property (i), by quality of information I mean the accuracy of each signal in predicting the unknown parameter. By quantity of information I mean the total number of signal outcomes drawn simultaneously[11]. Then the first property requires that the signals' accuracy is strictly decreasing in the sample size; the economic interpretation is that, the less resources are allocated to interpret each signal, the smaller is the accuracy of each signal in predicting the unknown parameter.

Property (ii) ensures that the signals' accuracy depends only on the total number of signals drawn and not on the specific set of parameter that signals aim to represent. This property plays an important role as the set of possible ability values of actual senders may be different for different equilibria. Indeed, notice that the fact that $\{F(\cdot \mid \theta, m)\}$ have the MLRP, implies that if a sender with ability, say, $\bar{\theta}$ applies also senders with ability above $\bar{\theta}$ apply. This allows for the existence of a partially separating equilibrium in which the segment of actual senders is $[\alpha, 1]$, where α denotes the ability of the sender who is indifferent between sending or not sending an application. Therefore the senders' strategic behavior affects the set of possible ability values, which, in case of a partially separating equilibrium, is $\Theta_\alpha = [\alpha, 1]$. Consequently, the support of family of distributions from which signal outcomes are drawn also depends on α, and I will henceforth denote it by $\{F_\alpha(\cdot \mid \theta, m)\}$. Property (ii) ensures that that the only effect that α has on the distribution $F_\alpha(\cdot \mid \cdot)$ is that of rescaling it according to the ability support of actual senders, $[\alpha, 1]$.

In the following a more specific information-processing technology, which has both properties (i) and (ii), is introduced.

[11] The same definition of quantity of information is provided in Moscarini and Smith (2002).

$(A.2)$ Let $\Theta_\alpha = [\alpha, 1]$ be the set of possible ability values, where $\alpha \in [0, 1)$. Then, in stage 3 signal outcomes are drawn from the family of distributions $\{F_\alpha (\cdot \mid \theta, m)\}$, where

$$F_\alpha (x \mid \theta, m) = I_{[\alpha,1]} (x) (1 - \pi (m)) \left(\frac{x - \alpha}{1 - \alpha} \right) + I_{[\theta,1]} (x) \pi (m) + I_{(1,\infty)} (x) \quad (3)$$

and where $\pi (\cdot)$ is a twice differentiable function such that $\pi (1) \equiv \pi_1 \leq 1$, $\pi' (\cdot) < 0$, $\pi'' (\cdot) > 0$ and $\lim_{m \to \infty} \pi (m) = 0$.

In the following pictures the distribution $F_\alpha (\cdot \mid \theta, m)$ and the corresponding probability $f_\alpha (\cdot \mid \theta, m)$ are depicted for the case in which $\alpha = 0$.

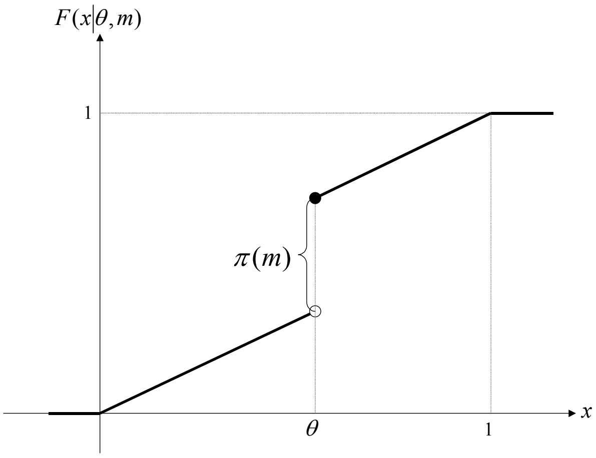

Figure 1(a)

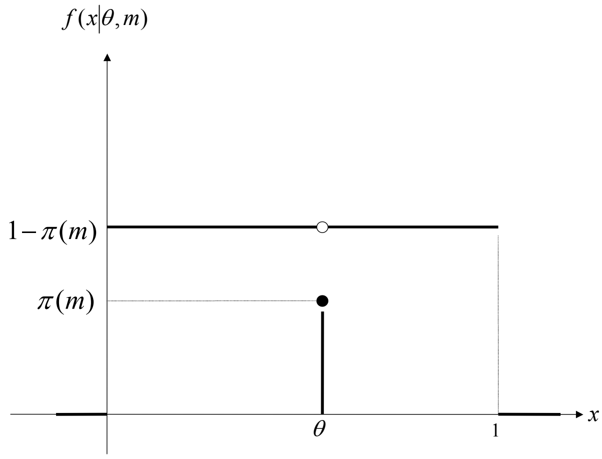

Figure 1(b)

$F_\alpha (\cdot \mid \theta, m)$ is a mixture distribution: the distribution of the signal conditional on ability θ is, with probability $\pi (m)$, a degenerate distribution with all probability mass at $x = \theta$, and, with probability $1 - \pi (m)$, a uniform distribution on $[\alpha, 1]$. Thus, $\pi (\cdot) \in (0, \pi_1]$ is the revealing probability and denotes the accuracy of each signal in predicting the unknown ability: the higher the value of $\pi (\cdot)$ the higher the accuracy, and $\pi (\cdot) \to 0$ captures the situation in which signals become completely uninformative. Since $\pi (\cdot)$ is decreasing in m, eventually approaching zero, the signals' accuracy, as well as the MLRP, vanishes with the number of signals that the Receiver processes. To see it, notice that, if $m \to \infty$, then $\pi (m) \to 0$ and $f_\alpha (x \mid \theta, m) \sim I_{[\alpha,1]} (x) \frac{1}{1-\alpha}$ for any θ. In other words, as $m \to \infty$, all signals become *equivalent* and *neutral* to the decision maker. It is the monotonic behavior of $\pi (\cdot)$ which captures the *trade-off between the quality and the quantity of information.* The fact that the information-processing is *not sensitive to the ability of actual senders* is captured by the fact that the only effect of α on the distribution $F_\alpha (\cdot \mid \cdot)$ is that of rescaling it according to the ability support of actual senders, $[\alpha, 1]$. The implications of such property will be discussed at the end of section 3 and section 4. In the remaining of the paper it is assumed that the information-processing technology is the one defined by $(A.2)$.

3 The problem of the Receiver.

Since the payoff of the Receiver depends on the support of actual senders, $\Theta_\alpha = [\alpha, 1]$, henceforth, it will be denoted by $u_\alpha^r(\hat{x}, m) \equiv E[\theta \mid \hat{x}, m, \theta \geq \alpha]$. The following lemma characterizes the functional form of the terminal payoff of the Receiver.

Lemma 3.1 *The payoff of the Receiver is*

$$u_\alpha^r(\hat{x}, m) = \pi(m)\hat{x} + (1 - \pi(m))\left(\frac{\alpha + 1}{2}\right) \tag{4}$$

The interpretation of (4) is straightforward. The information perspective of the Receiver is a mirror image of that of the sender, as he observes x but not θ, moreover, $\pi(\cdot)$ and $1 - \pi(\cdot)$ represent respectively the probability that a signal is fully informative and the probability that a signal is completely uninformative. Therefore, the Receiver perceives the ability of a sender as being equal to the corresponding signal outcome, with probability $\pi(\cdot)$, and as being equal to the prior average ability of actual senders with probability $1-\pi(\cdot)$. Notice also that (4), considered as a function of signal x, is a straight line with slope $\pi(m)$. Thus, the more signals are processed, the flatter is $u^r(\hat{x}, m)$; this behavior captures the fact that the informativeness of each signal decreases as more signals are simultaneously processed. Finally, the fact that $\lim_{m\to\infty} u^r(\hat{x}, m) = (\alpha + 1)/2$ is the feedback on the Receiver's side of the gradual degeneration of the MLRP. In the following the optimal strategy of the Receiver is determined.

Proposition 3.1 *The dominant strategy for the Receiver is* $\{m^*(n), \hat{x}^*\}$, *where* $\hat{x}^* = \max\{\hat{x}_1, .., \hat{x}_m\}$ *and* $m^*(n) = \min\{n, \bar{m}\}$, *with* $\bar{m} \in (1, \infty)$. *Moreover,* $\bar{m}$ *is unique and is given by the following condition*

$$\frac{|\pi'(\bar{m})|}{\pi(\bar{m})} = \frac{2}{(\bar{m}^2 - 1)} \tag{5}$$

I will provide a sketch of the proof of proposition 3.1, and some general remarks. The Receiver solves the following maximization problem

$$\max_{m,\hat{x}} u_\alpha^r(\hat{x}, m)$$

$$\text{subject to} \quad : \begin{cases} m \leq n \\ \hat{x} \in \hat{X}_m \end{cases}$$

He chooses the sample size m in stage 2 and selects one out of m signal outcomes in stage 4. Consider first stage 4. Since $u_\alpha^r(\hat{x}, m)$ is strictly increasing in $\hat{x}$ (this is a direct consequence of the MLRP), no matters how many (finite) signals the Receiver processes, he always selects the sender whose application produced the highest signal outcome. This proves $\hat{x}^* = \max\{\hat{x}_1, .., \hat{x}_m\}$. Denote by $\bar{x}_{(m)} \equiv E_x[\max\{x_1, ..x_m\}]$

the expected value of the maximum of m signals, and consider now stage 2, when the receiver chooses m in order to maximize the ex-ante expected utility

$$\begin{aligned} E_x\left[u_\alpha^r\left(x^*, m\right)\right] &= u_\alpha^r\left(\bar{x}_{(m)}, m\right) \\ &= \pi\left(m\right)\bar{x}_{(m)} + \left(1-\pi\left(m\right)\right)\left(\frac{\alpha+1}{2}\right) \end{aligned} \tag{6}$$

Notice first that, $u_\alpha^r\left(\bar{x}_{(1)}, 1\right) = u_\alpha^r\left(\left(\alpha+1\right)/2, 1\right) = \left(\alpha+1\right)/2$, that is, processing only one signal implies selecting one sender randomly and, therefore, getting ex-ante payoff equal to the prior expected ability of actual senders. Moreover, $u_\alpha^r\left(\bar{x}_{(m)}, m\right)$ approaches $\left(\alpha+1\right)/2$ as $m \to \infty$; this is the case because, when infinitely many signals are processed, each signal becomes completely uninformative. This implies that the ideal number of signals to process, namely $\bar{m}$, is always finite. Moreover, the well-behavior of function $\pi\left(\cdot\right)$ ensures that such number is unique and I will now provide the intuition behind its characterization. The function $u_\alpha^r\left(\bar{x}_{(m)}, m\right)$ captures the trade-off between quality and quantity of information that the Receiver faces when he has to choose m. The argument $\bar{x}_{(m)}$ increases in m, that is, the more signals are processed the higher is the chance of observing a very high signal. This is the *sample size* effect which has, ceteris paribus, a positive impact on the terminal payoff. However, the Receiver can increase the chance of observing a high signal only at the cost of reducing its quality. This is the detrimental effect of *decreasing accuracy* which is captured by the fact that $u_\alpha^r\left(\cdot, m\right)$ becomes flatter the larger is m (see fig. 2(a)). Therefore the ideal number of signals to process, $\bar{m}$, is determined by the balancement of the two effects. Condition (5), which characterizes $\bar{m}$, has a clear cost benefit interpretation since it is obtained by equating the marginal benefit of processing one additional application to the marginal cost of doing so. Notice from (6) that

$$\begin{aligned} MB\left(m\right) &= \pi\left(m\right)\frac{d\left(\bar{x}_{(m)}\right)}{dm} \\ MC\left(m\right) &= \left|\pi'\left(m\right)\right|\left(\bar{x}_{(m)} - \frac{\alpha+1}{2}\right) \end{aligned}$$

that is: the marginal benefit of processing one additional application is the marginal increase of the expected value of the maximum signal, weighted by the accuracy of the signal; the marginal cost of processing one additional application is the (positive) difference between the expected maximum signal and the prior average ability, weighted by the marginal decrease of the signal's accuracy. From the discussion above it follows that $u_\alpha^r\left(\bar{x}_{(m)}, m\right)$, considered as a function of m, is bell-shaped and reaches its unique maximum at $\bar{m} \in (0, \infty)$ (see figure 2(b)).

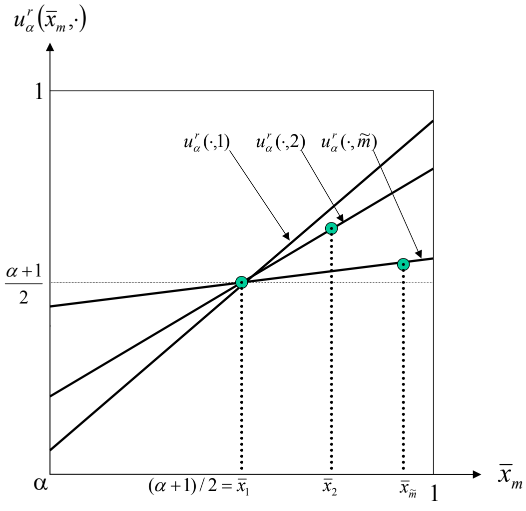

Figure 2(a)

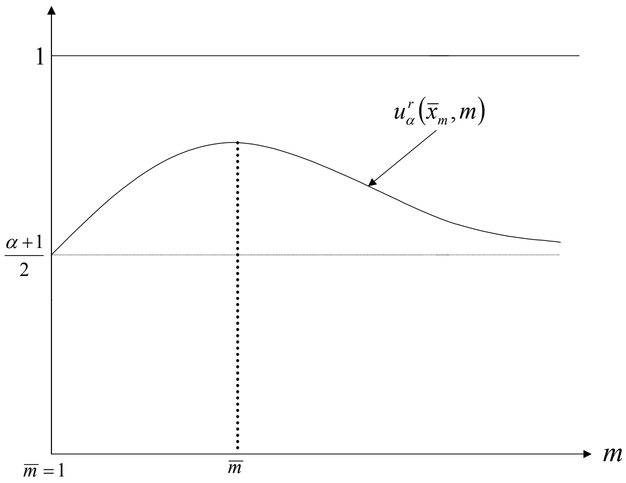

Figure 2(b)

Notice that $\bar{m}$ denotes the number of applications that the Receiver is, at most, willing to process. This implies that, if the number of applications received, n, is larger than $\bar{m}$ the Receiver is overloaded, as $n - \bar{m}$ applications are simply neglected as a result of the maximizing behavior.

Two aspects of proposition 3.1 are worth noticing. First, the optimal behavior of the Receiver does not depend on α: this is a direct consequence of the fact that the *information-processing technology is not sensitive to the ability of actual senders*. Second, $\bar{m}$ is fully determined by condition (5), which follows directly from assumption $(A.2)$. However, it is important to stress that, in order to allow for the possibility of information overload, assumption $(A.1)$ is sufficient. To see it notice that as m becomes large the Receiver views all signal as *equivalent* and *neutral*, and therefore each signal yields expected utility equal to the prior average ability. Consequently also under $(A.1)$ the optimal number of signals to process is finite. Therefore the only convenience of assumption $(A.2)$ is that it allows to have a closed-form solution to the maximization problem of the Receiver which does not depend on α and which has a clear cost-benefit interpretation.

4 The problem of the sender.

If $\Theta_\alpha = [\alpha, 1]$, when n senders apply and m applications are processed, the expected utility of an arbitrary sender $\theta \in [\alpha, 1]$ is

$$u_\alpha^s(\theta, n, m) = \Phi_\alpha(\theta \mid n, m) V - C$$

where $\Phi_\alpha(\theta \mid n, m) \equiv \Phi(\theta \mid n, m, \theta \geq \alpha)$. Denote by $c = C/V$ the cost of applying relative to the value of the vacancy, then sender θ finds it optimal to send if and only if $\Phi_\alpha(\theta \mid n, m) \geq c$. In order to obtain the vacancy, sender θ must go through a two-steps procedure. First, the application of θ must be taken into account by the Receiver, second, provided that this is the case, the signal outcome of θ must be larger than the outcomes of all other applications processed. It is therefore useful to express $\Phi_\alpha(\theta \mid n, m)$ in the following way

$$\Phi_\alpha(\theta \mid n, m) = \gamma(n, m)\, \phi_\alpha(\theta, m) \tag{7}$$

where $\gamma(n, m)$ denotes the probability that one's application is taken into account in the first place when n senders apply and m applications are actually processed, and $\phi_\alpha(\theta, m)$ is the probability that the signal outcome of sender θ is larger than the signal outcomes of the other $m-1$ senders processed. Trivially,

$$\gamma(n, m) = \begin{cases} 1 & \text{if} \quad n \leq m \\ m/n & \text{if} \quad n > m \end{cases} \tag{8}$$

while the nature of $\phi_\alpha(\theta, m)$, is made clear by the following lemma

Lemma 4.1 *If $m \geq 1$ senders are processed by the Receiver, the probability that the signal outcome of sender $\theta \in [\alpha, 1]$ is the largest one is*

$$\phi_\alpha(\theta, m) = \pi(m) \left(\frac{\theta - \alpha}{1 - \alpha} \right)^{m-1} + (1 - \pi(m)) \frac{1}{m} \tag{9}$$

Expression (9) has a clear analogy to expression (4). The first term of (9) says that, with probability $\pi(\cdot)$, the signal is a true representation of ability and, therefore the probability that sender θ assigns to the event of his signal being larger than that of the other $m-1$ senders is simply the $\Pr(\theta \geq \theta_1; ..; \theta \geq \theta_{m-1})$. The second term of (9) states that, with probability $1 - \pi(\cdot)$, each signal, independently of the underlying ability, is uniformly distributed on $[\alpha, 1]$; thus, each signal has equal chance of being larger than the other. More importantly, from the expression of $\phi_\alpha(\theta, m)$ it is also clear what is the effect of the decreasing signals' accuracy on the senders' side: as $m \to \infty$, $\phi_\alpha(\theta, m)$ converges in probability to $1/m$, thus, for large samples, the probability of having the largest signal outcome is purely determined by chance. It is important to stress that this phenomenon also holds under the weaker assumption $(A.1)$. If the Receiver allocates his finite resources on infinitely many signals, all signals become *equivalent* (as the MLRP vanishes). Consequently, the Receiver is not able to discriminate between different signal outcomes and, thus, he randomly selects one of them.

One last property of $\phi_\alpha(\theta, m)$ is worth noticing. Consider the transformation $g(\theta, \alpha) = \alpha + (1-\alpha)\theta$ whose role is that of mapping any $\theta \in [0, 1]$ to a new support, $[\alpha, 1]$, by keeping the relative location of θ fixed. For example, consider the location of the sender with the lowest ability on $[0, 1]$, that is $\theta = 0$; then plugging $\theta = 0$ into $g(\cdot, \alpha)$ yields the location of the sender with the lowest ability on the new support

$[\alpha, 1]$, that is, $g(0, \alpha) = \alpha$. Similarly, consider the location of the median sender when the actual ability support is $[0, 1]$, that is $\theta = 1/2$; then plugging $\theta = 1/2$ into $g(\cdot, \alpha)$ yields $g\left(\frac{1}{2}, \alpha\right) = (\alpha + 1)/2$, which is the location of the median sender when the actual ability support is $[\alpha, 1]$. Notice form (9) that it is always the case that $\phi_{\alpha=0}(\theta \mid m) = \phi_\alpha(g(\theta, \alpha) \mid m)$. Such property is a direct implication of the fact that the *information-processing technology is not sensitive to the ability of applying senders* and, basically, states that, for a given m, the probability that a sender assigns to the event of his signal outcome being larger than that of an arbitrary opponent, depends only on his relative location on the ability support and not on the particular support chosen.

5 Equilibrium analysis.

In the first part of this section I will characterize the equilibria that emerge for specific values of the primitives of the model. In the second part I will provide a comparative static analysis. In order to be able to compare equilibria for different levels of the information-processing technology I assume a specific functional form of $\pi(\cdot)$ which depends also on a technology parameter k.

Assumption $(A.3)$

$$\pi(m, k) = \frac{k}{k+m}, \text{ where } k \in (0, \infty) \tag{10}$$

The parameter k denotes how good is the information-processing technology, and higher values of k imply a better technology. The main convenience of (10) is that k can be interpreted as the amount of resources available to process information, and different values of k allow to capture the entire range of possible information technology levels. For any m, $\lim_{k\to 0} \pi(m, k) = 0$, thus, if no resources are available, the information-processing technology is completely useless. On the contrary, $\lim_{k\to\infty} \pi(m, k) \to 1$ for any m, thus, if infinitely many resources are at disposal, the technology is perfect. Moreover, the function (10) is increasing in k, meaning that, the more resources available, the more is the information that can be extracted from a fixed number of signals drawn. Now $\pi(\cdot)$ depends also on k, thus, the maximum number of applications that the Receiver is willing to process, $\bar{m}$, will also depend on k and, henceforth, will be denoted by $\bar{m}(k)$. Since a better technology allows to extract more information from the same number of signals it trivially follows that, $\bar{m}'(k) > 0$.

As already noticed, the fact that signals have the MLRP implies that more able senders have higher probability of being selected. Then, since the cost of sending (normalized to the value of the vacancy) c is constant there are two types of equilibria that can emerge: a pooling equilibrium (PE) and a partially separating equilibrium (PSE). The following proposition characterizes the equilibria of the model.

Proposition 5.1 *Let $\pi(\cdot)$ be given by (10). Let be $N > \bar{m}(k)$ and define $\mu(k) \equiv 1/(k + \bar{m}(k))$ and $\rho(k, N) \equiv (\bar{m}(k)/N)\,\mu(k)$. Then:*

$(P.5.1.a)$ *If $c \in [0, \rho(k, N))$, a PE with neglected applications arises.*

$(P.5.1.b)$ *If $c \in [\rho(k, N), \mu(k))$, a unique PSE with neglected applications arises.*

$(P.5.1.c)$ *If $c \in [\mu(k), 1/(1+k))$, a unique PSE without neglected applications arises.*

$(P.5.1.e)$ *If $c \in [1/(1+k), 1]$ only one sender with expected ability $1 - 2/N$ applies.*

First, the inequality $N > \bar{m}(k)$ ensures that the market size is large enough for information overload to be a potential problem. If, on the contrary, it was $N \leq \bar{m}(k)$ then, even if the entire mass of senders, N, applies, the information processing technology would never be at full capacity and an equilibrium with information overload would never emerge.

The proposition has a clear graphical interpretation. Let α be the ability threshold level of a PSE. For any PSE, that is, for any threshold level α, it is possible to calculate (see proof of proposition 5.1) the probability that a sender located at α assigns to the event of being selected in equilibrium. Such probability depends on the primitive of the model N and k and is denoted by $\Psi(\alpha, N, k)$. The qualitative behavior of $\Psi(\cdot, N, k)$, where $N > \bar{m}(k)$, is shown in figure 3

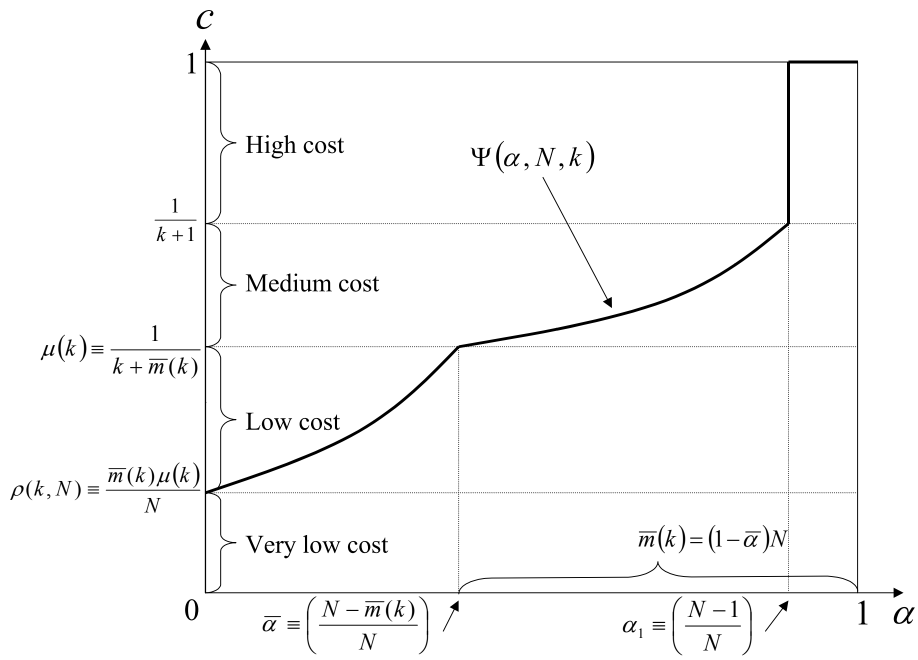

Figure 3

The strict monotonicity of $\Psi(\cdot, N, k)$ is crucial as it implies that a PSE for which the threshold level is α is determined by the condition $\Psi(\alpha, N, k) = c$, while the PE is determined by the condition $\Psi(0, N, k) > c$. Thus, the function $\Psi(\cdot, N, k)$ defines the cost regions for which qualitatively different equilibria emerge and, for any threshold level α, the number of actual senders endogenously determined in equilibrium is given by $n = (1-\alpha) N$. Then, notice that $\bar{\alpha}$ is the threshold level for which the number of senders applying in equilibrium equals the maximum number of applications that the Receiver is willing to process; indeed, $(1-\bar{\alpha}) N = \bar{m}(k)$. Therefore, $\mu(k) \equiv 1/(k + \bar{m}(k))$ is the probability that the sender located at $\bar{\alpha}$ assigns to the event of being selected in a PSE for which the threshold level is actually $\bar{\alpha}$. Indeed,

$$\begin{aligned}\Psi(\bar{\alpha}, N, k) &= \phi_{\bar{\alpha}}(\bar{\alpha}, \bar{m}(k)) \\ &= \frac{1-\pi(\bar{m}(k), k)}{\bar{m}(k)} = \frac{1}{k+\bar{m}(k)}\end{aligned}$$

Similarly, $\rho(k, N) \equiv (\bar{m}(k)/N)\mu(k)$ is the probability that the sender located at $\alpha = 0$ is selected in a pooling equilibrium as

$$\begin{aligned}\Psi(0, N, k) &= \frac{\bar{m}(k)}{N}\phi_0(0, \bar{m}(k)) \\ &= \frac{\bar{m}(k)}{N}\frac{1-\pi(\bar{m}(k), k)}{\bar{m}(k)} = (\bar{m}(k)/N)\mu(k)\end{aligned}$$

Finally, since $(1-\alpha_1) N = 1$, α_1 denotes the threshold level for which the mass of actual senders in equilibrium is unitary. When this is the case the random matching function ensures that each sender is assigned to one receiver and, therefore, each sender knows that will be selected for sure. This is the reason why, on the one hand $\Psi(\alpha, N, k) = 1$ for any $\alpha \geq \alpha_1$ and, on the other, the Receiver, by selecting the only applicant at his disposal, gets expected utility $(\alpha_1 + 1)/2 = 1 - 2/N$.

The following corollary follows directly form proposition 5.1.

Corollary 5.1 *For any $k \in (0, \infty)$ and $N > \bar{m}(k)$, a market is overloaded iff $c < \mu(k)$. Moreover $\mu'(k) < 0$ and* $\lim_{k\to\infty} \mu(k) = 0$.

This corollary supports the intuition that market congestion emerges whenever the cost of transmitting information is low, relatively to available resources needed to interpret information. The better the technology, the lower the cost must be, in order for information overload to emerge.

I now study how a marginal change in the market size, N, and in the technology level, k, affects the threshold level α and the number of senders applying in equilibrium, n. To avoid trivialities the attention is restricted to the PSE only. In a PE the sender located at zero is strictly better-off by sending, thus a marginal change in the exogenous parameters will not affect his behavior.

Proposition 5.2 *Let α and n be the equilibrium threshold level and the number of senders applying in a PSE. Then,*

$(P.5.2.a)$ $\partial\alpha/\partial k > 0$ *and* $\partial n/\partial k < 0$.

$(P.5.2.b)$ $\partial\alpha/\partial N > 0$ *and* $\partial n/\partial N = 0$.

The intuition behind $(P.5.2.b)$ is not surprising: a better technology allows the Receiver to rank applicants' quality more precisely, which discourages low-quality employees from applying. On the contrary, the result of $(P.5.2.a)$ is very interesting. It states that an increase in the market size does not translate in a larger number of actual applicants but only in higher applicants' ability. This is possible because competition discourages low ability employees from applying and thus, in equilibrium, only the better workers, who have more chances to be hired, are willing to bear the cost of sending an application. Therefore, when the market equilibrium implies separation, the self selection mechanism of applicants competing for a position is fully preserved. The following pictures show how the cost regions for which different equilibria arise change in k and N.

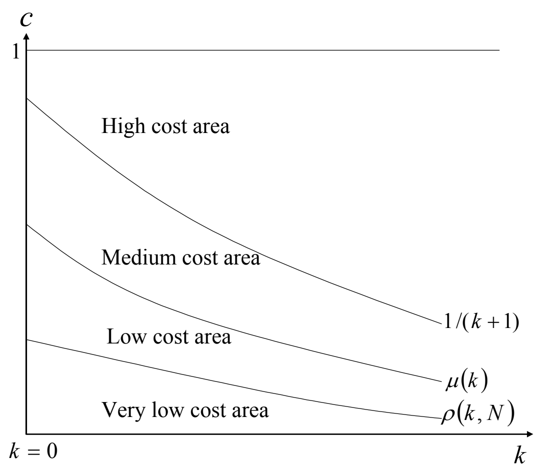

Figure 4(a)

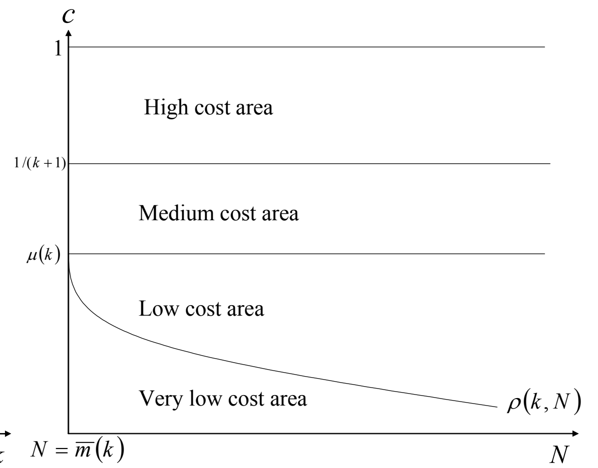

Figure 4(b)

On one hand, when the technology improves all cost regions, except the high-cost one, shrinks. Thus, for any sending cost c there exists a technology level that allows the Receiver to screen the best sender. On the other hand, a larger market size causes the very-low cost area to shrink and the low-cost to get larger. Thus, even for extremely low (but positive) sending cost, there exists a sufficiently large market size which ensures that the self selection mechanism plays a role. The reason is that, even though c lies initially in the very-low-cost area, which implies no separation, it will eventually lie in the low-cost area and, then, separation will occur.

The self selection mechanism has two very important implications on the functioning of the market. The first is that, from the Receiver point of view, a larger market is always better because the senders applying in equilibrium are of better quality. The second implication plays a role in determining the amount of information that the market generates, and it is addressed in the next section.

6 Market congestion and efficiency

In this section the analysis focuses on markets that suffer from information congestion, that is, markets where the cost of transmitting information is low relative to the information-processing technology. Implicit here is the idea that the cost at which information can be transmitted is not a strategic variable, but is exogenously given by the environment. If, on the contrary, the Receiver could use the cost as a screening device (e.g. the Receiver could determine the sending cost by imposing an application fee) it would put it high enough to ensure that only the best sender finds worthwhile applying.

So far I have assume that the Receiver can neglect applications. Clearly this assumption does not capture the reality of many monopsony markets[12]. However, the model I propose can be easily accommodated to represent a situation in which neglecting applications is not an option. This can be done by restricting the strategy space of the Receiver to be simply that of selecting one applications from those received. The resulting model is much simpler as it implies that the Receiver maximizes expected utility $u_{\alpha}^{r}(\hat{x}, n)$ only with respect to $\hat{x}$, and that, the probability that a sender assigns to the event of being selected depends only on the total number of actual senders, $\phi_{\alpha}(\theta, n)$. Therefore, in this section, I will define and compare the economic cost of market congestion for two different scenarios: one in which applications can be neglected, and one in which this is not possible.

The first question I address here is how the amount of information generated, n, depends on the market size, N.

Proposition 6.1 *Let be $N > \bar{m}(k)$. For any $k \in (0, \infty)$ and $c \in (0, \mu(k))$ the market is overloaded and*

$(P.6.1.a)$ *if the Receiver can neglect applications, the excessive amount of information is*

$$n(N) = \begin{cases} N & \text{if } \bar{m}(k) < N \leq \frac{1}{c}(1 - k\mu(k)) \\ \frac{1}{c}(1 - k\mu(k)) & \text{if } \frac{1}{c}(1 - k\mu(k)) < N < \infty \end{cases} \tag{11}$$

$(P.6.1.b)$ *if the Receiver cannot neglect applications, the excessive amount of information is*

$$n(N) = \begin{cases} N & \text{if } \bar{m}(k) < N \leq \frac{1}{c}(1 - kc) \\ \frac{1}{c}(1 - kc) & \text{if } \frac{1}{c}(1 - kc) < N < \infty \end{cases} \tag{12}$$

First of all, the proposition states that, irrespective of the fact that the Receiver is allowed or not allowed to neglect applications, the information generated is not

[12] For instance, there are many situations in which the agent in charge of processing applications is legally constraint to pay attention to all applicants. For example scientific journals read all the submissions they receive.

strictly increasing in the market size. On the contrary, it is bounded, and its maximum level depends on c and k. It is the self selection mechanism that is responsible for this result. A positive cost on the supply side identifies a hard bound on the amount of information produced by the market: for N sufficiently large, separation occurs, consequently the maximum number of applicants in equilibrium is constant and depends only on c and k. Notice also that the technology level identifies a hard bound (that is $\bar{m}(k)$) on the maximum amount of information that the market is willing to absorb. When $c < \mu(k)$, the former bound is larger than the latter, thus, the market is structurally subject to information congestion, the extent of which, is also bounded. Therefore the maximum magnitude of information overload, is $\frac{1}{c}(1 - k\mu(k)) - \bar{m}(k)$ if the receiver can neglect applications, and $\frac{1}{c}(1 - kc) - \bar{m}(k)$ if, on the contrary, applications cannot be neglected.

Since $c < \mu(k)$, from (11) and (12) it follows that the extent of market congestion is larger when the Receiver cannot neglect applications. Two effects play a role in shaping this result. First, if some applications are neglected in equilibrium, senders anticipate that there is a chance that their application might not be taken into account in the first place. Second, when applications are neglected, the signals' accuracy is larger compared to the case in which all excessive applications are processed, and this helps the Receiver to sort out relatively bad candidates. Both effects have a discouraging impact on senders and, therefore, a market in which neglecting applicants is an option induces less people to apply.

I now investigate what are the economic costs of information congestion. In an overloaded market in which some applications are neglected the inefficiency of information congestion arises only on the supply side. The Receiver, by ignoring some information, protects himself from the decreasing accuracy effect. However, since sending applications is costly, a market equilibrium with information overload implies too high sending costs for the society. Consider now an overloaded market in which all information generated must also be processed. Here, the side effects of information congestion are more severe and affect both the demand and the supply side. On one hand, the excessive number of applications processed decreases the utility of the Receiver[13]. On the other hand, the inefficiency is also present on the senders side as too many senders bear the cost of applying. Moreover, in a scenario in which applications cannot be neglected, the magnitude of the overload is larger compared to the situation in which ignoring applicants is an option. Therefore, here, the sending cost for the society are even higher. The next lemma is needed to support what follows.

Lemma 6.1 *Let be* $\Pi(k) = \pi(\bar{m}(k), k)$. *Then* $\Pi'(k) > 0$.

I now consider the total sending costs for the society. If applications can be

[13]This is not true if the market size is very large. Indeed, the self-selection mechanism implies that, if $N \to \infty$, the threshold level α approaches 1, and, thus, irrespectively on the number of applications processed, the utility of the receiver also approaches 1.

neglected and the market is overloaded to its maximum extent, the total sending cost for the society is

$$
\begin{aligned}
TC(c,k) &= c\left[\frac{1}{c}(1-k\mu(k))\right] \\
&= (1-k\mu(k)) = \left(1-\frac{k}{k+\bar{m}(k)}\right) = 1-\Pi(k) \qquad (13)
\end{aligned}
$$

If, on the contrary, applications cannot be neglected, the sending cost for the society is

$$
TC(c,k) = c\left[\frac{1}{c}(1-kc)\right] = 1/c-k \qquad (14)
$$

From (13) and (14) it follows that a marginal increase in the technology level always decreases the total sending cost. On one hand, a better technology implies that the Receiver can assess applicants ability more precisely, thus low ability applicants are discouraged from applying and the number of actual senders in equilibrium decreases. On the other hand, it also increases the maximum number of applications that the Receiver is willing to process. Consequently the extent of the overload decreases and, with it, also the total sending cost. Interestingly enough, in an equilibrium in which applications in excess are neglected, the total sending cost does not depend on the cost of sending one single application. The reason for this surprising result is that the elasticity of the number of actual senders with respect to the cost of sending one single application is always unitary: as the cost of sending one application increases the number of applicants in equilibrium decreases in the same proportion and therefore the total cost does not change. If, on the other hand, neglecting applications is not an option, the total sending costs are decreasing in the cost of sending one single application. Thus, in this second scenario, the elasticity of the number of actual senders with respect to the cost of sending one single application is negative and larger than one in absolute value.

7 Conclusion.

The evidence of everyday life shows that there are many situations in which people receive a too large amount of information compared to the one they are actually willing to process. In this paper I address the information overload issue in the specific situation in which many applicants compete to obtain one position and the employer screens among applications in order to select the best applicant. With this set up information overload can be defined as an equilibrium outcome in which some applications are neglected by the economic agent in charge of screening applicants. It has been shown that a large market is not directly responsible for market congestion. On the contrary, more competition has a beneficial effect because the self-selection mechanism is preserved. The results I obtain show that information overload occurs when the application cost is low relative to the information-processing technology level. This supports the intuitive idea that information overload is more likely to be present in environments in which the cost of transmitting information is low, while there are few resources to interpret information. I also note that in some instances neglecting applications is not possible (e.g. because of a legal constraint) and, therefore, I also consider a scenario in which all the information received must be processed. It turns out that the possibility of neglecting excessive information, to some extent, decreases the inefficiency of market congestion.

The general set up of the model is very natural and borrows standard concepts from the literature on economics of information. The crucial ingredient of the model is the assumption that when infinitely many applications are processed, then each application does not provide any information about the ability of the corresponding applicant. This is modelled by imposing that the family of density functions from which signal outcomes are drawn has the MLRP and that the MLRP vanishes as the number of signals drawn becomes large. The more specific information-processing technology that I use displays a clear trade-off between the quality and the quantity of information and is convenient for the possibility of obtaining closed form solutions with straightforward economic interpretations. More specifically, it allows to interpret neglecting applications as a consequence of maximizing behavior where the marginal utility of a larger sample size (which is the higher chance of observing a very good application) is compared to its opportunity cost (which is the marginal decrease in the capability of ranking the applicants according to their true abilities). Yet, the main results are robust and hold for very general information-processing technologies.

References

[1] **Akerlof, G.** (1970). The market of lemons: Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, **89**, 488-500.

[2] **Boal, W. M. and Ransom, M. R.** (1997). Monopsony in the Labor Market. *Journal of Economic Literature*, **35**, 86-112.

[3] **Cho, I.-K. and Kreps, D.** (1987). 'Signalling and Stable Equilibria', *Quarterly Journal of Economics*, **102**, 179-221.

[4] **Ficco, S. and Karamychev, V.** (2004) . 'Information Overload in Multi-Stage Selection Procedures', Tinbergen Institute Discussion Paper series.

[5] **Hahn, M., Lawson, R. and Young, G.** (1992). 'The effect of Time Pressure and Information Load on Decision Quality', *Psychology and Marketing*, **9,** 365-379.

[6] **Jacoby, J., Speller, D. E. and Kohn, C. A.** (1974). 'Brand Choice Behavior as a Function of Information Load'. *Journal of Marketing Research*, **11**, 63-69.

[7] **Janssen, M.C.W.** (2002) . 'Catching Hipos: Screening, Wages, and Competing for a Job', *Oxford Economic Papers,* **54 (2),** 321-*33.*

[8] **Lipman, B. L.** (1995). 'Information Processing and Bounded Rationality: A Survey', *The Canadian Journal of Economics*, **28,** 42-67.

[9] **Milgrom, P. R.** (1981). 'Good News and Bad News: Representation Theorems and Applications', *The Bell Journal of Economics*, **12**, 380-391.

[10] **Miller, G. A.** (1956). 'The Magical number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information', *Psychological Review*, **63**, 81-97.

[11] **Moscarini, G. and Smith, L.** (2001). 'The Optimal Level of Experimentation', *Econometrica*, **69**, 1629-1644.

[12] **Moscarini, G. and Smith, L.** (2002). 'The Law of Large Demand for Information', *Econometrica*, **70**, 2351-2366.

[13] **Nöldeke, G. and van Damme, E.** (1990) . 'Signalling in a Dynamic Labour Market', *Review of Economic Studies,* **57,** 1-23.

[14] **Spence, A. M.** (1973) . 'Job Market Signalling', *Quarterly Journal of Economics,* **90,** 225-43.

[15] **Weiss, A.** (1995) . 'Human Capital vs. Signalling Explanations of Wages', *Journal of Economic Perspectives,* **9,** *133-153.*

[16] **Zandt van, T.** (2001) . Information Overload in a Network of Targeted Communication', *2 May 2001,* available at http://zandtwerk.insead.fr/papers/Overload.pdf.

8 Appendix

8.0.1 Proof of lemma 3.1.

For notation simplicity, the subscript α of the distribution (3) will be omitted. I first show that $f(x) = I_{[\alpha,1]}(x)\frac{1}{1-\alpha}$. Since $F(\cdot \mid \theta, m)$ has a discrete jump at $x = \theta$, the difference $F(\theta \mid \theta, m) - \lim_{x\to\theta^-} F(x \mid \theta, m) = \pi(m)$ is the probability mass assigned to the event $x = \theta$, thus

$$\begin{aligned} f(x \mid \theta, m) &= I_{\{\theta\}}(x)\,\pi(m) + \frac{d}{dx}F(x \mid \theta, m) \\ &= \pi(m)\,I_{\{\theta\}}(x) + (1-\pi(m))\,I_{\{[\alpha,1]\setminus\{\theta\}\}}(x)\,\frac{1}{1-\alpha} \end{aligned}$$

Now

$$\begin{aligned} f(x \mid m) &= \int_{-\infty}^{+\infty} f(x \mid \theta, m)\,f(\theta)\,d\theta, \\ &= \frac{1}{1-\alpha}\int_{-\infty}^{+\infty} f(x \mid \theta, m)\,d\theta \\ &= \frac{1}{1-\alpha}\left[\pi(m)\int_{-\infty}^{+\infty} I_{\{\theta\}}(x)\,d\theta + (1-\pi(m))\int_{-\infty}^{+\infty} I_{\{[\alpha,1]\setminus\{\theta\}\}}(x)\,\frac{1}{1-\alpha}d\theta\right] \\ &= \frac{\pi(m)}{1-\alpha}\int_{-\infty}^{+\infty} I_{\{\theta\}}(x)\,d\theta + \frac{1-\pi(m)}{1-\alpha}\int_{\alpha}^{1}\frac{1}{1-\alpha}d\theta \\ &= \frac{\pi(m)}{1-\alpha}\int_{-\infty}^{+\infty} I_{\{\theta\}}(x)\,d\theta + \frac{1-\pi(m)}{1-\alpha} \end{aligned}$$

Since $I_{\{\theta\}}(x) = 1$ when $\theta = x$, and since θ is uniformly distributed on $[\alpha, 1]$, $\int_{-\infty}^{+\infty} I_{\{\theta\}}(x)\,d\theta = \int_{\alpha}^{1}\frac{1}{1-\alpha}d\theta = 1$. It then follows that $f(x \mid m) = f(x) = I_{[\alpha,1]}(x)\frac{1}{1-\alpha}$. Now, $f(\theta \mid x, m) = \frac{f(\theta)f(x|\theta,m)}{f(x)}$ and, since $f(x) = f(\theta)$, we have $f(\theta \mid x, m) = f(x \mid \theta, m)$ and therefore

$$f(\theta \mid x, m) = \pi(m)\,I_{\{x\}}(\theta) + (1-\pi(m))\,I_{\{[\alpha,1]\setminus\{x\}\}}(\theta)\,\frac{1}{1-\alpha}$$

It then trivially follows that

$$\begin{aligned} E_\alpha[\theta \mid x, m] &= \pi(m)\,x + (1-\pi(m))\int_{-\infty}^{+\infty} I_{\{[\alpha,1]\setminus\{x\}\}}(\theta)\,\frac{\theta}{1-\alpha}d\theta \\ &= \pi(m)\,x + (1-\pi(m))\int_{\alpha}^{1}\frac{\theta}{1-\alpha}d\theta \\ &= \pi(m)\,x + (1-\pi(m))\left(\frac{\alpha+1}{2}\right) \end{aligned}$$

■

8.0.2 Proof of proposition 3.1.

I will use the following notation.

$$\begin{aligned} x_{(m)} &\equiv \max\{x_1,...,x_m\} \\ \bar{x}_{(m)} &\equiv E_x[\max\{x_1,...,x_m\}] \end{aligned}$$

First, the fact that $x^* = x_{(m)}$ is trivial given that $E_\alpha[\theta \mid x, m]$ is strictly increasing in x. Then, the reduced form payoff in stage 2 is

$$\begin{aligned} E_x\left[u^r\left(x_{(m)}, m\right)\right] &= \pi(m) E_x[\max\{x_1,...,x_m\}] + (1-\pi(m))\left(\frac{\alpha+1}{2}\right) \\ &= \pi(m)\bar{x}_{(m)} + (1-\pi(m))\left(\frac{\alpha+1}{2}\right) \end{aligned}$$

I now show that $\bar{x}_{(m)} = \frac{\alpha+m}{m+1}$. $F_{x_{(m)}}(z) = \Pr(x_1 \le z, .., x_m \le z) = \Pi_{i=1}^m [F_{x_i}(z)]$. Since signal outcomes are i.i.d. with $f(x) = I_{[\alpha,1]}(x)\frac{1}{1-\alpha}$, it follows that $F_{x_{(m)}}(z) = \left(\frac{z-\alpha}{1-a}\right)^m$ and $f_{x_{(m)}}(z) = \frac{m}{1-\alpha}\left(\frac{z-\alpha}{1-a}\right)^{m-1}$. Thus

$$\begin{aligned} \bar{x}_{(m)} &= \int_\alpha^1 f_{x_{(m)}}(z)\, zdz = \int_\alpha^1 \frac{m}{1-\alpha}\left(\frac{z-\alpha}{1-a}\right)^{m-1} zdz \\ &= \frac{\alpha+m}{m+1} \end{aligned}$$

Then, the reduced form pay-off is

$$\begin{aligned} E_x\left[u^r\left(x_{(m)}, m\right)\right] &= E_\alpha\left[\theta \mid \bar{x}_{(m)}, m\right] = \pi(m)\left(\frac{\alpha+m}{m+1}\right) + (1-\pi(m))\left(\frac{\alpha+1}{2}\right) \\ &= \left(\frac{\alpha+1}{2}\right) + \pi(m)\left(\frac{\alpha+m}{m+1} - \frac{\alpha+1}{2}\right) \\ &= \left(\frac{\alpha+1}{2}\right) + \pi(m)\frac{(1-\alpha)(m-1)}{2(m+1)} \end{aligned}$$

Notice that $E_x\left[u^r\left(x_{(1)}, 1\right)\right] = \lim_{m\to\infty} E_x\left[u^r\left(x_{(m)}, m\right)\right] = (\alpha+1)/2$, thus $\bar{m} \in (1, \infty)$. Let be $\gamma(m) = \pi(m)\frac{(1-\alpha)(m-1)}{2(m+1)}$, then

$$\gamma'(m) = \left(\frac{1-\alpha}{2}\right)\left(\frac{\pi'(m)(m-1)}{(m+1)} + \frac{\pi(m)\, 2}{(m+1)^2}\right)$$

which yields the f.o.c.

$$\frac{|\pi'(\bar{m})|}{\pi(\bar{m})} = \frac{2}{\bar{m}^2 - 1} \tag{15}$$

The LHS of (15) captures the marginal cost of a larger sample size, while the RHS captures its marginal benefit. To show that $\bar{m}$ is unique, notice that, around $\bar{m}$, condition (15) can be approximated as follows

$$\begin{aligned}\frac{|\pi'(\bar{m})|}{\pi(\bar{m})+\pi'(\bar{m})(m-\bar{m})} &= \frac{2}{\bar{m}^2-1}\\ \frac{|\pi'(\bar{m})|}{\pi(\bar{m})} &= \frac{2}{(\bar{m}^2-1)+2(m-\bar{m})}\end{aligned}$$

The ratio $|\pi'(m)|/\pi(m)$ decreases in m at the rate of $1/m$, therefore, for any $m>\bar{m}$, the RHS of (13) is always larger then the LHS of (13) and the uniqueness of $\bar{m}$ follows. ■

8.0.3 Proof of lemma 4.1.

For notation simplicity, the subscript α of the distribution (3) will be omitted. Let y be the signal of sender θ, and $\{x_1,..,x_{m-1}\}$ the set of signals of the others $m-1$ senders. Let be $x_{(m-1)} \equiv \max\{x_1,..,x_{m-1}\}$. Recall from proof lemma 3.1 that the signal of an arbitrary sender is uniformly distributed on $[\alpha,1]$, that is,

$$F_x(z)=\frac{z-\alpha}{1-\alpha}$$

In the prof I will make use of the distribution of signals conditional on ability,

$$F_x(z\mid\theta,m)=I_{[\alpha,1]}(x)(1-\pi(m))F_x(z)+I_{[\theta,1]}(x)\pi(m)+I_{(1,\infty)}(x)$$

and the distribution of the maximum between $m-1$ signals coming from arbitrary senders

$$F_{x_{(m-1)}}(z)=[F_x(z)]^{m-1}$$

Then,

$$\begin{aligned}\phi_\alpha(\theta\mid m) &\equiv \Pr\left(x_{(m-1)}\le x\mid m,\theta\ge\alpha\right)\\ &= \int_{-\infty}^{+\infty}(1-F_x(z\mid\theta,m))f_{x_{(m-1)}}(z)\,dz\\ &= 1-\int_{-\infty}^{+\infty}F_x(z\mid\theta,m)f_{x_{(m-1)}}(z)\,dz\\ &= 1-\left((1-\pi(m))\int_\alpha^1 F_x(z)f_{x_{(m-1)}}(z)\,dz+\pi(m)\int_\theta^1 f_{x_{(m-1)}}(z)\,dz\right)\\ &= 1-\left(\frac{(1-\pi(m))(m-1)}{(1-\alpha)^m}\int_\alpha^1(z-\alpha)^{m-1}dz+\frac{\pi(m)(m-1)}{(1-\alpha)^{m-1}}\int_\theta^1(z-\alpha)^{m-2}dz\right)\\ &= 1-\left((1-\pi(m))\left(\frac{m-1}{m}\right)+\pi(m)\left(1-\left(\frac{\theta-\alpha}{1-\alpha}\right)^{m-1}\right)\right)\\ &= \pi(m)\left(\frac{z-\alpha}{1-\alpha}\right)^{m-1}+(1-\pi(m))\frac{1}{m}\end{aligned}$$

■

8.0.4 Proof of proposition 5.1.

Let α be the threshold level of a PSE, let $n = (1-\alpha) N$ be the number of senders applying in equilibrium. let $\Psi(\alpha, N, k)$ be the probability that the sender located at α assigns to the event of being selected in equilibrium. Notice first that, in equilibrium, it must be $n \geq 1$. If $n < 1$, then the mass of applying senders is smaller than the mass of receivers, and some receivers do not receive any application. Therefore provided $n < 1$, some senders have an incentive to apply because they know they will be selected for sure. Recalling that $N > \bar{m}(k)$ and noticing that, by definition, $\Psi(\alpha, N, k) \equiv \Phi_\alpha(\alpha \mid n, m^*(n))$, it follows that

$$\begin{aligned}\Psi(\alpha, N, k) &= \begin{cases} 1 & \text{if } 1 = n \\ \phi_\alpha(\alpha, n) & \text{if } 1 < n \leq \bar{m}(k) \\ \frac{m(k)}{n}\phi_\alpha(\theta, \bar{m}(k)) & \text{if } \bar{m}(k) < n < N \end{cases} \\ &= \begin{cases} 1 & \text{if } \alpha = \frac{N-1}{N} \\ \frac{1}{k+(1-\alpha)N} & \text{if } \frac{N-\bar{m}(k)}{N} \leq \alpha < \frac{N-1}{N} \\ \frac{\bar{m}(k)}{(1-\alpha)N}\frac{1}{k+\bar{m}(k)} & \text{if } 0 \leq \alpha < \frac{N-\bar{m}(k)}{\bar{m}(k)} \end{cases} \end{aligned} \quad (16)$$

Since $\mu(k) \equiv 1/(k+\bar{m}(k))$ and $\rho(k,N) \equiv (m(k)/N)\mu(k)$, the functional form (16) proves the proposition. ■

8.0.5 Proof of corollary 5.1.

The fact that a market is overloaded when $c < \mu(k)$ is trivial and follows directly from (16). I then show that $\mu'(k) < 0$.

$$\mu(k) = 1/(k+\bar{m}(k))$$

$\bar{m}(k)$ is given by the f.o.c. (15)

$$\begin{aligned}\frac{|\pi'(\bar{m}, k)|}{\pi(\bar{m}, k)} &= \frac{2}{\bar{m}^2 - 1} \\ \frac{1}{k+\bar{m}} &= \frac{2}{\bar{m}^2 - 1} \\ &\implies \bar{m}(k) = 1 + (2(k+1))^{1/2} \end{aligned} \quad (17)$$

therefore, $\mu'(k)$ and $\lim_{k\to\infty} \mu(k) = 0$. ■

8.0.6 Proof of proposition 5.2.

Two cases are considered.

Case 1: $n \leq \bar{m}(k)$. The equilibrium condition is

$$\frac{1}{k+(1-\alpha)N} - c = 0 \tag{18}$$

and implicit differentiations on (18) yield $\partial\alpha/\partial k = 1/N > 0$ and $\partial\alpha/\partial N = (1-\alpha)/N > 0$. It then follows that

$$\begin{aligned}
\partial n/\partial k &= -(\partial\alpha/\partial k) N = -1 < 0 \\
\partial n/\partial N &= (1-\alpha) - N(\partial\alpha/\partial N) = (1-\alpha) - N\left(\frac{1-\alpha}{N}\right) = 0
\end{aligned}$$

Case 2: $n > \bar{m}(k)$. The equilibrium condition is

$$\begin{aligned}
\frac{\bar{m}(k)}{(1-\alpha)N(k+\bar{m}(k))} - c &= 0 \\
\frac{\bar{m}(k)\mu(k)}{(1-\alpha)N} - c &= 0
\end{aligned} \tag{19}$$

Recalling (17), and applying implicit differentiations on (19), yields

$$\begin{aligned}
\partial\alpha/\partial k &= -\frac{(1-\alpha)N}{N\bar{m}(k)\mu(k)}\left[\bar{m}'(k)\mu(k) + \bar{m}(k)\mu'(k)\right] \\
&= \frac{(1-\alpha)N}{N\bar{m}(k)\mu(k)}\left[\frac{2+k+(2(k+1))^{1/2}}{(2(k+1))^{1/2}\left(1+k+(2(k+1))^{1/2}\right)^2}\right] > 0 \\
\partial\alpha/\partial N &= \frac{1-\alpha}{N} > 0
\end{aligned}$$

It then follows

$$\begin{aligned}
\partial n/\partial k &= -(\partial\alpha/\partial k) N < 0 \\
\partial n/\partial N &= (1-\alpha) - N(\partial\alpha/\partial N) = (1-\alpha) - N\left(\frac{1-\alpha}{N}\right) = 0
\end{aligned}$$

■

8.0.7 Proof of proposition 6.1.

If the receiver can neglect applications, then $\Psi(\alpha, N, k)$ is given by (16). Since $c < \mu(k)$, and $\bar{m}(k) > N$, if

$$\begin{aligned}
0 &< c < \rho(k, N) \equiv \frac{\bar{m}(k)}{N(k+\bar{m}(k))} \\
N &< \frac{\bar{m}(k)}{c(k+\bar{m}(k))} = \frac{1}{c}\left(1 - \frac{k}{k+m(k)}\right) = \frac{1}{c}(1 - k\mu(k))
\end{aligned}$$

the PE arises and $n = N$. If,

$$\begin{aligned} c &\geq \rho(k, N) \equiv \frac{\bar{m}(k)}{N(k + \bar{m}(k))} > 0 \\ N &\geq \frac{\bar{m}(k)}{c(k + \bar{m}(k))} = \frac{1}{c}\left(1 - \frac{k}{k + m(k)}\right) = \frac{1}{c}(1 - k\mu(k)) \end{aligned}$$

a PSE arises and, from $(P.5.2.b)$, n is constant for any marginal increases in N. Therefore the (bounded) amount of information is given by the number of senders applying in the PSE for which the threshold level is $\alpha = 0$, that is, n such that

$$\begin{aligned} c &= \frac{\bar{m}(k)}{n(k + \bar{m}(k))} \\ n &= \frac{\bar{m}(k)}{c(k + \bar{m}(k))} = \frac{1}{c}(1 - k\mu(k)) \end{aligned}$$

If, on the other hand, the Receiver cannot neglect applications, then $\Psi(\alpha, N, k)$ is simply

$$\begin{aligned} \Psi(\alpha, N, k) &= \begin{cases} 1 & \text{if } \ 1 = n \\ \phi_\alpha(\alpha, n) & \text{if } \ 1 < n \leq N \end{cases} \\ &= \begin{cases} 1 & \text{if } \ \alpha = \frac{N-1}{N} \\ \frac{1}{k + (1-\alpha)N} & \text{if } \ 0 \leq \alpha < \frac{N-1}{N} \end{cases} \end{aligned}$$

Thus, if

$$\begin{aligned} 0 &< c < \frac{1}{k + N} \\ N &< \frac{1}{c}(1 - kc) \end{aligned}$$

the PE arises and $n = N$. If, on the contrary,

$$\begin{aligned} \frac{1}{k + N} &\leq c \\ N &\geq \frac{1}{c}(1 - kc) \end{aligned}$$

a PSE arises and n is constant for any marginal increases in N. Therefore the (bounded) amount of information is given by the number of senders applying in the PSE for which the threshold level is $\alpha = 0$, that is, n such that

$$\begin{aligned} c &= \frac{1}{k + n} \\ n &= \frac{1}{c}(1 - kc) \end{aligned}$$

■

8.0.8 Lemma 6.1.

$$\begin{aligned}\Pi(k) &\equiv \frac{k}{k+\bar{m}(k)} \\ &= \frac{k}{k+1+(2(k+1))^{1/2}}\end{aligned}$$

Thus,

$$\Pi'(k) = \frac{2+k+(2(k+1))^{1/2}}{(2(k+1))^{1/2}\left(1+k+(2(k+1))^{1/2}\right)^2} > 0$$

■