Network and Belief Formation

Francisco Gerónimo Ugarte Bedwell*

October 28, 2009

Abstract

The paper discusses network formation, in a setting where players can minimize the uncertainty about possible states of the world in order to play an optimal strategy according to their beliefs. They do so by connecting to other individuals who then share their information with them. The study is done considering social networks, according to the idea that these are formed in order to minimize the uncertainty individuals have to face. The paper considers the existence of noise in the transmission of information within the network, and analyzes which kind of networks emerge under different attitudes towards uncertainty of individuals. The basic criterion used to analyze formation is pairwise stability of the network. Moreover, it characterizes some criteria to analyze how they relate to the real distribution of the event. This modelling approach will hopefully shed some light on why and how individuals decide to establish some link between them in a social sphere, why societies have a certain architecture according to the attitude towards uncertainty of the individuals that conform it, and why different societies have different beliefs.

*Ph.D in Economics student, Tel Aviv University
1 Introduction

In the past years, the study of networks in economics has shed light on problems that were not solved in a satisfying way by the preexisting literature. Models of network formation have explained the individual decisions of establishing links or not in a society, analyzing the kind of architectures that emerge after connection decisions are made. Several analytical tools have been developed for this study, including solution concepts like stability and efficiency of a network, which are the benchmark for analyzing formation. There are several theories in diverse areas of the social sciences dealing with this question, and one of these considers individuals who wish to obtain information about the world they live in as the main reason for establishing a social link with other people. This idea constitutes the main motivation of this paper and will be explained in further detail when reviewing the existing literature.

The paper introduces case-based decision theory into a networks framework through a model of social links formation in which the behavior of each individual depends on her relative attitude towards uncertainty. There is uncertainty concerning the occurrence of states of the world and there are individuals who infer a probability distribution for them from cases available to them. They have the possibility of connecting to other individuals in order to minimize their uncertainty by acquiring the information others possess in the form of cases. Moreover, the information of people being farther away in the social sphere of an individual are weighted less in the formation of an ex-post belief. For the formation stage, the connection decision of the agent will depend on her degree of uncertainty aversion and on the number of cases her potential connections have observed, with utility functions varying across types to reflect their different attitudes towards uncertainty.

Consider an individual $j$’s decision of establishing a connection or not in a network. Let us assume that once it is formed, she is far away in the network from individual $i$. This is equivalent to saying that $j$ almost does not have any information about the cases of $i$ (or no information at all, if there is no path in the network linking $i$ to $j$), or that $j$ is socially distant enough from $i$ so as to make $j$ not ponder the information of $i$ in a significant way. Individual $j$’s information would be different if she decides to connect to $i$. Throughout the
paper, individuals do not possess any information on cases observed by others until they connect, and hence the connection decision does not depend on it. The trade-off between potential gains in information and incurring in a connection cost is analyzed in order to drive network stability and efficiency results.

In general, the extra information about the distribution makes the probabilities (not necessarily additive) that each individual had assigned to states of the world change, and hence, their decisions in equilibrium may differ. For example, an agent can decide to connect to someone, and the extra information can shift her bets to a state of the world that she ponders as more likely to happen after updating her beliefs with the new information. Throughout the paper, agents are described as facing a game where they do not know with uncertainty what the move of nature will be, and hence, they infer beliefs about the state of the world in which they will be located. Strategy profiles are taken to be optimal in the Bayesian sense, depending on the relative likelihood of nodes within an information set.

In the discussion, if an individual has a higher degree of uncertainty aversion, she would be willing to connect less since she judges the information to be less reliable as a tool for reducing uncertainty. While computing the expected value of a bet given her information, she would always increase her expected utility in a lesser way by knowing more about the probability distribution of the event in the "real world" than a less uncertainty averse individual. Similarly, less uncertainty averse people would want to connect more. Individuals will decide to connect only if the cost of the acquired information is not higher than the additional value it yields them, computed through Choquet Expected Utility.

One of the main contributions of this paper is introducing a framework to explain the motivation of individuals to connect as a result of their wish to minimize the uncertainty they have about the world. The main results of the paper are showing that different types of social architectures emerge according to the attitudes towards uncertainty of the individuals who conform a particular society, while pinpointing the influence of network architecture on ex-post beliefs and the conditions for reaching a consensus (network or component-wise). A condition for guaranteeing the irrelevance of content in information is proposed as an analytical result, together with a clear relationship between the decision problem individuals
face and the computation of CEU through what we call network-weighted $\varepsilon$-contaminated capacities (a type of non-additive probabilities). Other interesting results involve proving the equivalence between belief functions and these special form of capacities as well as stating and applying an alternative notion of network efficiency, which we call statistical efficiency.

The results are useful not only for the study of social and economic networks but also for the analysis of information transmission in branches of knowledge such as psychology, sociology, telecommunications, etc. To our knowledge, there is no paper that discusses network formation as a response to uncertainty minimization, which is an explanation widely accepted by sociologists of why people establish links in a society.

The paper will be organized as follows: Section 2 presents the relevant literature about networks (specifically formation) and decisions under uncertainty. Section 3 presents the model in detail and its main results. In section 4, a simple example is presented in order to illustrate some of the expected results of the paper. Section 5 talks about possible applications of this framework, both in social networks and in other kinds of informational structures. Section 6 concludes.
2 Literature

Introducing case-based decision theory in a network setting has only been done by Gilles, James, Barkhi and Diamantaras (2007). In their paper, they model network formation using tools from case-based decision theory, by introducing a utility function which depends on the realization of a random network and the payoffs that the neighbours obtain. They include a similarity function that describes how much nodes in the graph resemble the neighbors in question. In other words, the social network constitutes the available memory of agents. After simulations, they draw conclusions on stability and connectedness.

In contrast, the model presented in this paper integrates these two fields in a different manner, using cases as the primitives for information and having a specific utility function which depends on a qualification of the information of other nodes and on the potential network structure. The utility function does not vary with connections or random realizations of the graph. Moreover, the paper will describe how posterior beliefs or distributions are affected by the network architecture and will characterize the resulting social beliefs.

Other than this, no other paper has incorporated case-based decision theory to network formation. Thus, the existing literature relevant for this paper is that of social networks and that of decision theory as an attempt to explain information transmission through individuals. The present review follows this line of thought, with a final summary in order to place the paper and its contributions within the existing economic literature.

2.1 Networks

The beginning of the study of information transmission through social networks can be found in the graph theoretical study by Baker and Shostak (1972) on the propagation of gossip and broadcasts. They, and many who followed, focused on the way information could reach everyone in a network, according to some network criterion, i.e. the minimum number of links required for information to reach all nodes in the graph. These models are revised thoroughly in the survey by Hedetniemi, Hedetniemi and Liestman (1988). In these models,
however, no explicit characterization is made of what the cost or the utility obtained from
the connection represent in relation to the motivation connection of each node in a graph.

The benchmark models which consider the effects on the utility or cost of an individual by
establishing a connection in a network are the connections model and the co-author model
in the paper by Jackson and Wolinsky (1996). They describe the spillover effects within
a network (in the sense of a positive externality) and how individuals decide to connect
among themselves in order to maximize their utility. They call this the connections model,
whose main aim is to analyze network stability and efficiency, providing a characterization
of stable and efficient networks depending on parameters found in the model (e.g. the decay
in the value given by each step and the value each individual yields to other nodes in its
connected component). They also present the co-author model, where we the effects of the
network can be seen in the case of two individuals deciding to be connected or not, when
connections have to be done by mutual consent. In this case, there exist both a positive
and a negative externality. This model is worth mentioning since within the positive effect
there is some intrinsic idea of a sharing of information among the co-authors. Bala and
Goyal (2000) present a directed version of the connections model, where there need not be
mutual agreement for creating the link. Moreover, they assume only one individual sponsors
the link, which has some relevant stability and efficiency consequences, using Nash networks
to discuss stability, comparing it to the pair-wise stability notion proposed by Jackson and

Regarding information transmission, a paper by Noble, Davy and Franks (2004) considers
the effects of network architecture on the transmission of information. In their study, they
find that skewer networks are less efficient in propagating information. They also contrast
their result with the degrees of skewness found in nature and speculate about the possible
reasons for networks to be skewed. Sarangi, Kannan and Ray (2005) present a model where
the formation of Information Networks is analyzed. In their model, each individual pays
for all the information they acquire through indirect links, analyzing stable and efficient
networks. Their paper builds on the Nash networks described by Bala and Goyal (2000),
with agents having different endowments of information, and costs that depend on both the
distance from the source and the value of the information itself.

As for applied work, there have been models dealing with networks and crucial players in them, which are useful in explaining networks that create a negative effect on society, which is the case of the crime networks described in the paper by Ballester, Calvò-Armengol and Zenou (2006). They relate the Bonacich centrality measure in a network with the crime effort exerted by a node in it. Other interesting applications are described in the survey by Jackson (2003).

Informational networks have been studied in the frame of communication theory and linguistics, as seen in the paper by Berger and Calabrese (1975), where the Uncertainty Reduction Theory is stated. This theory in relational development claims that strangers create social links among themselves in order to reduce the uncertainty they have regarding each other. Other studies have focused on how individuals want to learn about the world they live in by gathering information through a social connection, like Borgatti and Cross (2003) who model the probability of seeking information according to knowing what the others know, valuing it, being able to access their thinking in a timely manner and knowing that it is not too costly to access. Their paper consists of an empirical study and states results regarding the relationship between network architecture and information seeking processes.

In our study, the informational structure determines a connection behavior of the agents in a society, depending on their ex-ante information and on their potential gains from connecting to a particular set of nodes (agents) in a graph (network). Stability and efficiency will be analyzed and a relationship between centrality of an individual and her contribution to the ex-post beliefs will be stated.

2.2 Uncertainty

Subjective expected utility maximization was at first axiomatized based on the work of Ramsey (1931) and De Finetti (1937). The Savage (1954) Axioms on Expected Utility constitute the benchmark for Expected Utility Theory. The famous Ellsberg (1961) paradox showed how the Savage axioms did not hold empirically. In his example, Ellsberg shows how
individuals do not act according to the Bayesian approach and that there is some additional component of the assessment of uncertainty. Namely, individuals end up with non-additive probabilities that are justified by an aversion to more "ambiguous" events, given the same expected payoff.

This paradox inspired the modelling of non-additive subjective probabilities. Anscombe and Aumann (1963) provide another derivation of expected utility theory that requires the presence of lotteries with objective probabilities. In their paper they assume an action is no longer a function from states to outcomes, but from states to probability distributions on the set of outcomes. Hence, the consequences of these actions are lotteries themselves. All these works reduced uncertainty to risk by using subjective probabilities.

One notable approach for modelling expected utility is that of Schmeidler (1989) who considers non-additive probabilities, or capacities and computes a Choquet expected utility over states of the world using those capacities. Basically, the main idea is that capacities are the relevant measures of the likelihood of an event and hence, they constitute the statistical ingredient for computing the expectation. Gilboa and Schmeidler (1989) then formulate another approach with Maxmin Expected Utility, which considers computing an expected utility using multiple priors. The computation considers a maximization over the worst possible outcome on each state of the world. These two papers constitute the benchmark on Decision Theory, and we will refer to the theories in them as CEU and MEU respectively in what follows.

Since these seminal papers, a common problem that has tried to be solved in the literature is the differentiation between the ambiguity of an event and the attitude towards ambiguity of the decision makers. Ghirardato, Maccheroni and Marinacci (2004) provide a behavioral characterization of the ambiguity a decision maker perceives and build a representation of ambiguity attitudes based on this revealed ambiguity. Klibanoff, Marinacci, and Mukerji (2005) present a model of smooth preferences, where they characterize expected utility as an expectation over expectations, hence separating the risk and ambiguity components of preferences. They define ambiguity aversion of an individual using the degree of convexity a function on the outside expectation has.
Modelling belief computation and updating is of great importance in the literature. Gayer (2007) defines evaluation of probabilities, according to a memory process where individuals overvalue the probabilities or undervalue them, according to some specific functional forms of memory that can be considered optimistic or pessimistic, hence resulting in different attitudes towards uncertainty due to how individuals store frequencies in their memory. One result that will be used throughout the paper to discuss belief updating is the characterization by Lo (1998). In his paper, he defines $\varepsilon$-contaminated beliefs, a special case of the class of belief functions, which are a special case of convex capacities of Schmeidler (1989).

For our model, the degree of ambiguity aversion will be included in the primitives of the model through the definition of a capacity and hence, will also have updating consequences. The ambiguity of an event does not play a central role per se on the analysis, since the ambiguity of an event will stay the same as dictated by the game structure together with the move of nature and the difference will be a consequence of the attitude towards uncertainty each individual possesses. We will be more specific about this point when defining the model. For now, it is worth mentioning that individuals do not decide based on their perceived ambiguity of an event.

Although the modelling approach in this paper is unprecedented, the use of networks for dealing with information updating is not. In his paper, Pearl (1986) discusses the subject of information transmission using a network setting. He establishes the relationships between logical propositions (nodes) in a network and analyzes propagation and fusion of new information in what he calls belief networks. The links in his model represent the relationship between those propositions judged as directly related among themselves. He also addresses the question of existence of a particular representation (tree structure) and how it implies the revelation of the complete tree topology through the analysis of pairwise relationships between nodes in the graph.

Following the literature, we assume in this paper that ambiguity averse individuals obtain an increasing marginal utility from additional cases available for the formation of a belief on a particular event. The converse is assumed for ambiguity loving individuals, while ambiguity neutral individuals have a constant marginal utility for all cases.
3 The Model

The model presented in this section consists of two main aspects: the decision of individuals in the network sense (this is, whether to establish a link to another individual or not), and the way that beliefs are updated given a certain network architecture. So individuals decide to connect or not, knowing that they will acquire information in the form of a subjective probability distribution of the other individuals on event set $\mathcal{A}$.

In order to solve the model, first the way in which beliefs are updated is presented, followed by the decision connection that already takes it into consideration. For this matter, capacities over states of the world that individuals possess and their attitudes towards ambiguity will play a key role in determining their actions. So first, let us look at the way beliefs are formed by individuals.

3.1 Preliminaries

Let $\Omega = \{\omega_1, ..., \omega_K\}$ be a set of states of the world with $K \geq 2$ and let $\Sigma$ be the associated algebra with power set $2^\Omega$. Let $C$ be a non-empty set of cases, with $c^i \in M \times \Sigma$, where $m$ is a list of variables that describe a context and $a \in \Omega$ is the event that occurred. For now, we consider cases where uncertainty was revealed completely. In case it had not, it would suffice to consider $a$ as the smallest event that was observed.

Each individual observes the world and acquires a history of cases $C_i$ of length $\lambda_i$ with $\lambda_i \in \{\mathbb{N}, 0\}$. The observed history of realizations results in beliefs about an event $\beta_i(x;C_i)$ with $x \in \Omega$. Let us define as well the real probabilities associated with the different states of the world as $p : \Omega \to [0,1]$. Let $q \in Q = \{q_1, ..., q_N\}$ be the set of actions available for an individual before the uncertainty is resolved, and let $\pi : \Omega \times Q \to \mathbb{R}$ be the function that associates a payoff to every action on every possible state of the world. We define $q^*(\omega_k) \in Q$ as the action that maximizes payoffs in state of the world $\omega_k$. For the moment we assume that that $q^*(\omega_k)$ is single-peaked. Hence, as a result of ex-ante optimization we obtain that the optimal strategy for each individual is given by

$$ s^i() = \begin{cases} q^*(\omega_k) & \text{if } \beta_i(\omega_k;C_i) \geq \beta_i(\omega_j;C_i) \quad \forall \omega_j \subset \Omega \end{cases} $$
Considering the above-mentioned strategies, we can also compute the probability that an individual will succeed in choosing a certain action. Let us also make clear that the actions taken will depend on beliefs of individuals about a particular state of the world, while the probabilities of success depend on the objective probabilities of those states. Namely, the probability of success of action $q_y$ is given by $p(\sqrt{q_y}) = \max \{p(\omega_k) : q^*(\omega_k) = q_y\}$. Each individual knows this and hence, she wants to know what the distribution of states of the world is, according to her cases and to information she can potentially acquire from other individuals.

Let $i = \{1, \ldots, I\}$ be individuals in a society and let matrix $G_{I \times I} = [g_{ij}]$ describe a social network, where $g_{ij} = 1$ if there exists a link between individuals $i$ and $j$ and $g_{ij} = 0$ if it does not. Let $g_i = \{j \in G : g_{ij} = 1\}$ be a list of the direct connections of $i$ and $g_i^D = \bigcup_{j \in g_i^{D-1}} g_j^{D-1}$ be the list of all individuals connected to $i$ in a path of length $D$. To include all the individuals in the network connected to $i$ it suffices to consider paths of length $I$ and denote the list $G_i = \bigcup_{j \in g_i^{I-1}} g_j^{I-1}$.\(^1\)

Let $d_{ij}$ represent the geodesic distance between individuals $i$ and $j$ and let $\gamma(d_{ij})$ be a discount function $\Gamma : \{\mathbb{N}, 0, \infty\} \rightarrow (0, 1]$. Initially, we will consider the specific discount function $\Gamma(d_{ij}) = \delta^{d_{ij}}$, where $\delta \in [0, 1]$ used in Jackson and Wolinsky (1996) to describe the connections model in a network under the presence of positive externalities.\(^2\)

We alter this function to reflect the dynamics of information through the network and make it individual, therefore defining $\Gamma_i(d_{ij}) = [1 - \varepsilon_i(N_i)]^{d_{ij}}$ as the relevant discount function. The term $\varepsilon_i(N_i)$ is a contamination function and depends on both $i$’s degree of uncertainty aversion and on the total amount of information available to her. Namely, as the number of entries in the dataset of $i$ increases, the contamination function decreases. It is also true that $\varepsilon_i(N) \geq \varepsilon_j(N) \ \forall N$. The contamination function is also portrays decreasing marginal contamination. In other words, as more information is included in an individual’s dataset, an extra entry on the dataset will affect the overall preciseness of information less than the one before. To illustrate, we can consider the graphs of the contamination function for two

---

1. Considering paths of length $I - 1$ would suffice.
2. The modelling choice using a positive externalities framework is due to the fact of similar spillover effects from information as the ones present in those models.
individuals with different degrees of uncertainty aversion.

In the analysis, the discount is applied to the datasets acquired from other individuals. Define individual $i$’s original dataset as the list of cases $D_{ii} = [C_i]$ and the dataset she obtains from $j$, $D_{ij} = [C_j]$. Also, let us define the concatenation of two datasets $D_{ii}$ and $D_{ij}$ as $D_{ii} \circ D_{ij} = [C_i, C_j]$. In the model, the datasets $i$ obtains from other individuals are contaminated according to the social distance from her. The contaminated concatenated dataset is given by $D_i = \circ_{j \in g_i} (1 - \varepsilon_i(\sum_{k \in g_i} N_k)^{d_{ij}} \times D_{ij})$. The set of all resulting contaminated concatenated datasets is defined as the vector of $I$ lists $D = [D_i]$. Note that the further defines a mapping from the network to the resulting datasets for each individual. Namely, let $\Delta : G \rightarrow D$.

Having defined the datasets, we can now impose some desirability properties on the beliefs that result from the processing of information. In order to do this, we consider the axioms in Billot, Gilboa, Samet and Schmeidler (2005) and adapt them to a our setting.

**Axiom 1 (Contaminated Combination Axiom)** Consider datasets $D_{ii}$ and $[1 - \varepsilon(N_i + N_j)]^{d_{ij}} \times D_{ij}$. If $D_{ii}$ induces beliefs $\beta_i(\omega_j; C_i)$ and $[1 - \varepsilon(N_i + N_j)]^{d_{ij}} \times D_{ij}$ induces beliefs $\tilde{\beta}_i(\omega_j; C_j)$, then the beliefs $\tilde{\beta}_i(\omega_j; C_i)$ induced by $D_i = D_{ii} \circ \left[1 - \varepsilon_i(\sum_{k \in g_i} N_k)^{d_{ij}} \times D_{ij}\right]$ belong to the
convex hull of $\beta_i(\omega_j; C_i)$ and $\tilde{\beta}_i(\omega_j; C_j)$. In other words $\exists \lambda \in [0,1]$ such that $\tilde{\beta}_i(\omega_j; C_i) = \lambda \cdot \beta_i(\omega_j; C_i) + (1 - \lambda) \cdot \tilde{\beta}_i(\omega_j; C_j)$.

Aa in Billot, Gilboa, Samet, and Schmeidler, additional observations of the same case always add to the total information. Additionally, the way that $\lambda$ is determined in equilibrium will depend on the relative lengths of each concatenated dataset. This is, in the definition above, if $N_i = N_j$ we would have $\lambda = 1 - \lambda = 1/2$. In general, this property can be stated as follows. Given $D_{i1}, ..., D_{ik}$ of lengths $N, ..., N_k$ respectively, the weights $\lambda_1, ..., \lambda_k$ in the concatenation $D_i = \circ_{j \in g_i} D_{ij}$ are defined as

$$\lambda_{ij} = \frac{N_j}{\sum_{j \in g_i} N_j}$$

Intuitively, this definition implies that the only thing that matters to differentiate between datasets is the contamination according to network architecture, and that, once contaminated, the computation is still made as an average of the cases available to each individual $i$.

For the next axiom, consider the set of all permutations $\Pi_{\lambda_i}$ that includes all bijections $\pi_i : \{1, ..., \lambda_i\} \rightarrow \{1, ..., \lambda_i\}$. For $D_i$ of length $\lambda_i$ and a permutation $\pi_i \in \Pi_{\lambda_i}$, let $\pi_i D_i$ be the permuted database, that is $\pi_i D_i$ of length $\lambda_i$ is defined by $(\pi_i D_i)^m = D^{\pi_i(m)}$ for $m \leq \lambda_i$.

**Axiom 2 (Invariance Axiom)** For every $\lambda_i \geq 1$, every $D_i$ of length $\lambda_i$ and every permutation $\pi_i \in \Pi_{\lambda_i}$, $p(D_i) = p(\pi_i D_i)$.

This axiom states that cases that appear later in particular dataset have the same influence on the computation of probabilities. Moreover, if an individual knows the number of cases of all others, the weights expressed in the combination axiom are set according to them. In particular, if all the individuals in the connected component of $i$ have $N_j$ elements in their datasets, then $\lambda = 1/\#g_i$, i.e. each dataset (contaminated) is weighted in the same way for computing the subjective probability of success.

By means of these two axioms, we can state a relationship between the probabilities that will result from information exchange once the network is formed.
Lemma 1 (Subjective Probability Proximity) Each subjective probability of succeeding in the game for individual $i$, $p_i$, is given by:

$$v_i(A) = \sum_{j \in g_i} \frac{[1 - \varepsilon_i(N)]^{d_{ij}+1} \times p_j(A)}{\#g_i}$$

Proof in the Appendix.

For the sake of the analysis, it will be useful to show the equivalence between this subjective probability and capacities (non-additive probabilities), which will be the relevant measure used to compute expected utility.

Definition 1 (Capacity) A real-valued set function $v$ on $\Sigma$ is called a capacity if it satisfies the following conditions:

- $v(\emptyset) = 0$
- $v(\Omega) = 1$
- For all $A$ and $B$ in $\Sigma$: $A \subset B \implies v(A) \leq v(B)$

Proposition 1 $v_i(A)$ is a capacity.

Proof in the appendix.

Definition 2 ($\varepsilon$-contaminated beliefs) A belief function $\mu$ is said to be $\varepsilon$-contaminated if it has the following parametric specification: there exist an event $a \subseteq \Omega$, a probability measure $p^* \in M(a)$ and a real number $\varepsilon \in [0,1]$ such that

$$\mu = \{(1 - \varepsilon)p^* + \varepsilon p : p \in M(a)\}$$

where $M(a)$ denotes the set of all probability measures on $a$. 
Each individual knows about the length of the database each one of her potential connections has, and doesn’t know the information included on the cases until connection has happened. The mechanism through which this evolves into a posterior distribution is the acquisition of information from the individuals that she will be connected to in the network. This is, individuals chose to connect to other individuals in order to minimize the uncertainty they have, and then once the information has been revealed, each individual computes her posterior probability distribution taking into account the \( N_j \) cases of the individuals that she is connected to.

Before describing the maximization problem, let us define some graph theoretical concepts that will be useful for the analysis. First let \( G^N \) denote the complete graph, i.e. the set of all subsets of \( I \) of size 2. Let \( ij \) denote the subset of \( I \) that contains only \( i \) and \( j \). Let \( G + ij = G \cup \{ ij \} \) and \( G - ij = G \setminus \{ ij \} \). Let \( K(G) = \{ i | \exists j \text{ such that } ij \in G \} \) and \( k(G) \) be the cardinality of \( K(G) \). A path in \( G \) connecting \( i_1 \) and \( i_N \) is a set of distinct nodes \( \{ i_1, i_2, ..., i_N \} \subseteq K(G) \) such that \( \{ i_1 i_2, i_2 i_3, ..., i_{N-1} i_N \} \subseteq G \).

**Definition 3 (Component of a Graph)** The graph \( G' \subseteq G \) is a component of \( G \), if \( \forall i, j \in K(G') \) and \( i \neq j \), there exists a path in \( G' \) connecting \( i \) and \( j \), and for any \( i \in K(G') \) and \( j \in K(G) \), \( ij \in G \) implies that \( ij \in G' \).

**Definition 4 (Value of a Graph)** The value of a graph is represented by \( \nu : \{ G | G \subseteq G^N \} \rightarrow \mathbb{R} \). The set of all such functions is \( \mathcal{V} \). In our analysis the value will be an aggregate of individual utilities, \( \nu(G) = \sum_i U_i(G, \lambda_j) \).

In order to analyze formation, we will use the notions of pairwise stability and efficiency of a network architecture analogue to the ones presented in Jackson and Wolinsky in a graph theoretical framework.

**Definition 5 (Pairwise Stability)** An graph \( G \) is pairwise stable with respect to \( U \) and \( \nu \) if

(i) for all \( ij \in G \), \( U_i(G, \nu) \geq U_i(G - ij, \nu) \) and \( U_j(G, \nu) \geq U_j(G - ij, \nu) \)
and

(ii) for all \( ij \notin G \), if \( U_i(G, \nu) < U_i(G + ij, \nu) \) then \( U_j(G, \nu) > U_i(G + ij, \nu) \).
**Definition 6 (Strong Efficiency)** A graph $G \subseteq G^I$ is strongly efficient if $v(G) \geq v(G')$ for all $G' \subseteq G^I$.

The definition of a strongly efficient network architecture is also presented here, even though an alternative concept will be used to analyze efficiency in information transferring.

We now describe the utility maximization problem of an agent through means of the Choquet Expected Utility. First let us note that CEU in this setting will be defined as:

$$CEU_i(G) = \left[ \pi \sqrt{\pi} - \pi \right] \times v_i(\sqrt{\cdot}; G) + \pi \chi$$

, since the only possible states considered are succeeding and failing. In the above expression $\pi \sqrt{\pi}$ and $\pi \chi$ stand for the payoffs of success and failure, respectively. On the other hand, this does not restrict per se the number of states of the world of the underlying decision problem. Moreover, individuals will have to pay a cost $c$ for establishing a direct link with another agent in the network and hence acquire her information. For this matter, let us also define the set of direct connections of $i$ as $\tilde{g}_i : j \in G : d_{ij} = 1$ and hence, the total cost of $i$ becomes $c \# \tilde{g}_i$. With this, we can see that the Choquet Expected Utility (net of costs) of individual $i$ can be defined as:

$$\max_{g_i} \left[ \pi \sqrt{\pi} - \pi \chi \right] \times \frac{\sum_{j \in \tilde{g}_i} [1 - \varepsilon_i(N)]^{d_{ij} + 1} \times E p_j(\sqrt{\cdot})}{\# \tilde{g}_i} + \pi \chi - c \# \tilde{g}_i$$

**Definition 7 (CEU-Maximization with Network-Weighted Capacities)** Let $S \in [0, 1]$ be a constant that depends on the number of states of the world in the game. Each individual will solve the following maximization problem:

$$\max_{g_i} \left[ \pi \sqrt{\pi} - \pi \chi \right] \times S \times \frac{\sum_{j \in \tilde{g}_i} [1 - \varepsilon_i(N)]^{d_{ij} + 1}}{\# \tilde{g}_i} + \pi \chi - c \# \tilde{g}_i$$

*Proof in the appendix.*

After formation has occurred, information individuals acquire the histories of other individuals and update their information. Since we will be dealing with information updating,
it would be useful to characterize a criterion for evaluating the quality of the information resulting from the process, i.e. to have some notion of distance between what individuals get from the information they have, including what they get from the others and what the real distribution of the event in question would be.

Even though we can characterize pairwise stable networks, the solution for some values of the parameters may not be unique, and hence, we introduce an additional desirability property by means of the following definitions:

**Definition 8 (Kullback-Leibler Divergence)** Let $\phi_S$ and $\phi^*$ be probability distributions. The Kullback-Leibler Divergence (or relative entropy) of $\phi_S$ from $\phi^*$ is given by:

$$D_{KL}(\phi^* \| \phi_S) = \int_{\Omega} \phi^*(A) \log \frac{\phi^*(A)}{\phi_S(A)}$$

**Definition 9 (Statistical Efficiency)** Let $\phi^*(A)$ be the "real" distribution of an event. We say that social belief $\phi_S$ is statistically more efficient than $\phi'_S$ if $|\phi^*(A) - \phi_S| < |\phi^*(A) - \phi'_S|$. In what follows, we will compute this inequality using the Kullback-Leibler divergence, i.e. $\phi_S$ is statistically more efficient than $\phi'_S$ if $D_{KL}(\phi_S) < D_{KL}(\phi'_S)$.

Through this notion, we will be able to find a unique architecture that renders posterior beliefs that are closer to the real distribution of the event, among the class of pairwise stable networks. This notion of efficiency is used in what follows, since we believe it to be more useful for talking about the transmission of information, and we can see its usefulness by means of the following lemma:

**Proposition 2** Let $G$ be a network with social belief $\phi_S$ (a stationary vector of $\Phi(A; G)$). Let $G^{PW}(U_i(\cdot))$ be the set of all pairwise stable networks for $U_i(\cdot)$. Then $\exists G \in G^{PW}(U_i(\cdot))$ such that $D_{KL}(\phi^*(A) \| \phi_S)$ is minimal.

*Proof in the appendix.*

Another result that will be used for stating the results of the paper deals with the eigenvalues of the belief updating structure. This result will be used to discuss the case of information trading through statistics.
3.1.1 Results on Network Formation

**Proposition 3 (Pairwise Stable Networks)** Characterization of PWS networks (omitted for this submission, will be sent and presented when accepted)

- 
- 
- 
- 
- 
- 
- 

**Proposition 4 (Efficient Networks)** Characterization of SE networks (omitted for this submission, will be sent and presented when accepted)

- 
- 
- 
- 
- 
- 
- 

**Proposition 5 (Statistically Efficient Networks)** Characterization of SEN (omitted for this submission, will be sent and presented when accepted)

- 
- 
- 
- 
- 
- 
-
4 An Illustration

Consider the game:

where each individual can play pure strategies \( s = \{U, NU\} \) or more generally, mixed strategies \( \sigma = \{x, 1 - x : x \in [0, 1]\} \) where \( x \) represents the probability of her playing \( U \). The Bayesian equilibrium of the game is given by

\[
\sigma_i^* = \begin{cases} 
\{1, 0\} & \text{if } p(R) \geq p(NR) \\
\{0, 1\} & \text{if } p(NR) \geq p(R)
\end{cases}
\]

From here, we can compute the probability of an individual being right (i.e. reaching final nodes that yield a payoff \( \pi(\sqrt{\cdot}) \)) by the following way

\[
p_i(\sqrt{\cdot}) = p(R) \times x + p(NR) \times (1 - x)
\]

and assuming that individual \( i \) plays her optimal strategy

\[
p_i(\sqrt{\cdot} | \sigma_i = \sigma_i^*) = \max\{p(R), p(NR)\}
\]

Individuals are able to acquire information from other individuals in a social network. Through this process, they are basically sharing their datasets and hence, the probabilities of an individual being correct in her predictions is related to the one of other players in her connected component. Specifically, the relationship between the probabilities will be given by.

\[
v_i(\sqrt{\cdot}) = \frac{\sum_{j \in g_i} [1 - \varepsilon_j(N)]^{d_{ij} + 1} \times E[p_i(\sqrt{\cdot} | \sigma_i = \sigma_i^*)]}{\#g_i}
\]
It is important to notice that this relationship comes from the fact that the datasets satisfy the contaminated invariance axiom (and the combination axiom as well CHECK) (G&S)

Substituting for the probability of being right assessed by other individuals, we obtain

\[ v_i(\Omega) = \frac{\sum_{j \in g_i} [1 - \varepsilon_i(N)]^{d_{ij} + 1} \times E[p(R), 1 - p(R)]}{\# g_i} \]

**Proposition 6** \((v_i(\cdot) \text{ is a capacity})\) Insert here / / /

(a) \(v_i(\Omega) = 1\)

(b) \(v_i(\emptyset) = 0\)

(c) \(A \subseteq B \Rightarrow v_i(A) \leq v_i(B)\)

Moreover, we develop the expectation operator and get

\[ v_i(\sqrt{\cdot}) = \frac{\sum_{j \in g_i} [1 - \varepsilon_i(N)]^{d_{ij} + 1} \times f^1[p(R), 1 - p(R)] f(p(R)) dp(R)}{\# g_i} \]

At this point, if we assume a uniform prior, i.e. \(p(R) \sim U(0, 1)\), and recalling that

\[ \int_0^1 \max \{x, 1 - x\} dx = \frac{3}{4} \]

we can simplify the expression and obtain:

\[ v_i(\sqrt{\cdot}) = \frac{3}{4} \frac{\sum_{j \in g_i} [1 - \varepsilon_i(N)]^{d_{ij} + 1}}{\# g_i} \]

Having reached this capacity, we can compute Choquet Expected Utility as follows:

\[ CEU = [\pi \sqrt{\cdot} - \pi X] \times 3 \frac{4}{4} \times \frac{\sum_{j \in g_i} [1 - \varepsilon_i(N)]^{d_{ij} + 1}}{\# g_i} + \pi X \]

, and state the connection problem of the individual as:

\[ \max CEU - c \]

or more specifically,

\[ \max \frac{3}{4} \times [\pi \sqrt{\cdot} - \pi X] \times \frac{\sum_{j \in g_i} [1 - \varepsilon_i(N)]^{d_{ij} + 1}}{\# g_i} + \pi X - c \]

**Proposition 7** (Pairwise Stable Networks) • The empty network if \(c \geq \)

20
• An averse-star network if $c \leq c \leq$

• An averse-complete network if $c \leq c \leq$

• A minimally connected CP if $c \leq c \leq$

• A connected CP if $c \leq c \leq$

• The complete network if $c \leq$

(results omitted from this version, will be sent and presented when accepted)

5 Applications

(Section omitted for this submission, will be presented and sent when accepted. The main ideas of this section are the ones described below)

Communication applications.

Information seeking in various settings. Bets, Symmetric payoff settings, stocks.

Examples found in the literature: uncertainty papers, network papers. Relationship of these results to the setting here.

Not to use with one source examples, like broadcasting.
6 Conclusions

The model presented in the paper managed to introduce a model of information-seeking and belief updating in a society, using social networks as the framework, with individuals that acquired all the information from their connections and behaved as Bayesian agents, and also with individuals that acquired only the distribution computed by their connections, and updated their beliefs with an $\varepsilon$-contamination. The resulting belief structure in society differs competely, with the latter specification incorporating a hysteria effect in society. The paper characterized the conditions under which both approaches are equivalent, which would correspond to having a society where ambiguity loving individuals are enough (in number or degree) to compensate for ambiguity averse ones. Empirically, we know that people tend to be averse to ambiguity. Since we characterized a "balanced society" according to eigenvectors in the belief updating structure, only one ambiguity loving individual would be necessary in order to have the possibility of stable social beliefs, when individuals exchange their distributions, given she loves ambiguity enough to compensate for the rest of society.

Proving that the beliefs functions we are dealing with are indeed capacities allows us to work with them in ways other authors have done before, for the analysis of decisions under uncertainty. It is important to notice that the capacities described in the model contain a distance-weighting factor, that could be interpreted differently outside network analysis.

Another result of the model is the relationship between network architecture and social beliefs. In general, we can see that stable networks with the same individuals per component may imply different distributions, depending on how connections are made (e.g. a wheel-network, a star-network). This result is important for two reasons: first, it can be used as a criterion to reduce the set of pairwise stable networks and look for the one that satisfies statistically efficiency; and secondly, it helped us to analyze how the centrality of an individual in a network plays a role in the computation of social beliefs. In particular, the information of a more central individual is weighted higher by more individuals and is incorporated to the final beliefs accordingly.

Another result of the paper states that networks with higher degrees of connectivity
are related to societies being farther away from the truth after the computation of social beliefs. This result is closely related to the notion of statistical efficiency. Since more averse individuals want to be more connected and get more information, the resulting network is more connected. On the other hand, the resulting network embeds a higher contamination and hence, a higher deviation from the real distribution of the event.

Moreover, for network formation analysis, the paper presents a characterization of stable and efficient networks for heterogenous agents. The multiplicity of pairwise stable networks encountered for certain values of the discount factor can be reduced to a unique solution by requiring statistical efficiency. The efficiency notion is not analyzed thoroughly, since we use our alternative definition of efficiency.
7 Bibliography


