Intuitive Priors*

Jawwad Noor

February 8, 2021

Abstract

A probability measure over a multi-dimensional state space is an Intuitive Belief if it is an aggregation of pairwise assocations, which have been formed on the basis of past experience in the environment. Associations are shown to correspond to an analog of pointwise mutual information, and a separability property in beliefs is shown to characterize the model. The formation of associations is modelled as an extension of machine learning. Intuitive Beliefs are shown to exaggerate correlations in low probability states, exhibit the Disposition Effect documented in behavioral finance, and exhibit belief patterns observed in the psychology literature on overconfidence.

JEL classification: C45, D01, D90

Keywords: Belief formation, intuition, system 1, associations, neural networks, machine learning, overconfidence, disposition effect.

1 Introduction

Consider a researcher attending a seminar who finds, without being able to articulate why, that "something does not feel right" about the speaker's results. On reflection, she determines whether her feeling was right or not.

This commonplace experience illustrates the *intuitive process*, a mode of thinking and information-processing that differs substantially from deliberation. The intuitive process is unconscious and precedes deliberation. While the results of deliberation are consciously perceived in the mind, intuitive assessments are perceived only in the particular language of *gut feeling*. The basis for intuition is one's *experience* in their environment. For instance, the researcher's past exposure to various models gave rise to a sense of how different assumptions relate with different conclusions, and this formed the basis for her feeling that something was not right about the speaker's results. Her gut feeling may well have been incorrect, since the intuitive process does not combine impressions with the same logical discipline that deliberation might: the famous mathematician Paul Erdos reportedly did not believe (presumably at an intuitive level) the reasoning underlying the solution to the Monty Hall problem until he was shown a computer simulation (Vazsonyi 1999). But her gut feeling may well have been correct. Indeed, intuition is the source of creativity and raw material for the deliberative process: researchers will often "see" a theorem or an empirical hypothesis before they can establish it rigorously.

***Noor is at the Department of Economics, Boston University, 270 Bay State Road, Boston MA 02215. Email: jnoor@bu.edu.** **The author acknowledges valuable comments from Roland Benabou, Yi-Hsuan Lin and Bart Lipman and helpful feedback from seminar audiences at Boston University and Princeton University. Expert research assistance has been provided by Alonso Ahumada and Yi-Hsuan Lin. The author is also grateful for the hospitality of the Cowles Foundation and the Department of Economics at Yale where part of this research was completed. The usual disclaimer applies.**

This paper studies intuition on the presumption that it is an economically relevant phenomenon. If intuition is indeed shaped by regularities in the environment, then it must in fact serve as an informative, albeit noisy, signal about the environment. Such a signal may be valuable for agents, such as venture capitalists or financial investors, operating in complex, high-dimensional environments where cognitive constraints are surely binding. Perhaps for this reason, the marketing literature notes that investors often factor in their intuitive assessments when making decisions (Salas et al 2010, Hensman and Sadler-Smith 2011, Huang and Pearce 2015, Huang 2018). Anecdotally, investors speak of the importance of having a "feel for the market", and investors may not pursue an opportunity if they are "not feeling it". The idea that decisions may be based on a feeling without articulated reasons is also reminiscent of Keynes (1936, pp 161-162)'s belief that investors were driven by a "spontaneous urge to action", which he referred to as "animal spirits". Although the idea of intuition-as-information is a motivation, the subject of this paper is of a more foundational nature: we explore the structure and formation of intuition. The nature of intuitive reasoning in a dynamic setting is studied in Noor (2020b).

Limiting our scope to a static belief over a multi-dimensional state space, we propose a formal model of Intuitive Beliefs and its formation. We hypothesize that the building blocks of intuition are *associations* – the brain's natural tendency to form *connections* between observations in the world (Betch 2007, Morewedge and Kahneman 2010). Associations are at play, for instance, in memory retrieval cues (such as when a song on the radio evokes a memory), thinking patterns (such as stereotyping), habits (such as a morning ritual of reading the news and having coffee) and phobias (such as when the sight of a spider evokes fear). Such associations are continually strengthened or weakened by experience through the process of associative learning, which is influenced by factors such as frequency, repetition, similarity and salience (see Wasserman and Miller, 1997, for a review of the psychology literature). Due to the enduring nature of associations and the influences that determine its strength, associations are not necessarily "rational". For instance, a piece of news running on a loop on the television can strengthen an association with each repetition, even though the agent is observing just one data point. A person raised by an arachnophobic parent may associate house spiders with danger, despite being presented with evidence to the contrary.

Although accounts of intuition vary (Betch 2007), the relevance of associations and associative learning for intuition is clearly articulated in research in psychology (see Section 6) In this paper we abstract from all features except the associative nature of intuition, and conceptualize intuition as a *simple aggregation of associations*. While deliberation may build on a set of nontrivial relationships between variables, we conceptualize intuition as a summary of simple relationships, in a sense to be made precise.

Model: Our modelling choices take inspiration from Artificial Intelligence (AI). For a preview of our model, take as the primitive a probability measure p over some finite set Ω, consisting of multi-dimensional states of the world, $x = (x_1, .., x_N)$. Intuitive Beliefs take the following form for any state $x = (x_1, .., x_N) \in \Omega$,

$$p(x_1, .., x_N) = \frac{1}{Z} exp \left[\sum_{i<j} a(x_i, x_j) + \sum_{i} b(x_i) \right].$$

In assessing the likelihood of state $x = (x_1, .., x_N)$, the model sums the (real-valued) associations $a(x_i, x_j)$ between each pair x_i, x_j (requiring $i < j$ in the summation is to avoid double-counting) and the background association $b(x_i)$ between each x_i and prior information Ω. For instance, suppose that a manager is evaluating how well three different personality types (corresponding to dimensions 1, 2 and 3) would perform in a group. Specifically, suppose that she is evaluating the probability of the state $(good_1, good_2, good_3)$ that the personality types each perform well together. Her Intuitive Belief will aggregate the strength of the pairwise associations $a(good_i, good_j)$ she has between a good performance by each personality types i, j of workers (perhaps by observing pairs of workers with these personality types working together in the past), along with the strength of the association

$b(good_i)$ between the performance of personality type i and her background information (for instance, she may be aware from experience that the nature of the task is more conducive to productivity for some personality types than others). In the model, the belief $p(good_1, good_2, good_3)$ is a normalized exponential function of the sum of these 6 associations.

The network is identified with the tuple (a, b). The functional form is reminiscent of the density of a multivariate Gaussian distribution, which is given by:

$$f(x) = \frac{1}{\sqrt{2\pi^n|\Lambda|}} exp[-(x-\mu)^T \Lambda^{-1}(x-\mu)],$$

for real-valued vectors $x \in \Omega$, where $\mu \in \mathbb{R}^N$ is the mean of x and Λ is the variance-covariance matrix. The states in our model are not real-valued, but the model is Gaussian in spirit in that beliefs are based on an additive aggregation of pairwise relationships. The model also corresponds to the reduced form of the Boltzmann machine (Hinton and Sejnowski 1983, Ackley et al 1985), an energy-based stochastic neural network used in AI. The simple network architecture does not involve any signals or directed links as in more familiar networks such as feedforward networks, and has proved particularly tractable for our purposes.

Belief Formation: We model the formation of beliefs as the *training* of the weights in (a, b) with respect to an objective probability q over Ω. We take inspiration from machine learning. Applied as is, machine learning involves finding an Intuitive Belief p that minimizes relative entropy with respect to q. However, this training problem does not always have a solution because of the parsimony of our model: we formulate networks that only have “visible” nodes – nodes whose values correspond to observables in the environment – whereas neural networks in AI routinely use “hidden” nodes which have no interface with the environment and provide the free parameters needed for their purposes. This motivates our extension of machine learning, where we imagine that beliefs are trained in an environment where the agent’s experience is limited to randomly observing subsets of dimensions at a time.

Results: The main results of this paper are as follows.

1. Although the representation is plagued with non-uniqueness, we find a *unique canonical* representation that admits a sharp identification of the associative network (a, b). Specifically, we define a notion p^g of “geometric marginal” (a geometric analog of the usual notion of a marginal) and show that an Intuitive Belief p is always represented by a network where:

$$exp[a(x_i, x_j)] = \frac{p^g(x_i x_j)}{p^g(x_i)p^g(x_j)} \text{ and } exp[b(x_i)] = p^g(x_i),$$

that is, where associations $a(x_i, x_j)$ are a measure of correlation reminiscent of *pointwise mutual information*[1] used in information theory, and background associations $b(x_i)$ are a notion of marginal belief.

2. We study inductive inference in the model: we show how a belief formed on a limited number of states extends to a belief over all states. In particular, an agent who has formed beliefs on a limited number of states may nevertheless hold beliefs about states that she has never observed, since her beliefs on the limited number of states may form associations that generate her beliefs about unobserved states. We also characterize the model, showing that its empirical content lies in a separability property.

3. We formulate a notion of training with respect to some objective probability measure q over Ω. We provide sufficient conditions for the existence and uniqueness of a solution and its characterization. We show that when the agent’s experience is limited only to “simple” observations,

[1]Pointwise mutual information between x_i, x_j is defined by $ln\frac{p(x_i x_j)}{p(x_i)p(x_j)}$, which involves 2-dimensional and 1-dimensional marginal distributions of p.

trained Intuitive Beliefs admit a representation where

$$exp[a(x_i, x_j)] = \frac{q(x_i x_j)}{q(x_i)q(x_j)} \text{ and } exp[b(x_i)] = q(x_i),$$

that is, associations can in fact be viewed as an encoding of objective pointwise mutual information and background association as an encoding of the objective marginal. This provides an alternative to the canonical representation, and serves as a sharp mapping from the objective distribution q to the agent's beliefs p.

4. With such a mapping in hand, we study the model's properties in a simple 3-dimensional environment. We show how some patterns of correlation in the data may be exaggerated by beliefs, and how others may not even register. We show that the association $\frac{q(x_i x_j)}{q(x_i)q(x_j)}$ can be particularly strong between rarely occurring instances of x_i and x_j, as if the co-occurrence of two rare instances is "salient" for Intuitive Beliefs. This feature can result in a significant overestimation of low probability states. The model can accommodate a Disposition Effect documented in behavioral finance, whereby investors tend to sell winning stock to soon and hold on to losing stock for too long. Relatedly, the model can accommodate patterns in overconfidence documented by psychology, where subjects tend to be underconfident in their performance on easy problems but tend to be overconfident in their performance on hard problems.

This paper ties together several strands of research to produce a theory of Intuitive Beliefs and its formation: it articulates a notion of intuition based on cognitive psychology, formalizes it by adapting a suitable neural network model and the notion of machine learning from AI, and studies the model in a way that is relevant for application in economics. On a technical level the paper is an exploration of the intersection of economics and AI.

Section 2 presents our model and Section 3 explores its empirical content. Section 4 presents our model of belief formation. Sections 5 presents some simple applications. Section 6 concludes with a discussion of related literature. All proofs are collected in the Appendix.

2 Intuitive Beliefs: Model

2.1 Primitives

For $1 < N < \infty$, the set $\Gamma = \{1, ..., N\}$ of *sources or elements of uncertainty* is a finite subset of $\mathbb{N}_+$ with generic elements $i, j, k, ...$ For each source of uncertainty $i \in \Gamma$, the finite abstract set $\Omega_i = \{x_i, y_i, z_i ...\}$ consists of all possible *elementary states* of source i, and is referred to as the *elementary state space* for source i. The (*full*) *state space* is the product space given by

$$\Omega := \prod_{i \in \Gamma} \Omega_i,$$

with generic element $x = (x_1, ..., x_N)$. An *event* in the full state space is, as usual, an element of some algebra on Ω:

$$\phi \neq \Sigma \subset 2^{\Omega},$$

with generic elements $\mathbf{x}, \mathbf{y}, \mathbf{z}$... For any $I \subset \Gamma$, we adopt the notation $\Omega_I := \prod_{i \in I} \Omega_i$ and $\Omega_{-I} := \prod_{i \notin I} \Omega_i$. Similarly, a state $x_I z_{-I}$ specifies $x_I \in \Omega_I$ and $z_{-I} \in \Omega_{-I}$.

To illustrate the setup, consider the quarterly earnings announcements of a set of companies $\Gamma = \{1, .., N\}$. The possible earnings of company i are given by $\Omega_i = \{x_i, y_i, ..\}$. A state $x = (x_1, ..., x_N)$ is a vector of earnings announcements. The event that companies $I \subset \Gamma$ announce earnings $x_I \in \Omega_I$ is given by $\mathbf{x} = \{x_I z_{-I} \in \Omega : z_{-I} \in \Omega_{-I}\}$.

We model beliefs as a standard probability measure:

Definition 1 *(Beliefs) A belief over* Ω *is a probability measure* p *over* (Ω, Σ)*, that is, a set function that assigns* $p(\mathbf{x}) \in [0,1]$ *for each event* $\mathbf{x} \in \Sigma$ *and satisfies (i)* $p(\Omega) = 1$*, (ii)* $p(\phi) = 0$ *and (iii)* $p(\mathbf{x} \cup \mathbf{y}) = p(\mathbf{x}) + p(\mathbf{y})$ *for any* $\mathbf{x}, \mathbf{y} \in \Sigma$ *s.t.* $\mathbf{x} \cap \mathbf{y} = \boldsymbol{\phi}$.

A belief has full support if $p(x) > 0$ *for all* $x \in \Omega$*. The set of all beliefs on* Ω *is denoted by* $\Delta(\Omega)$.

Beliefs are the primitive of our model. We treat beliefs as observable behavioral objects since they can be derived from betting preferences (Savage 1954) and are routinely elicited in experiments (Schotter and Trevino 2014).

Experiments routinely show that subjects' beliefs are non-probabilistic in that they violate additivity and do not even monotonically assign a higher probability to larger events (Tversky and Kahneman 1974). Nevertheless, we restrict attention to probability measures in this paper as part of a systematic development of our theory. We extend our theory to non-probabilistic intuitive beliefs in Noor (2020a). While this paper deals with a static prior belief, the extension to a dynamic setup is pursued in Noor (2020b).

2.2 Model

An association is a psychological connection between observations. We distinguish between two kinds of associations: the connection between any pair of elementary states x_i and x_j, and that between an elementary state x_i and background information Ω.

Definition 2 *An associative network* (a, b) *on* Ω *is defined by*

(i) a set of nodes given by Γ*, where each node* $i \in \Gamma$ *takes values in* Ω_i*,*

(ii) an association function that assigns to each distinct $i, j \in \Gamma$ *and* $x_i, x_j \in \cup_{k\in\Gamma}\Omega_k$ *a symmetric*[2] *associative weight* $a(x_i, x_j) \in \mathbb{R} \cup \{-\infty\}$*, which we write as* $a(x_i x_j)$*,*

(iii) a background association function b *that assigns* $b(x_i) \in \mathbb{R} \cup \{-\infty\}$ *to each* $x_i \in \cup_{k\in\Gamma}\Omega_k$.

An associative network (a, b) *is real-valued if* (a, b) *are real-valued functions.*

Associations are pairwise and undirected in the model, and serve as the building blocks for beliefs over multi-dimensional states $x \in \Omega$ in the model. The fact that associations are undirected is a simplification, but perhaps without loss of generality in a static setting where information is fixed at Ω.[3] Associations can take on positive or negative real-values, or a value of $-\infty$. The meaning of positive or negative associations will be explored in the sequel. The interpretation of $a(x_i x_j) = -\infty$ is that the occurrence of x_i is maximally associated with the non-occurrence of x_j, and vice versa (due to symmetry). A state of the world involving such $x_i x_j$ will be viewed as impossible in our model.

We model intuition as the aggregation of associations:

Definition 3 *Beliefs* p *over* Ω *are Intuitive Beliefs (IB) if there exists an associative network* (a, b) *and a real number* $Z > 0$ *such that for any* $x \in \Omega$*,*

$$p(x) = \frac{1}{Z} exp \left[\sum_{i<j} a(x_i x_j) + \sum_{i\in\Gamma} b(x_i) \right].$$

The set of all Intuitive Beliefs over Ω *is denoted* $\Delta_{IB}(\Omega)$.

The interpretation of the model was provided in the Introduction. The scalar Z is a function of (a, b), since $Z = Z \times p(\Omega) = \sum_{z\in\Omega} exp \left[\sum_{i<j} a(z_i z_j) + \sum_{i\in\Gamma} b(x_i) \right]$. Therefore, Intuitive Beliefs

[2] That is, $a(x_i x_j) = a(x_j x_i)$.

[3] In Noor (2020b), we allow for information to vary, but the impact of updating is captured adequately by incorporating the dependence of information into background association b alone.

can equivalently be written as $p(x) = \frac{exp\left[\sum_{i<j} a(x_i x_j) + \sum_{i\in\Gamma} b(x_i)\right]}{\sum_{z\in\Omega} exp\left[\sum_{i<j} a(z_i z_j) + \sum_{i\in\Gamma} b(z_i)\right]}$ for any $x \in \Omega$. The use of the exponential function has no meaning beyond implying that $p(x)$ can be described alternatively as multiplicatively aggregating subjective building blocks:

$$p(x) = \frac{1}{Z} \prod_{i<j} \xi(x_i x_j) \times \prod_{i\in\Gamma} \zeta(x_i),$$

where $\xi(x_i x_j) = exp[a(x_i x_j)]$ and $\zeta(x_i) = exp[b(x_i)]$. When there are two or less dimensions, any probability measure can be written as an Intuitive Belief: for instance if $N = 2$ then the model takes the form $p(x_i x_j) = \frac{1}{Z} exp\left[a(x_i x_j) + b(x_i) + b(x_j)\right]$, and any p can be replicated by appropriately chosen a while setting $b = 0$. Intuitive Beliefs have content when $N > 2$.

We already noted in the Introduction that when elementary states Ω_i consist of real numbers, the model subsumes the (density of the) multivariate Gaussian distribution. Nevertheless, in general we can imagine that Intuitive Beliefs view the world through a Gaussian lens, in that they build the joint distribution of variables from pairwise relationships. The functional form also corresponds to the reduced form of the Boltzmann machine studied in AI – see Section 6.2 for details.

2.3 Modelling Higher Intelligence[4]

Intuitive Beliefs reduce the world to basic pairwise relationships and build the association between elementary states $x = (x_1 ... x_n)$ *additively*:

$$a(x_1 ... x_n) := \sum_{i<j} a(x_i x_j) + \sum_{i\in\Gamma} b(x_i).$$

An agent who deliberates would conceive of potentially complicated relationships that may not correspond to a simple additive aggregation of associations. For instance, consider a market with price, demand and supply dimensions, that is, $\Gamma = \{p, D, S\}$, and suppose each dimension can take only two values $\Omega_i = \{high, low\}$. An agent who assesses the market in terms of price theory rules out the possibility of states such as $x = (high_p, low_D, high_S)$, even if the marginal probability of each pair $(high_p, low_D)$, $(high_p, high_S)$ and $(low_D, high_S)$ is positive. In contrast, for Intuitive Beliefs, if each pair is viewed as possible, then each pairwise association is real-valued, and their additive aggregation implies that x must be possible:

$$a(x_i x_j) > -\infty \text{ for all } i, j \in \Gamma \implies p(x) > 0.$$

While our aim is to describe agents who cannot conceive of complex relationships spanning multiple dimensions, we provide comments on how our model might be extended to permit more complex relationships between an arbitrary number of dimensions less than N.

1. One way to accommodate this is to simply merge some dimensions. For instance, an agent may have a theory for how the US, UK and Chinese economies work individually. Here $\Gamma = \{US, UK, C\}$ and for each economy i, the elementary state space $\Omega_i = \prod_{m=1}^{M} \Omega_i^m$ is itself a multi-dimensional space reflecting the various markets Ω_i^m in that economy. The 3-dimensional Intuitive Belief then describes an agent who does not have a theory for how the three economies function together and relies on an additive aggregation of pair-wise associations between the economies.

2. Another direction is to build non-additive relationships into the definition of an associative network. For instance, suppose there are 4 or more dimensions and we wish to model an agent who can form nontrivial associations on up to 3 dimensions only. Ignoring background associations and allowing for 3-way associations, we can write

$$p(x_1 x_2 x_3 x_4) = \frac{1}{Z} exp\left[\sum_{i<j<k} a(x_i x_j x_k)\right],$$

[4]This section can be omitted without loss of continuity.

so that beliefs about high dimensional states patch together overlapping 3-way associations.

3. Finally we note that, in AI, non-additivity is generated by the use of *hidden nodes*, that is, nodes whose values are completely "subjective" in the sense of not being visible to an observer. The nodes in the Intuitive Belief model are called *visible nodes* because their values correspond to observable elementary states. To illustrate how hidden nodes can be used, introduce a single new dimension h into our framework (refer to it as a "hidden dimension") and let its elementary state space be given by $\Omega^h = \{x_h, y_h, ..\}$. Suppose there exists an associative network on $\Omega \times \Omega^h$, and for simplicity assume that there are no background associations and there are no associations between visible dimensions, that is, $b = a(x_i x_j) = 0$. Then Intuitive Beliefs take the form $p(x, x_h) = \frac{1}{Z} exp\left[\sum_{i\in\Gamma} a(x_i x_h)\right]$, where the only assocations are between visible x_i and the hidden x_h. Consider a belief over the observable dimensions Ω that corresponds to the marginal of p on Ω:[5]

$$p^m(x) = \frac{1}{Z} \sum_{z_h \in \Omega^h} exp\left[\sum_{i\in\Gamma} a(x_i z_h)\right].$$

In such a model, the relationship between x_i and x_j comes from the individual associations with each hidden x_h, and is not independent of the other elementary states x_{-ij}. To see how this can capture the above agent who understands price theory, imagine that the hidden node captures the equilibria her theory of the world admits. To obtain $p(high_p, low_D, high_S) = 0$, there must exist no $z_h \in \Omega^h$ such that $a(high_p, z_h) + a(low_D, z_h) + a(high_S, z_h) < -\infty$.

3 Results

In this section we study the empirical content of the model. We making the simplifying assumption that beliefs have full support: $p(x) > 0$ for all $x \in \Omega$. It is easy to see that an Intuitive Belief has full support if and only if it is represented by a real-valued network (a, b).

3.1 Behavioral Meaning of Association

It is perhaps not surprising that the weights (a, b) in the network can be changed in possibly many ways without changing the belief it represents.[6] The leads to our main foundational challenge: to find a *canonical* way of defining a network and representing beliefs.

3.1.1 Preliminaries: Marginals and PMI

Take any probability measure $p \in \Delta(\Omega)$ (not necessarily an Intuitive Belief) and let K_I be the cardinality of Ω_I for any nonempty set of dimensions $I \subset \Gamma$. Since Ω has a product structure,

$$K_I = \prod_{i\in I} K_i,$$

that is, the cardinality of Ω_I is the product of the cardinality of each Ω_i, $i \in I$. Denote the marginal belief on Ω_I by

$$p^m(x_I) = \sum_{z_{-I}\in\Omega_{-I}} p(x_I z_{-I}) \qquad x_I \in \Omega_I.$$

It will be useful to observe that marginals can in fact be viewed as a *normalized mean* of $p(x_I z_{-I})$ averaged over all $z_{-I} \in \Omega_{-I}$:

$$p^m(x_I) = \frac{1}{Z_I} \sum_{z_{-I}\in\Omega_{-I}} \frac{1}{K_{-I}} p(x_I z_{-I})$$

[5]This corresponds to a "Restricted Boltzmann machine with one hidden node", where "Restricted" refers to the restriction that $a(x_i x_j) = 0$ in the layer of visible nodes.

[6]The characterization of non-uniqueness is presented in Theorem 5 in the appendix.

with an appropriate normalizing constant Z_I. That is, a marginal tells us how high $p(x_I z_{-I})$ tends to be on average as we vary $z_{-I} \in \Omega_{-I}$. For expositional reasons, however, we will write marginals equivalently as a normalized mean of $p(x_I z_{-I})$ over all $z \in \Omega$ instead of all $z_{-I} \in \Omega_{-I}$:[7]

$$p^m(x_I) = \frac{1}{Z_I} \sum_{z \in \Omega} \frac{1}{K} p(x_I z_{-I}) \qquad x_I \in \Omega_I.$$

Next, recall the measure of correlation $\varphi(x_i, x_j)$ between two variables known as *pointwise mutual information* (PMI) familiar from information theory:

$$exp[\varphi(x_i, x_j)] = \frac{p^m(x_i x_j)}{p^m(x_i) p^m(x_j)}.$$

Clearly, if p exhibits statistical independence (that is, $p(x) = \prod_{i \in \Gamma} p^m(x_i)$ for all $x \in \Omega$) then there is no correlation $\frac{p^m(x_i x_j)}{p^m(x_i) p^m(x_j)} = 1$. Similarly, positive and negative correlations are captured by numbers respectively higher and lower than 1.

3.1.2 Unique Canonical Representation

We saw that a marginal $p^m(x_I)$ can be viewed as an arithmetic average of $p(x_I z_{-I})$ over all $z \in \Omega$. Define a *normalized geometric marginal* – or *geo-marginal* for short – by the normalized geometric mean of $p(x_I z_{-I})$ over all $z \in \Omega$:[8]

$$p^g(x_I) := \frac{1}{Z_I} \prod_{z \in \Omega} p(x_I z_{-I})^{\frac{1}{K}} \qquad x_I \in \Omega_I.$$

We show in Appendix A that geo-marginals are attractive mathematical objects since their properties mirror those of arithmetic marginals. We also show that marginals and geo-marginals coincide under statistical independence.

For any $x_i, x_j \in \cup_{k \in \Gamma} \Omega_k$, define *geometric PMI* by

$$exp[a_g(x_i x_j)] = \frac{p^g(x_i x_j)}{p^g(x_i) p^g(x_j)}.$$

This is a natural geometric counterpart of PMI. Since marginals and geo-marginals coincide under statistical independence, if p exhibits statistical independence then $exp[a_g(x_i x_j)] = 1$, as in the arithmetic case.

We provide behavioral meaning to the notion of an associative network by identifying associative weights with geometric PMI, and associative bias with uni-dimensional geo-marginals. We refer to such associative networks as:

Definition 4 *An associative network* (a_g, b_g) *over* Ω *is* <u>*canonical*</u> *if for all* $x_i, x_j \in \cup_{k \in \Gamma} \Omega_k$,

$$exp[a_g(x_i x_j)] = \frac{p^g(x_i x_j)}{p^g(x_i) p^g(x_j)} \text{ and } exp[b_g(x_i)] = p^g(x_i).$$

[7]Since $K_I = \prod_{i \in I} K_i$, observe that

$$p^m(x_I) = \frac{1}{Z_I} \sum_{z_{-I} \in \Omega_{-I}} \frac{1}{K_{-I}} p(x_I z_{-I}) = \frac{1}{Z_I} \sum_{z_{-I} \in \Omega_{-I}} \frac{K_I}{K} p(x_I z_{-I}) = \sum_{z_I \in \Omega_I} \left[\sum_{z_{-I} \in \Omega_{-I}} \frac{1}{K} p(x_I z_{-I}) \right] = \frac{1}{Z_I} \sum_{z \in \Omega} \frac{1}{K} p(x_I z_{-I}),$$

where $Z_I = \sum_{y_I \in \Omega_I} \sum_{z_{-I} \in \Omega_{-I}} \frac{1}{K_{-I}} p(y_I z_{-I}) = \frac{1}{K_{-I}} = \sum_{y_I \in \Omega_I} \sum_{z \in \Omega} \frac{1}{K} p(y_I z_{-I})$.

[8]As in the case of arithmetic marginals, this is equivalent to taking a normalized geometric mean of $p(x_I z_{-I})$ over all $z_{-I} \in \Omega_{-I}$ rather than all $z \in \Omega$.

We show that:

Theorem 1 *A belief p with full support is an Intuitive Belief if and only if it is represented by a canonical associative network* (a_g, b_g).

The Theorem confirms that Intuitive Beliefs support the behavioral definition of an associative network. With these behavioral definitions, we can give meaning to a positive (resp. negative) association between x_i and x_j in terms of $\frac{p^g(x_i x_j)}{p^g(x_i)p^g(x_j)} > 1$ (resp. < 1). The existence of a canonical representation implies that the model lends itself to the same kind of analysis as in standard models in economics, where components of the representation can be measured empirically and be used to compare across individuals. This solution to the substantive non-uniqueness issue plaguing the AI-style network-based representation is one of the main contributions of this paper.

By inserting the behavioral expressions into the representation we obtain a useful reduced form of the model:

$$p(x) = \frac{1}{Z} \times \left[\prod_{i<j} \frac{p^g(x_i x_j)}{p^g(x_i)p^g(x_j)} \right] \times \prod_{i \in \Gamma} p^g(x_i) \qquad x \in \Omega.$$

The reduced form implies that we can imagine that the agent arrives at her beliefs by taking a (geometrically) statistically independent belief $\prod_{i \in \Gamma} p^g(x_i)$ as a starting point, and augmenting this with correlations $\frac{p^g(x_i x_j)}{p^g(x_i)p^g(x_j)}$ between each pair of elementary states.

The proof of Theorem 1 first requires a thorough understanding of the uniqueness properties of the model, which is fleshed out in Theorem 5 in the Appendix. This permits us to study conveniently normalized subclasses of representations. Of particular interest are "z-normalized representations" (Lemma 9) where we fix any $z \in \Omega$ and normalize the representation so that $a(z_i z_j) = b(z_i) = 0$ for all distinct $i, j \in \Gamma$ and $a(x_i z_j) = 0$ for all x_i. We show that such representations can be identified by the relationships $a(x_i x_j) = \frac{p(x_i x_j z_{-ij})p(z)}{p(x_i z_{-i})p(x_j z_{-j})}$ and $b(x_i) = \frac{p(x_i z_{-i})}{p(z)}$. Since a geometric mean of any set of representations is still a representation for p, the canonical representation is obtained from the geometric mean of all the z-normalized representations, where z is varied over all of Ω.

3.2 Intuitive Inductive Inference

In our model, if the agent has an association between "animal" and "alibinism" and between "animal" and "swans", then she may intuitively come to believe in the existence of an albino swan even if she has never seen one, that is, she may assign strictly positive probability to the state $(animal, albino, swan)$. We demonstrate a result on the nature of such inductive inference in our model, which will later prove useful in our study of belief formation (Section 4).

Definition 3 defines Intuitive Beliefs only for a domain with a product structure, but in order to study inductive inference we will need to consider an agent who forms a belief over an arbitrary subset of states. For any arbitrary $\phi \neq D \subset \Omega$, we say that the belief $p(.|D)$ is an *Intuitive Belief over D* if it is obtained by conditioning some Intuitive Belief p over Ω on D, that is, if there exists $p \in \Delta_{IB}(\Omega)$ such that:

$$p(x|D) = \frac{p(x)}{\sum_{y \in D} p(x)} \qquad x \in D.$$

Refer to p as the *extension* of $p(.|D)$. In principle, $p(.|D)$ has many extensions, and in particular may include the trivial extension that assigns probability 0 outside D. Denote the set of Intuitive Beliefs over D by $\Delta_{IB}(D)$. Denoting the cardinality of D by $K(D)$, for any belief $p(.|D)$ and dimensions $I \subset \Gamma$, the geo-marginal $p^g(x_I|D)$ is naturally defined by taking a normalized geometric mean of $p(x_I z_{-I}|D)$ over all $z \in D$:

$$p^g(x_I|D) := \frac{1}{Z_I} \prod_{z \in D} p(x_I z_{-I}|D)^{\frac{1}{K(D)}} \qquad x_I \in \Omega_I.$$

Imagine that the agent has formed a belief $p(\cdot|D)$ on some $D \subset \Omega$. The next result characterizes its extension to Ω under a richness condition which essentially requires that each possible pair of elementary states $x_i, x_j \in \cup_{k\in\Gamma}\Omega_k$ appears in states in D. To state this richness condition, fix any set of elementary states $\phi \neq S_i \subset \Omega_i$ for each dimension $i \in \Gamma$, and let $S = \prod_{i\in\Gamma} S_i \subset \Omega$. We say that D is *rich* if for each distinct $i, j \in \Gamma$,

$$\Omega_{ij} \times S_{-ij} \subset D, \tag{1}$$

that is, if for each $x_i, x_j \cup_{k\in\Gamma} \Omega_k$ and $z_{-ij} \in S_{-ij}$ the state $(x_i x_j z_{-ij})$ is in D.

Theorem 2 *A full-support Intuitive Belief $p(\cdot|D)$ over a rich subdomain $D \subset \Omega$ has a unique extension, p over Ω. Moreover, this extension has full support and satisfies:*

$$p(x) = \frac{1}{Z} \times \left[\prod_{i<j} \frac{p^g(x_i x_j|D)}{p^g(x_i|D)p^g(x_j|D)}\right] \times \prod_{i\in\Gamma} p^g(x_i|D) \qquad x \in \Omega.$$

The result establishes that when the agent forms a full-support belief for the states in D, then she has only one way of extending her belief to states outside D. The uniqueness is driven by the richness condition which ensures that beliefs over D fix the values of all the building blocks required to construct a belief over Ω. The result also shows that the extension has full-support, that is, the agent believes in the existence of states that she has not observed. Finally, the result tells us that we can compute her extended beliefs using the canonical representation wrt D.

3.3 Characterization: Separability in Estimated Associations

Recall that the geo-marginal $p^g(x_I) := \frac{1}{Z_I} \prod_{z\in\Omega} p(x_I z_{-I})^{\frac{1}{K}}$ makes use of all $z \in \Omega$. Consider, instead, that we define an "estimate" of $p^g(x_I)$ by using fewer z. We provide a strong result on what an analyst can learn about beliefs p from estimates derived in a particular way. The result highlights a separability property that in fact characterizes Intuitive Beliefs.

Consider a data restriction on dimension $i \in \Gamma$ given by a set of elementary states $\phi \neq S_i \subset \Omega_i$, and let $S = \prod_{i\in\Gamma} S_i \subset \Omega$. For any dimensions $i, j \in \Gamma$, consider the subdomain of Ω that applies the data restriction to all dimensions outside i, j:

$$\Omega_{ij} \times S_{-ij},$$

the cardinality of which is denoted $K(\Omega_{ij} \times S_{-ij})$. Compute a normalized geometric mean of $p(x_I z_{-I})$ using only $z \in \Omega_{ij} \times S_{-ij}$, that is, for any $I \subset \Gamma$, define the *estimated geo-marginal* by:

$$\hat{p}^{S_{-ij}}(x_I) := \frac{1}{Z_I} \prod_{z\in\Omega_{ij}\times S_{-ij}} p(x_I z_{-I})^{\frac{1}{K(\Omega_{ij}\times S_{-ij})}} \qquad x_I \in \Omega_I.$$

Define the corresponding *estimated geometric PMI* for $(x_i, x_j) \in \Omega_{ij}$ by

$$exp[\hat{a}_{S_{-ij}}(x_i x_j)] = \frac{\hat{p}^{S_{-ij}}(x_i x_j)}{\hat{p}^{S_{-ij}}(x_i)\hat{p}^{S_{-ij}}(x_j)}.$$

Consider the following relationship between geo-PMI and estimated geo-PMI:

Definition 5 *Beliefs p over Ω with full support satisfy* <u>*Relative Associative Separability*</u> *(RAS) if for any $\phi \neq S = \prod_{i\in\Gamma} S_i$ and any distinct $i, j \in \Gamma$, there exists $\zeta_{S_{-ij}} > 0$ s.t.*

$$\frac{\hat{p}^{S_{-ij}}(x_i x_j)}{\hat{p}^{S_{-ij}}(x_i)\hat{p}^{S_{-ij}}(x_j)} = \zeta_{S_{-ij}} \times \frac{p^g(x_i x_j)}{p^g(x_i)p^g(x_j)}, \qquad (x_i, x_j) \in \Omega_{ij}.$$

For any distinct dimensions $i, j \in \Gamma$, the restriction S_{-ij} shifts $\frac{p^g(x_i x_j)}{p^g(x_i)p^g(x_j)}$ to $\frac{\hat{p}^{S_{-ij}}(x_i x_j)}{\hat{p}^{S_{-ij}}(x_i)\hat{p}^{S_{-ij}}(x_j)}$. RAS requires that the shift takes the form of scaling $\frac{p^g(x_i x_j)}{p^g(x_i)p^g(x_j)}$ by a constant $\zeta_{S_{-ij}}$ that does not depend on $x_i x_j$. Stated differently, the relative geo-PMI comparing the association between $x_i x_j$ with that between $y_i y_j$ is in fact *equal* to the corresponding relative estimated geo-PMI: for all $x_i x_j$, $y_i y_j$,

$$\frac{p^g(x_i x_j)}{p^g(x_i)p^g(x_j)} / \frac{p^g(y_i y_j)}{p^g(y_i)p^g(y_j)} = \frac{\hat{p}^{S_{-ij}}(x_i x_j)}{\hat{p}^{S_{-ij}}(x_i)\hat{p}^{S_{-ij}}(x_j)} / \frac{\hat{p}^{S_{-ij}}(y_i y_j)}{\hat{p}^{S_{-ij}}(y_i)\hat{p}^{S_{-ij}}(y_j)}.$$

There are two useful observations to be made about RAS. First, in order to determine the true relative strength $\frac{p^g(x_i x_j)}{p^g(x_i)p^g(x_j)} / \frac{p^g(y_i y_j)}{p^g(y_i)p^g(y_j)}$ of associations along two dimension i, j, we do not need all the data on p over Ω: by RAS, it suffices to know the values of p on $\Omega_{ij} \times S_{-ij}$, even if S_{-ij} is a singleton. The second observation is that RAS is in fact a separability property in beliefs: while the estimated association $\frac{\hat{p}^{S_{-ij}}(x_i x_j)}{\hat{p}^{S_{-ij}}(x_i)\hat{p}^{S_{-ij}}(x_j)}$ changes with S_{-ij}, the ratios of such estimated associations equal $\frac{p^g(x_i x_j)}{p^g(x_i)p^g(x_j)} / \frac{p^g(y_i y_j)}{p^g(y_i)p^g(y_j)}$ and are completely *independent* of S_{-ij}. That is, the evaluation of relative estimated associations along dimensions i, j is independent of what values states take outside those dimensions. This separability property in fact characterizes the model:

Theorem 3 *Consider a belief p over Ω with full support. Then p is an Intuitive Belief if and only if it satisfies Relative Associative Separability.*

The proof of necessity of RAS is based on the observation that the ratio $\frac{p(x_i x_j z_{-ij})p(z)}{p(x_i z_{-i})p(x_j z_{-j})}$ must be independent of z_{-ij} for Intuitive Beliefs. This is an expression of the *additive aggregation* of associations in our model, and is the key feature at play in RAS. The proof of sufficiency uses RAS to show how for any fixed $z \in \Omega$ and taking $S = \{z\}$, the beliefs $p(x_i x_j z_{-ij})$ and $p(x_i z_{-i})$ can be used to inductively construct $p(x_i x_j x_k z_{-ijk})$, $p(x_i x_j x_k x_l z_{-ijkl})$, ... by replacing $S = \{z\}$ with $S' = \{x_k z_{-k}\}, S'' = \{x_k x_l z_{-kl}\}$,..., invoking RAS each step of the way. The end result is an expression for $p(x)$ that is entirely in terms of $p(x_i x_j z_{-ij})$ and $p(x_i z_{-i})$, and takes the form of a reduced form of an Intuitive Belief.

The characterization paves a path for generalizing the model to allow for higher intelligence: Find the number C of dimensions so that, rather than looking at estimated pointwise mutual information, an appropriately defined multivariate notion $\frac{\hat{p}_D(x_{i_1} \ldots x_{i_C})}{\prod_{k=1}^{C} \hat{p}_D(x_{i_k})}$ exhibits the analog of RAS. This direction is left for future research.

4 Belief Formation

We have modelled a belief as a probablity measure that is constructed using a building blocks. While we view this feature as necessary in order to think of the belief as reflecting intuition, we also consider it necessary that the building blocks should be shaped by the environment. This leads us to model the formation of beliefs. We do so by thinking of belief formation in terms of the “training of a network” in AI. We first note a limitation in adopting the machine learning approach for our model, and then outline our extension of the machine learning approach.

4.1 Training Problem: Machine Learning

The *Kullback-Leibler (KL) divergence* or *relative entropy* for any two probability measures $p, q \in \Delta(\Omega)$ is defined by

$$KL(q||p) := \sum_{x \in \Omega} q(x) \times ln \frac{q(x)}{p(x)},$$

with the convention that $0 \times ln[0] = 0$. It is well known that KL-divergence is non-negative and strictly convex. Although it fails the triangle inequality, it is pervasively used as a notion of distance between distributions.

Of relevance for us is machine learning applied to stochastic networks, which can be described in the following terms. A stochastic network in AI can be identified with a probability distribution over states. A class of stochastic networks accordingly corresponds to some class $P \subset \Delta(\Omega)$ of probability distributions. Given an objective distribution q over Ω (denoting the frequency distribution over states computed using the given "training data"), machine learning entails finding $p \in P$ that is closest to q in the sense of:

$$min_{p \in P} KL(q||p). \tag{2}$$

A well-known result for the class of Boltzmann machines (Hinton and Sejnowski 1983, Ackley et al 1985) is that if p satisfies the corresponding first order conditions then for all $x_i, x_j \in \cup_{k \in \Gamma} \Omega_k$,

$$p^m(x_i x_j) = q^m(x_i x_j),$$

that is, the two-dimensional marginals generated by the network must equal the objective counterpart. It is remarkable that although p and q are defined over multi-dimensional vectors $x \in \Omega$, they are shown to be closest when their two-dimensional marginals match. This result is used in machine learning for a training algorithm (known as "gradient descent") which, roughly speaking, follows these steps: starting with any weights in the network, the i^{th} iteration of the algorithm adjusts each weight in the network in accordance with the distance between the two-dimensional marginals (more precisely, the partial derivative of KL-divergence wrt the weight), and the algorithm terminates when the marginals are within a pre-specified threshold.

We note that a solution to the training problem may not exist in general. Specifically, the set P may not be rich enough to replicate all the two-dimensional marginals of q. This does not pose an issue in AI for two reasons. First, P can be made arbitrarily rich by adding hidden nodes to the network (Section 2.3). Second, in practice, the training algorithm aspires to reach a threshold of closeness to q, and not exact minimization. In economics, it is preferrable that a solution exists and is unique, and admits a characterization.

The obvious extension of machine learning to our set-up is to posit:

Definition 6 *Intuitive Beliefs p are <u>globally trained</u> by q if p solves*

$$min\ KL(q||r)\ s.t.\ r \in \Delta_{IB}(\Omega).$$

The problems noted above are inherited here. The first order conditions define a system of nonlinear equations. Because the equations are nonlinear, there may not exist a solution even if there are more building blocks in (a, b) than equations. To make matters worse, the system is over-identified if $N > 1$.[9] Therefore there is a need for more parameters in our model. There are at least two routes one can take. One is to extend Intuitive Beliefs to includes hidden nodes as in Section 2.3. We leave this to future research, and instead consider a different route that has some psychological plausibility: place cognitive limitations on the agent's ability to observe the dimensions of any state.

4.2 Extended Training Problem

We begin by supposing that there is a probability measure q over Ω that reflects the objective frequency of occurrence of different states (in Section 4.4 we show that other interpretations are possible). We assume that q satisfies:

[9]There are $\prod_{i<j}[K_i K_j] + \frac{N(N-1)}{2}$ equations generated by the first order conditions defined for each distinct $x_i, x_j \in \cup_{k \in \Gamma} \Omega_k$, and the fact that marginals along any i, j must sum to 1). We can normalize $b = 0$ (by the uniqueness result proved in the appendix, Theorem 5) and so p is determined by a network $(a, 0)$ that has $\sum_{i<j}[K_i K_j]$ parameters.

Assumption 1 *The distribution q has full support.*

4.2.1 Subjective Information

In the belief formation stage, the agent repeatedly observes realizations of the state, with frequencies given by q. We assume that the agent does not necessarily perceive, or perhaps absorb, all the dimensions in any given observation. We suppose that, given a true state $x \in \Omega$, for each dimension i the agent either obtains the news that "the realization in dimension i is $x_i \in \Omega_i$", or obtains "no news about dimension i", which we denote by z_i^*. Thus, for each dimension i the agent receives a *signal* that lies in $\Omega_i^* = \Omega_i \cup \{z_i^*\}$. The full space of signals is given by:

$$\Omega^* = \prod_{i \in \Gamma} \Omega_i^*,$$

where a generic signal $(x_I z_{-I}^*) \in \Omega^*$ reveals the elementary states x_i on dimensions $i \in I \subset \Gamma$, and conceals the remaining. We imagine that information is being filtered in this manner by a subjective information structure σ, modelled as a state-dependent probability distribution.

Assumption 2 *There exists an information structure σ over Ω^* where, conditional on any state $x \in \Omega$ and for any $I \subset \Gamma$, the signal $(x_I z_{-I}^*) \in \Omega^*$ is generated with a probability $\sigma(I) \in [0,1]$ that is independent of x_{-I}.*

The key assumption on the information structure is that the configuration of dimensions randomly revealed is determined *independently of the true state.* To the extent that σ captures a limitation in perception or absorption of objective information, it is reasonable that the probability of perceiving dimensions I should not depend on realizations of the state that she does not perceive.

The agent observes only the signals obtained with strictly positive probability under σ, denoted by:

$$D^* := \{(x_I z_{-I}^*) \in \Omega^* : \sigma(x_I z_{-I}^*) > 0 \text{ for some } x \in \Omega \text{ and } I \subset \Gamma\},$$

where we suppress the dependence of D^* on σ to ease exposition. The information structure in conjunction with the objective probability measure q over Ω generates an objective probability measure over these signals: for each $(x_I z_{-I}^*) \in D^*$,

$$q(x_I z_{-I}^* | D^*) = \sigma(I) \times q^m(x_I). \tag{3}$$

where, as before, $q^m(x_I)$ denotes the marginal objective probability over Ω_I.

While the true distribution is q over Ω, the agent's beliefs are trained by the distribution she observes, namely, $q(\cdot|D^*)$ over $D^* \subset \Omega^*$.

4.2.2 Training

Unlike a standard agent in a learning environment, this agent does not have a prior belief over Ω nor any knowledge of the information structure to use for inference. Consequently, from the agent's perspective, the state space is in fact the space of signals:

$$\Omega^* = \prod_{i \in \Gamma} \Omega_i^*.$$

She must form beliefs over Ω^*, but she only has access to experience in the form of the distribution $q(\cdot|D^*)$ over D^*. The training problem involves three steps: first the agent must (i) use $q(\cdot|D^*)$ to *form* a belief $p(\cdot|D^*)$ over D^*, (ii) *extend* this to a belief $p(\cdot|\Omega^*)$ over the full subjective state space Ω^*, and (iii) her extended beliefs must *induce* some beliefs p over the objective state space $\Omega \subset \Omega^*$ which the analyst can also observe. We formalize these steps in the following way.

(i) Recall the definition of Intuitive Beliefs over a subset of the state space (Section 3), and consider the set $\Delta_{IB}(D^*)$ of all Intuitive Beliefs over D^*. We model the formation of a belief $p(.|D^*)$ over D^* as a solution to the minimization problem

$$min\ KL(\ q(.|D^*)\ ||\ r) \text{ over } r \in \Delta_{IB}(D^*), \tag{4}$$

Thus, to form a belief $p(.|D^*)$ over D^* is to minimize the KL-divergence between the objective probability $q(.|D^*)$ and Intuitive Beliefs in $\Delta_{IB}(D^*)$.

(ii) Once a belief $p(.|D^*)$ is formed, it is extended (in the sense of Section 3) possibly non-uniquely to an Intuitive Belief $p(.|\Omega^*)$ over all signals Ω^*.[10] That is, it is extended to some $p(.|\Omega^*)$ such that for all $x \in D^*$,

$$p(x|D^*) = \frac{p(x|\Omega^*)}{\sum_{y \in D^*} p(x|\Omega^*)}.$$

(iii) For any such extended belief $p(.|\Omega^*)$, the agent cannot not only express beliefs over partially revealed signals but also over *fully revealing* signals $x, y \in \Omega \subset \Omega^*$. It is natural to define her belief p over states Ω in terms of her beliefs $p(.|\Omega^*)$ over fully revealing signals $\Omega \subset \Omega^*$. The beliefs p over Ω that the analyst observes are (uniquely) *induced* by $p(.|\Omega^*)$ in the sense that for all $x \in \Omega$,

$$p(x) = \frac{p(x|\Omega^*)}{\Sigma_{y \in \Omega} p(y|\Omega^*)}.$$

We collect this process of training beliefs in:

Definition 7 *For a given information structure σ, Intuitive Beliefs p over Ω are trained by q over Ω if there exists (i) a solution $p(.|D^*) \in \Delta_{IB}(D^*)$ to (4) which (ii) extends to some $p(.|\Omega^*) \in \Delta_{IB}(\Omega^*)$ that (iii) induces p.*

When $D^* = \Omega$, this notion of training coincides with standard machine learning (Definition 6).

4.3 Results

Our first result extends a known result in the AI literature on Boltzmann machines (Hinton and Sejnowski 1983, Ackley et al 1985). Define the marginal belief on $x_i x_j$ by $p^m(x_i x_j|D^*) = \sum_{x_i x_j y_{-ij} \in D^*} p(x_i x_j y_{-ij}|D^*)$. Define $q(x_i x_j|D^*)$ analogously.

Proposition 1 *Consider an extended environment (q, σ) that satisfies assumptions 1 and 2. There exists an Intuitive Belief p trained by q if and only if there exists $p(.|D^*)$ s.t. for all distinct $i, j \in \Gamma$,*

$$p^m(x_i x_j|D^*) = q^m(x_i x_j|D^*), \qquad x_i, x_j \in \cup_{k \in \Gamma} \Omega_k^*.$$

The existence of a solution to the training problem is tied to the existence of a solution to the minimization problem (4). The existence of this solution is necessary and sufficient for the existence of a trained belief p, since the extension and induced beliefs in Definition 7 always exist.

[10] Interestingly, the requirement that $p(.|\Omega^*)$ is an Intuitive Belief makes it difficult to think of the agent as Bayesian with respect to some unobserved prior over Ω. For such an agent there would be some belief μ over Ω such that $p(x_I z^*_{-I}|\Omega^*) = \sigma(I) \times \mu(x_I)$ where $\mu(x_I) = \sum_{w_{-I} \in \Omega_{-I}} \mu(x_I w_{-I})$ is a marginal belief. In particular, when $I = \Gamma$ in this expression, we see that p and μ would be related by $\mu(x) = \frac{p_\sigma(x|\Omega^*)}{\sigma(\Gamma)}$. But then it must be true that for any $I \subsetneq \Gamma$,

$$p(x_I z^*_{-I}|\Omega^*) = \frac{\sigma(I)}{\sigma(\Gamma)} \times \left[\sum_{w_{-I} \in \Omega_{-I}} p(x_I w_{-I}|\Omega^*) \right].$$

In this equation, there is an Intuitive Belief on the left-hand side of the equation, and a sum of Intuitive Beliefs on the right-hand side. However, the sum of Intuitive Beliefs is not generically an Intuitive Belief.

Additional structure is obtained by imposing structure on σ. Note that some signals are "simpler" than others. Consider in particular signals that reveal at most two dimensions: z^*, $x_i z^*_{-i}$ and $x_i x_j z^*_{-ij}$. Refer to such signals as being *simple* (with respect to z^*) and denote the set of all such states by

$$D_{z^*} := \{(x_i x_j z^*_{-ij}) \in \Omega^* : x_i, x_j \in \cup_{k \in \Gamma} \Omega^*_k\}. \tag{5}$$

(This is a "rich" set of states with $S = \{z^*\}$ in the sense of Section 3.2). It will be natural to assume that the agent can at least observe all simple signals:

Assumption 3 *σ satisfies $D_{z^*} \subset D^*$.*

Proposition 2 *Consider an extended environment (q, σ) that satisfies assumptions 1, 2 and 3. If there exists an Intuitive Belief p over Ω trained by q, then it must be unique.*

The proof of the result proceeds as follows. Although the set of Intuitive Beliefs is not necessarily convex, we define a convex space A of normalized representations for Intuitive Beliefs in $\Delta_{IB}(D^*)$. For any $a \in A$ and Intuitive Belief $p_a \in \Delta_{IB}(D^*)$ that it represents, we say that the "distance" between q and a is $KL(q||p_a)$, and accordingly define an objective function $a \mapsto g(a) := KL(q||p_a)$. We prove that this function must be strictly convex in a under Assumption 3. Therefore the minimizer $a \in A$ of $g(a)$, if it exists, must be unique. This determines the uniqueness of a solution $p(\cdot|D^*)$ to (4), if one exists, and Theorem 2 then guarantees a unique extension to $p(\cdot|\Omega^*)$.

As noted earlier, the system of first order conditions is over-determined in the machine learning setup ($D^* = \Omega$) and in general does not yield a solution. However if D^* is a small enough set, then a solution can be guaranteed. A noteworthy special case is where:

Assumption 4 *σ satisfies $D_{z^*} = D^*$.*

That is, the agent's beliefs are trained only by the simple signals. It is plausible that simpler information is more readily grasped than more complex information, and if 3 dimensional relationships are too complex for the agent, then Assumption 4 is natural. We show that this assumption has a lot of bite: it guarantees the existence of a solution, and also yields a sharp characterization of our main object of interest – the trained Intuitive Belief p over Ω. Although our next result follows readily from previous results, we state it as a theorem due to its value as a characterization of trained Intuitive Beliefs.

Theorem 4 *In an extended environment (q, σ) that satisfies Assumptions 1, 2 and 4, there exists a unique Intuitive Belief p over Ω trained by q over Ω. It is represented by an associative network (a, b) defined by: for each $i, j \in \Gamma$ and $x_i x_j \in \cup_{k \in \Gamma} \Omega_k$,*

$$exp[a(x_i x_j)] = \frac{q^m(x_i x_j)}{q^m(x_i) q^m(x_j)} \text{ and } exp[b(x_i)] = q^m(x_i),$$

and admits the reduced form:

$$p(x) = \frac{1}{Z} \left[\prod_{i<j} \frac{q^m(x_i x_j)}{q^m(x_i) q^m(x_j)} \right] \times \prod_{i \in \Gamma} q^m(x_i) \qquad x \in \Omega.$$

The result provides a mapping from the objective distribution q to trained beliefs p. It tells us that trained beliefs are as if associations encode the objective PMI, and the background association encodes the objective uni-dimensional marginal. Observe the desirable property that trained beliefs are independent of the probabilities used by the subjective information structure σ, as they do not appear in any of the expressions.

While this representation for p is an alternative to the canonical representation (Theorem 1), it provides a direct means to study how the environment impacts the agent's intuitions. We consider some applications in Section 5.

4.4 Alternative Interpretations of q

According to the Availability Heuristic (Tversky and Kahneman 1974), the likelihood of a state reflects the *ease of recall* of instances of that state. We note that the ease of recall is related to association strength, which is influence by factors such as the frequency of observations, their salience, similarities with other states, etc. In the training problem we interpreted the objective distribution q as a frequency distribution, and thus our model exhibits the feature that frequently observed states will be assigned a higher belief. The interpretation of q can be extended to capture other features that relate to the Availability Heuristic. For instance, repetition of a state can also affect beliefs, and this can be modelled as follows. We have already presumed that each state x may occurs multiple times, but now denote each occurrence of x by $x(1), x(2), ...$ The news of the occurrence $x(m)$ may be repeated, and let $r(x(m))$ denote the frequency of this repetition. If x occurs at M times then let $r(x(m)) = 0$ for any larger number $m > M$. Then we can define the agent's repetition-adjusted experience q by the formula $q(x) = \frac{\sum_{m\in\mathbb{N}_+} r(x(m))}{\sum_{y\in\Omega}\sum_{m\in\mathbb{N}_+} r(y(m))}$.

5 Applications

We consider several simple applications that suggest that some of the patterns of beliefs and behavior observed in the literature may be explicable by the idea that agents turn to intuitive assessments when in complex environments such as financial markets.

5.1 The Setting

There are three companies $\Gamma = \{1, 2, 3\}$ and the performance of each company $i \in \Gamma$ is either high or low, $\Omega_i = \{h, l\}$. The state space consists of vectors of performance of the companies, $\Omega = \{h, l\}^3$. Consider a "base-line" distribution q_b over Ω that is uniform and satisfies statistical independence:

$$q_b(x_1x_2x_3) = \frac{1}{8}, \qquad (x_ix_jx_k) \in \Omega,$$

that is, each dimension is high or low with marginal probability $\frac{1}{2}$. Correlation is introduced with the help of an "extreme" distribution r over Ω that violates statistical independence. We presume that the agent observes the objective distribution q over Ω that α-mixes r and q_b for some $\alpha \in [0, 1)$:

$$q(x_1x_2x_3) = \alpha r(x_1x_2x_3) + (1-\alpha)\frac{1}{8}, \qquad (x_ix_jx_k) \in \Omega.$$

Since $\alpha < 1$ and q_b has full support, so does q.

We assume throughout that the agent's beliefs p about the joint performance of the companies are trained by the objective distribution q in the sense of Theorem 4. Therefore, trained beliefs satisfy

$$\begin{aligned} p(x_1x_2x_3) &= \frac{1}{Z}\left[\frac{q^m(x_1x_2)}{q^m(x_1)q^m(x_2)}\frac{q^m(x_1x_3)}{q^m(x_1)q^m(x_3)}\frac{q^m(x_2x_3)}{q^m(x_2)q^m(x_3)}\right] \times [q^m(x_1)q^m(x_2)q^m(x_3)] \\ &= \frac{1}{Z}\frac{q^m(x_1x_2)q^m(x_1x_3)q^m(x_2x_3)}{q^m(x_1)q^m(x_2)q^m(x_3)}, \end{aligned}$$

where $Z = \sum_{(y_1y_2y_3)\in\Omega} \frac{q^m(y_1y_2)q^m(y_1y_3)q^m(y_2y_3)}{q^m(y_1)q^m(y_2)q^m(y_3)}$.

As a benchmark, we first observe that if an agent with Intuitive Belief is trained by the statistically independent base-line q_b, then she is perfectly calibrated:

Proposition 3 *Suppose that Intuitive Beliefs p over Ω are trained by q_b over Ω. Then $p = q_b$.*

This is an immediate consequence of Theorem 4: if objective PMI satisfies $\frac{q_b^m(x_i x_j)}{q_b^m(x_i) q_b^m(x_j)} = 1$ for all x_i, x_j then $p(x) = \frac{1}{Z} \prod_{i \in \Gamma} q_b^m(x_i) = \frac{1}{8}$ for all $x \in \Omega$. Indeed, this is a general feature of the model: for any $N \geq 1$, statistical independence between any two dimensions in the environment is imbibed by trained beliefs. We show next that, in contrast, when there exist correlations, trained beliefs may diverge from the objective distribution.

5.2 Correlation and Miscalibration

Consider first the case where the companies' performance are positively correlated, which we model by presuming $r(h_1h_2h_3) = r(l_1l_2l_3) = 0.5$ in the extreme distribution. When mixed with the uniform base-line distribution q_b, the states $(h_1h_2h_3)$ and $(l_1l_2l_3)$ therefore have the highest objective probability according to q:

$$q(h_ih_jh_k) = q(l_il_jl_k) > q(h_ih_jl_k) = q(h_il_jl_k).$$

In this sense there is a clear pattern of correlation in q. We show in this case that trained beliefs exaggerate the probability of co-movement of the companies' performance. (There is no real distinction between "high" and "low", but we will see that correlations exhibited in the most likely states matter here).

Proposition 4 *Suppose that Intuitive Beliefs p over Ω are trained by q over Ω, which is defined by an arbitrary $\alpha \in (0, 1)$ and an extreme distribution that satisfies:*

$$r(h_1h_2h_3) = r(l_1l_2l_3) = 0.5.$$

Then, for any distinct $i, j, k \in \Gamma$, trained beliefs p satisfy:

$$p(h_ih_jh_k) = p(l_il_jl_k) > q(h_ih_jh_k) = q(l_il_jl_k)$$

$$> q(h_ih_jl_k) = q(h_il_jl_k) > p(h_ih_jl_k) = p(h_il_jl_k).$$

Therefore, trained beliefs are less dispersed than the objective distribution. Although Intuitive Beliefs cannot directly detect 3-way correlation in the environment, in this case the 3-way correlation translates into a strong pairwise correlation (PMI) between performances. When these strong pairwise correlations are combined to form p, the result is an overestimation of the probability of $(h_1h_2h_3)$ and $(l_1l_2l_3)$. In a setting where the agent can invest in each companys' stocks, Proposition 4 suggests that the agent would view returns as being more risky than they really are, thereby demanding a higher premium than a well-calibrated agent.

Next, consider the case where correlation is in some sense less clear in the extreme distribution r. Imagine that the companies are operating in a small market where either it is the case that all the companies have low earning, or that two companies have high earnings at the expense of a third. In particular, $r(h_ih_jh_k) = r(h_il_jl_k) = 0$ for any $i, j, k \in \Gamma$, and all other states have equal probability $r(h_ih_jl_k) = \frac{1}{4}$. When r is α-mixed, $\alpha \in (0, 1)$, with the uniform base-line distribution q_b, it must be that for all $i, j, k \in \Gamma$,

$$q(h_ih_jl_k) > \frac{1}{8} > q(h_ih_jh_k) = q(h_il_jl_k).$$

In this case:

Proposition 5 *Suppose that Intuitive Beliefs p over Ω are trained by q over Ω defined by an arbitrary $\alpha \in (0, 1)$ and an extreme distribution that satisfies: for all distinct $i, j, k \in \Gamma$,*

$$r(h_ih_jl_k) = r(l_il_jl_k) = \frac{1}{4}.$$

Then, for any distinct $i, j, k \in \Gamma$, trained beliefs p satisfy:

$$p(x_ix_jx_k) = \frac{1}{8} \qquad (x_ix_jx_k) \in \Omega.$$

Thus, trained beliefs are more dispersed than the objective distribution. To highlight the reason most transparently, consider what happens when p is trained directly by r defined in the proposition, rather than the noisy version of it given by q. Observe that r is such that for all distinct $i, j \in \Gamma$,

$$r^m(h_ih_j) = r^m(l_il_j) = r^m(h_il_j) = \frac{1}{4},$$

that is, the marginal distribution for any pair of dimensions is uniform. It follows that $\frac{r^m(x_ix_j)}{r^m(x_i)r^m(x_j)} = 1$ for all x_ix_j, that is, despite the 3 way correlation, there is no pairwise correlation. Indeed, $p(x_ix_jx_k) = \frac{1}{8}$ for all $x \in \Omega$, that is, trained beliefs exhibit statistical independence, failing to capture any of the correlation in r.

These propositions illustrate that trained beliefs depend on the pattern of pairwise correlations, and how the pattern of 3-way correlations matter in determining them.

5.3 The Disposition Effect

The Disposition Effect is the finding that investors sell winning stock too soon and hold on to losing stock for too long (see Shefrin and Statman, 1985, and subsequent literature). We show that Intuitive Beliefs can lead to such behavior.

Capture the case where the market is "trending upward" by assuming $r(h_1h_2h_3) > r(l_1l_2l_3)$. We consider how beliefs behave when the trend is "strong" in the sense that there is little noise in the true distribution q, captured by looking at α close to 1. We also look at the analogous case of a strong downward trend.

Proposition 6 *Suppose that for any $\alpha \in (0, 1)$, Intuitive Beliefs p over Ω are trained by q over Ω. If the extreme distribution satisfies: $r(h_1h_2h_3) + r(l_1l_2l_3) = 1$ and*

$$r(h_1h_2h_3) > r(l_1l_2l_3),$$

then for all α sufficiently close to 1, the trained beliefs p satisfy

$$p(h_ih_jh_k) < q(h_ih_jh_k) \text{ and } p(l_il_jl_k) > q(l_il_jl_k).$$

These inequalities are reversed if $r(h_1h_2h_3) < r(l_1l_2l_3)$.

Therefore, when the market is strongly trending upward, the agent's beliefs underestimate the true probability of upward movement, and overestimate the true probability of downward movement. An outside observer will find that the agent's willingness to hold the stocks is "too low" given the true distribution when the market is going up. Similarly, when the market is going down, the analyst will find the agent's willingness to be "too high".

The result is driven by how associations are formed for low vs high probability elementary states. Using the expressions derived in the proof, we see that

$$lim_{\alpha \to 1} \frac{q^m(h_ih_j)}{q^m(h_i)q^m(h_j)} = \frac{1}{r(h_1h_2h_3)} \text{ and } lim_{\alpha \to 1} \frac{q^m(l_il_j)}{q^m(l_i)q^m(l_j)} = \frac{1}{r(l_1l_2l_3)},$$

and therefore, when the high state is more likely, $r(h_1h_2h_3) > r(l_1l_2l_3)$, then the association between high performances are not as strong as they are for low performances. The idea is that when the marginal probability of l_i and of l_j is small, then the co-occurrence l_il_j is that much more *salient*, leading to a possibly high association $\frac{q^m(l_il_j)}{q^m(l_i)q^m(l_j)}$. The agent still assigns higher probability to the objectively more likely state, but the miscalibration of correlation leads to a miscalibration in beliefs.

5.4 Rare Events

Rietz (1988) and Barro (2006) suggest that several asset pricing puzzles (such as high equity premium, low risk-free rate, and volatile stock returns) can be explained by the potential for rare economic disasters. We show that Intuitive Beliefs can significantly overestimate low probability states.

This is in fact already illustrated in Proposition 6 in the special case where $r(l_1l_2l_3) = 0$. However, since the probability of multiple states go to zero there as $\alpha \to 1$, we verify the desired property for the case where only one state, say $(l_1l_2l_3)$, goes to 0. Specifically, we assume that r assigns probability 0 to $(l_1l_2l_3)$ and is otherwise uniform on the remaining states:

Proposition 7 *Suppose that for any $\alpha \in (0,1)$, Intuitive Beliefs p over Ω are trained by q over Ω. If the extreme distribution satisfies:*

$$r(x_1x_2x_3) = \frac{1}{7} \text{ for all } (x_1x_2x_3) \neq (l_1l_2l_3),$$

then there is $\lambda_1, \lambda_2 > 0$ s.t. for all α sufficiently close to 1,

$$p(l_1l_2l_3) > \lambda_1 > \lambda_2 > q(l_1l_2l_3).$$

The calculation in the proof reveals that while $q(l_1l_2l_3)$ approaches 0 as $\alpha \to 1$, the trained belief $p(l_1l_2l_3)$ approaches 3.7%.

5.5 Patterns in Overconfidence

The psychology literature studying overconfidence finds that subjects tend to be overconfident about their performance on difficult tasks and underconfident about their performance on easy tasks (Moore and Healy 2008). This is accommodated as a special case of Proposition 6. Reinterpret the setting as follows: each $i \in \Gamma$ corresponds to a question on a test, and the performance on any question can be high or low. To model an easy test, let the extreme distribution satisfy $r(h_1h_2h_3) = 1$ (so that the most likely state is one where the agent gets all the questions right). Similarly to model a difficult test, let $r(l_1l_2l_3) = 1$. The Proposition 6 states that for a sufficiently easy task (that is, as $\alpha \to 1$), the agent underestimates her performance, $p(h_ih_jh_k) < q(h_ih_jh_k)$, and in a sufficiently difficult task she overestimates it, $p(h_ih_jh_k) > q(h_ih_jh_k)$.

The psychology literature finds robust evidence of overprecision of beliefs, that is, beliefs tend to be less dispersed than the objective distribution (Moore and Healy 2008). Proposition 4 shows that Intuitive Beliefs are overprecise when correlations in the objective distribution imply high pairwise correlation. Proposition 5 shows that overprecision is not a general property of Intuitive Beliefs, however, since beliefs may incorrectly infer global statistical independence from pairwise statistic independence.

6 Related Literature

6.1 Psychology

Associationism, a philosophical school with early expositors such as Hume (1748), recognized the creation of associations as the most basic function of the mind and sought to reduce all mental life to associations. It served as the foundation of behavioral psychology in the early 20th century until it gave way to the cognitive revolution in the mid 20th century. In its particular manifestation as conscious memory, assocations have been modelled as networks in cognitive psychology using spreading activation networks (Collins and Loftus 1975, Anderson 1983). More advanced modelling of associative memory was taken up in the study of artificial neural networks in AI (for instance Hopfield 1982), which also had an influence in psychology (Kahana 2020).

Betch (2007) defines intuition in the following terms: "Intuition is a process of thinking. The input to this process is mostly provided by knowledge...primarily acquired via associative learning. The input is processed automatically and without conscious awareness. The output of the process is a feeling that can serve as a basis for judgments and decisions". Morewedge and Kahneman (2010) posit that intuitive judgements are made through automatic, non-deliberative "System 1" processing, which makes use of heuristics and associative memory. Because it incorporates heuristics into information processing, we view System 1 as a potentially richer phenomenon than what our model captures. That said, future research might determine the extent to which a purely associative framework such as ours might be able to subsume the workings of heuristics.

Turning more specifically to beliefs, the dominant paradigm in the psychology literature on intuitive judgement under uncertainty is the celebrated Heuristics and Biases program of Kahneman and Tversky, which posits that people's beliefs are heuristic-based intuitive judgements (Tversky and Kahneman 1974). The three heuristics discussed in this literature are Availability, Representativeness and Anchoring and Adjustment heuristics, of which the first two are more closely related to this paper. According to the Availability heuristic, people assess likelihoods in terms of how many salient examples come to their minds. According to the Representativeness heuristic, people update their likelihoods based on the representativeness of the information for the proposition being evaluated, which are driven by similarity considerations. Salience and similarity are notions subsumed by the notion of associations and therefore can potentially be accommodated in our model. However, it should be reiterated that our model is about how intuitions are formed for complex objects (multi-dimensional states). The psychology experiments focus on intuitive judgements in simple, low-dimensional problems, and our model imposes no structure on beliefs when there are 2 or fewer dimensions.

6.2 AI

AI has studied a wide variety of neural networks, and our model is related to the specific class of energy-based neural networks (Hopfield, 1982). Unlike the familiar neural networks with signals running through a directed graph, energy-based networks utilize the notion of "energy" in the network (the inspiration comes from the classic Ising model used in Statistical Physics, where the electro-magnetic spin of atoms interact with each other to create energy). Our model is inspired more specifically by the Boltzmann machine (Hinton and Sejnowski 1983, Ackley et al 1985). Consider a network as in our model, and suppose nodes can either be activated or not, taking on only two values, $\Omega_i = \{0, 1\}$ (the values $\{-1, 1\}$ are also used). A state is thus a configuration of activations of the nodes. For any configuration $x \in \Omega$, the *energy* in the network is defined as

$$\Lambda(x) = -[\sum_{i<j} a(x_i x_j) + \sum_{i \in \Gamma} b(x_i)],$$

which is taken as a negative quantity due to the convention in Physics. The term $a(x_i x_j)$ is called a weight and $b(x_i)$ is called a bias. In the Boltzmann machine, the probability that any node i is on or off is a (logistic) function of the energy in the system: $p(i = on) = \frac{1}{1+exp[\frac{\Delta\Lambda_i}{T}]}$ where $\Delta\Lambda_i$ the change in energy in the system if node i is activated relative to when it is not, and T is a scalar called the "temperature" in the system. The configuration of activations of the nodes therefore evolves probabilistically. At a "thermal equilibrium", the probability of a configuration is given by the Boltzmann-Gibbs distribution $p(x) = \frac{1}{Z}exp[-\Lambda]$, where the scalar Z is referred to as the "partition function".

To understand how the Boltzmann machine is applied in practice, imagine each pixel on a screen as a node, where each pixel is either activated or not activated. Think of an image on the screen generating a particular configuration for visible nodes. Include an additional visible node i that takes values $\Omega_i = \{cat, not\ cat\}$. If one is interested in teaching a Boltzmann machine to recognize the image of a cat then, for any image on the screen, the Boltzmann machine (depending on the weights in the network, which include hidden nodes) produces a conditional probability that the image is a cat. The Boltzmann machine can be "trained" using different images and its weights adjusted by an algorithm until it predicts (using maximum likelihood) the cat image sufficiently well.

This functional form is the end point of the Boltzmann machine and the starting point for our model, in that our model is based only on the reduced form of the Boltzmann machine in its equilibrium. One can imagine that the details of the Boltzmann model can be used to provide *cognitive* foundations for Intuitive Beliefs, so that when the agent thinks about x_i she probabilistically also thinks about x_j and so on. A final point of departure from the Boltzmann machine, and AI more generally, is that our model is restricted to *visible nodes* (which take values that correspond to elements in the environment, such as elementary states), while AI pervasively makes use of *hidden nodes* which effectively add free parameters (see Section 2.3 for an illustration).

6.3 Economics

The formalization in economics of Heuristics and Biases largely takes the form of modelling particular findings (Rabin (2002) models the law of small numbers, Rabin and Vayanos (2010) model the gambler's and hot-hand fallacies, Benjamin et al (2019) model base-rate neglect, Gennaioli and Shleifer (2010) model the Representativeness heuristic). Mullainathan (2002) and Bodoh-Creed (2020) explicitly incorporate properties of associative memory (such as the fact that a memory is easier to recall the more frequently it is recalled). As models of "Bayesian updating with a bias", these papers presume an arbitrary unexplained prior, and regard a belief as intuitive to the extent that it exhibits an updating bias. The current paper does not provide a model of updating, and instead explore what it means for a prior to be intuitive.

Memory is also modelled in Gilboa and Schmeidler (2001) (in the context of choice under subjective uncertainty) and Bordalo et al (2019) (in the context of choice under certainty). Past experiences are consciously considered by the agent in the former, and automatically come to the agent's mind in the latter, and in both cases some notion of similarity determines the extent to which something is considered or recalled. Memory is modelled as a set of past cases or experiences in either model. In our model, the agent's past experiences are encoded in her associative network, and notions such as similarity are subsumed in the more abstract notion of association.

In the narrow literature on belief formation, Gilboa and Schmeidler (2003) write a case-based model of an agent whose likelihood assessment about a proposition is a similarity-weighted sum of support for that proposition offered by a set of past cases that the agent has access to. As motivation for their seminal work on case-based decision theory, Gilboa and Schmeidler (2001) cite Hume, an early Associationist philosopher, who states that "[f]rom causes which appear *similar* we expect similar effects" (Hume 1748). Spiegler (2016) uses Bayesian networks to describe an agent who uses data in conjunction with her causal theory of the world in order to form beliefs. Using a directed acyclic graph (DAG), a Bayesian network graphically represents the conditional independence between different variables, and uses the chain rule to aggregate conditional probabilities to form a prior. The agent draws objective conditional probabilities from some objective distribution q, and constructs her prior by the chain rule.[11] Posterior beliefs are calculated using Bayes' Rule.

[11] For instance, the graph $x_2 \to x_1 \leftarrow x_3$ represents the belief that x_1 is conditionally dependent on x_2 and x_3, which are conditionally independent of each other and x_1. The chain rule yields $p(x_1x_2x_3) = q(x_1|x_2x_3)q(x_2)q(x_3)$.

Similar to Spiegler (2016), we draw from the computer science literature, but we do so by adapting energy-based networks, which involve fully connected networks with symmetric weights and share little with Bayesian networks. Similar to Gilboa and Schmeidler (2003) and Spiegler (2016), we are interested in the mapping from data to beliefs. However, our interest is in the formation of intuitive beliefs whereas these papers are, at least in exposition, deliberative in spirit.[12] For an illustration of the difference between intuitive and deliberative beliefs we note that associations can be altered if the agent is faced with copies of one data point, whereas models of conscious decision-making or bounded rationality would treat the news as one data point.

A Appendix: Geometric Marginals

Consider any probability measure $p \in \Delta(\Omega)$, not necessarily an Intuitive Belief. Denote the cardinality of Ω_i by K_i. Then for any $\phi \neq I \subset \Gamma$, the cardinality of $\Omega_I = \prod_{i \in I} \Omega_i$ is $K_I := \prod_{i \in I} K_i$.

For any $\phi \neq I \subset \Gamma$, define the geometric I-marginal of $p \in \Delta(\Omega)$ by

$$p^g(x_I) := \frac{1}{Z_I} \prod_{z_{-I} \in \Omega_{-I}} p(x_I z_{-I})^{\frac{1}{K_{-I}}} \qquad x_I \in \Omega_I,$$

where $Z_I := \sum_{y_I \in \Omega_I} \prod_{z_{-I} \in \Omega_{-I}} p(y_I z_{-I})^{\frac{1}{K_{-I}}}$. The first lemma shows that we can equivalently, and conveniently, take a product over all $z \in \Omega$ instead. The simple proofs for the results are omitted.

Lemma 1 *For any $\phi \neq I \subset \Gamma$, and any $x_I \in \Omega_I$,*

$$p^g(x_I) = \frac{\prod_{z \in \Omega} p(x_I z_{-I})^{\frac{1}{K}}}{\sum_{y_I \in \Omega_I} \prod_{z \in \Omega} p(y_I z_{-I})^{\frac{1}{K}}}.$$

Recall that the marginal of a marginal defines a corresponding marginal of the original probability measure. The next lemma shows that the corresponding property holds for geometric marginals

Lemma 2 *For any $\phi \neq J \subset \Gamma$ and $j \in J$ let $I = J \backslash \{j\}$. Then for any $x_I \in \Omega_I$,*

$$p^g(x_I) = \frac{\prod_{z_j \in \Omega_j} p^g(x_I z_j)^{\frac{1}{K_j}}}{\sum_{y_I \in \Omega_I} \prod_{z_j \in \Omega_j} p^g(y_I z_j)^{\frac{1}{K_j}}} \qquad x_I \in \Omega_I.$$

Say that p exhibits statistical independence if $p(x) = \prod_{i \in \Gamma} p^m(x_i)$ for all $x \in \Omega$. Write $p^m(x_I) = \prod_{i \in I} p^m(x_i)$ for any $\phi \neq I \subset \Gamma$. The next lemma shows that marginals and geo-marginals coincide under statistical independence.

Lemma 3 *If p exhibits statistical independence then for any $\phi \neq I \subset \Gamma$ and $x \in \Omega$*

$$p^g(x_I) = p^m(x_I).$$

In particular

$$p^g(x_I) = \prod_{i \in I} p^g(x_i)$$

[12] Spiegler (2016)'s agent is fully Bayesian except that she draws from data using a possibly incorrect causal theory, whereas Gilboa and Schmeidler (2001, pg 9) write "our main interest is in descriptive and normative models of conscious decisions made by humans".

B Appendix: Uniqueness Theorem for Intuitive Beliefs

The proof of Theorem 1 relies on an understanding of the uniqueness properties of the representation. We present the uniqueness result here as a separate theorem since it is of independent interest.

B.1 Basic Identification Result

Consider any Intuitive Belief p and any representation (a, b).

Lemma 4 *For all $x, y, z \in \Omega$ and any $i \in \Gamma$,*

(i) $\frac{p(x)}{p(y)} = exp\left[\sum_{i<j}[a(x_i x_j) - a(y_i y_j)] + \sum_i [b(x_i) - b(y_i)]\right]$

(ii) $\frac{p(x_i z_{-i})}{p(y_i z_{-i})} = exp\left[\sum_{i \neq j \in \Gamma}[a(x_i z_j) - a(y_i z_j)] + [b(x_i) - b(y_i)]\right]$.

Proof. The first claim follows trivially from the model. The second follows just as easily: for any $x_i z_{-i}, y_i z_{-i} \in \Omega$,

$$\frac{p(x_i z_{-i})}{p(y_i z_{-i})} = exp[[\sum_{i \neq j \in \Gamma} a(x_i z_j) + \sum_{i \neq k < j \neq i} a(z_k z_j)] - [\sum_{i \neq j \in \Gamma} a(y_i z_j) + \sum_{i \neq k < j \neq i} a(z_k z_j)] + b(x_i) - b(y_i) + \sum_{i \neq j \in \Gamma} [b(z_i) - b(z_i)]$$

$$= exp\left[\sum_{i \neq j \in \Gamma} [a(x_i z_j) - a(y_i z_j) + b(x_i) - b(y_i)]\right],$$

as desired. ■

The next key result provides the connection between a and p in any representation.

Lemma 5 *For any Intuitive Belief p and any $x, z \in \Omega$,*

$$\frac{exp[a(x_i x_j) + a(z_i z_j)]}{exp[a(x_i z_j) + a(z_i x_j)]} = \frac{p(x_i x_j z_{-ij}) p(z)}{p(x_i z_{-i}) p(x_j z_{-j})}.$$

Proof. By the representation and the previous lemma, for any $x, z \in \Omega$ and $i, j \in \Gamma$,

$$\frac{p(x_i x_j z_{-ij})}{p(z_i z_j z_{-ij})} = exp[[a(x_i x_j) - a(z_i z_j)] + \left[\sum_{ij \neq k \in \Gamma} a(x_i z_k) - a(z_i z_k)\right] + \left[\sum_{ij \neq k \in \Gamma} a(x_j z_k) - a(z_j z_k)\right]$$

$$+ [b(x_i) - b(z_i) + b(x_j) - b(z_j)]]$$

$$= exp[[a(x_i x_j) - a(z_i z_j) - [a(x_i z_j) - a(z_i z_j)] - [a(x_j z_i) - a(z_j z_i)]]$$

$$+ \left[\sum_{i \neq k \in \Gamma} [a(x_i z_k) - a(z_i z_k)] + [b(x_i) - b(z_i)]\right] + \left[\sum_{j \neq k \in \Gamma} [a(x_j z_k) - a(z_j z_k)] + [b(x_j) - b(z_j)]\right]$$

$$= exp\left[a(x_i x_j) - a(x_i z_j) - a(x_j z_i) + a(z_j z_i)\right] \times \frac{p(x_i z_{-i})}{p(z_i z_{-i})} \frac{p(x_j z_{-j})}{p(z_j z_{-j})}$$

$$= \frac{exp[a(x_i x_j) + a(z_i z_j)]}{exp[a(x_i z_j) + a(z_i x_j)]} \times \frac{p(x_i z_{-i})}{p(z_i z_{-i})} \frac{p(x_j z_{-j})}{p(z_j z_{-j})}$$

which rearranges to

$$\frac{p(x_i x_j z_{-ij})}{p(x_i z_{-i})p(x_j z_{-j})} = \frac{p(z_i z_j z_{-ij})}{p(z_i z_{-i})p(z_j z_{-j})} \times \frac{exp[a(x_i x_j) + a(z_i z_j)]}{exp[a(x_i z_j) + a(z_i x_j)]}$$

$$= \frac{1}{p(z|\Omega)} \times \frac{exp[a(x_i x_j) + a(z_i z_j)]}{exp[a(x_i z_j) + a(z_i x_j)]},$$

thereby establishing the claim. ■

B.2 Uniqueness Theorem

We will often use the following convenient fact without explicit reference to it. Adopt the notation $\prod_{i<j\leq m} := \prod_{i,j:i<j\leq m}$. Take any scalars c_{ij} satisfying $c_{ij} = c_{ji}$.

Lemma 6 *For any $m \geq 2$ and strictly positive scalars c_{ij} defined for each distinct $i, j \leq m$, it must be that*

$$\prod_{i<j\leq m} c_{ij} = \prod_{i\leq m} \prod_{i\neq j\leq m} c_{ij}^{0.5}.$$

Proof. The term $\prod_{i<j\leq m} c_{ij}$ takes each i and multiplies the terms c_{ij} where $j > i$, and then takes the product over all i. We obtain the same expression if we take each i and multiply the terms $c_{ij}^{0.5}$ with all $j \neq i$, and then take the product over all i. ■

The identification theorem stated next starts with a representation (a, b) and shows to define all other representations from it. Suppose there is a "budget" k_i for every dimension $i \in \Gamma$. For any $x_i \in \Omega_i$, allocate a (possibly negative) amount $\zeta(x_i)$ to the background association to obtain a "shifted" background association (7). From the remaining balance, $k_i - \zeta(x_i)$, allocate a portion $\gamma_j(x_i)$ to each j distinct from i, in a manner that balances the budget: $\sum_{i\neq j\in\Gamma} \gamma_j(x_i) = k_i - \zeta(x_i)$. For any $x_i x_j$, "shift" the associative weight $a(x_i x_j)$ by $\gamma_j(x_i) + \gamma_i(x_j)$ to obtain a new associative weight (6). The Theorem reveals that (α, β) obtained in this way represents p. Conversely, any two representations must be related by such budget-balanced shifts.

Theorem 5 *Consider Intuitive Beliefs p with full support represented by some (a, b). Then p is also represented by (α, β) iff there exist real-valued functions $\gamma_j(x_i)$, $\zeta(x_i)$ and k_i satisfying*

$$\zeta(x_i) + \sum_{i\neq j\in\Gamma} \gamma_j(x_i) = k_i,$$

for all $i \in \Gamma$ and $x_i \in \Omega_i$, such that for any $x \in \Omega$ and any distinct $i, j \in \Gamma$,

$$\alpha(x_i x_j) = a(x_i x_j) + \gamma_j(x_i) + \gamma_i(x_j), \tag{6}$$

and

$$\beta(x_i) = b(x_i) + \zeta(x_i). \tag{7}$$

Proof. Suppose p has full support and that it is represented by a and α. We establish the proof of sufficiency in Steps 1-4:

Step 1: Show that p is represented by (a, b) iff it is represented by $(\alpha, 0)$ where

$$\alpha(x_i x_j) = a(x_i x_j) + \frac{1}{(N-1)}[b(x_i) + b(x_j)].$$

To establish this observe that
$\sum_{i<j} \alpha(x_i x_j) + \sum_i 0$

$= \sum_{i<j} a(x_ix_j) + \sum_{i<j} \frac{1}{(N-1)}[b(x_i) + b(x_j)]$
$= \sum_{i<j} a(x_ix_j) + \frac{1}{(N-1)}\frac{1}{2}\sum_i \sum_{i\neq j\in\Gamma}[b(x_i) + b(x_j)]$
$= \sum_{i<j} a(x_ix_j) + \frac{1}{(N-1)}\frac{1}{2}\sum_i[(N-1)b(x_i) + \sum_{i\neq j\in\Gamma} b(x_j)]$
$= \sum_{i<j} a(x_ix_j) + \frac{1}{(N-1)}\frac{1}{2}\sum_i[(N-2)b(x_i) + \sum_{j\in\Gamma} b(x_j)]$
$= \sum_{i<j} a(x_ix_j) + \frac{1}{(N-1)}\frac{1}{2}[(N-2)\sum_i b(x_i) + N\sum_{j\in\Gamma} b(x_j)]$
$= \sum_{i<j} a(x_ix_j) + \frac{1}{(N-1)}\frac{1}{2}[2(N-1)\sum_i b(x_i)]$
$= \sum_{i<j} a(x_ix_j) + \sum_i b(x_i)$.

Therefore $exp[\sum_{i<j}\alpha(x_ix_j)] = exp[\sum_{i<j} a(x_ix_j) + \sum_i b(x_i)]$. Moreover, $\Sigma_{y\in\Omega} exp[\sum_{i<j}\alpha(y_iy_j)] = \Sigma_{y\in\Omega} exp[\sum_{i<j} a(y_iy_j) + \sum_i b(y_i)]$. Consequently $\frac{exp[\sum_{i<j} a(x_ix_j)+\sum_i b(x_i)]}{\Sigma_{y\in\Omega} exp[\sum_{i<j} a(y_iy_j)+\sum_i b(y_i)]} = p(x|\Omega) = \frac{\sum_{i<j}\alpha(x_ix_j)}{\Sigma_{y\in\Omega} exp[\sum_{i<j}\alpha(y_iy_j)]}$, establishing the result.

Step 2: Show that, if p is represented by $(a, 0)$ and $(\alpha, 0)$ then there exists a function $(x, i, j) \mapsto \gamma_j^*(x_i)$ s.t. for any $x \in \Omega$ and $i, j \in \Gamma$,

$$a(x_ix_j) = \alpha(x_ix_j) + \gamma_j^*(x_i) + \gamma_i^*(x_j).$$

Suppose p is represented by $(a, 0)$ and $(\alpha, 0)$. Fix any $\bar{z} \in \Omega$ and for any $x \in \Omega$ define

$$\gamma_j^*(x_i) = a(x_i\bar{z}_j) - \alpha(x_i\bar{z}_j) - \frac{1}{2}[a(\bar{z}_i\bar{z}_j) - \alpha(\bar{z}_i\bar{z}_j)],$$

Since p has full support, the network must be real-valued. By Lemma 5,

$$\frac{exp\,[a(x_ix_j) + a(\bar{z}_i\bar{z}_j)]}{exp\,[a(x_i\bar{z}_j) + a(\bar{z}_ix_j)]} = \frac{p(x_ix_j\bar{z}_{-ij})p(\bar{z})}{p(x_i\bar{z}_{-i})p(x_j\bar{z}_{-j})} = \frac{exp\,[\alpha(x_ix_j) + \alpha(\bar{z}_i\bar{z}_j)]}{exp\,[\alpha(x_i\bar{z}_j) + \alpha(\bar{z}_ix_j)]]},$$

and it follows that

$$a(x_ix_j) = \alpha(x_ix_j) + [a(x_i\bar{z}_j) - \alpha(x_i\bar{z}_j)] + [a(\bar{z}_ix_j) - \alpha(\bar{z}_ix_j)] - [a(\bar{z}_i\bar{z}_j) - \alpha(\bar{z}_i\bar{z}_j)]$$

$$= \alpha(x_ix_j) + [a(x_i\bar{z}_j) - \alpha(x_i\bar{z}_j)] + [a(\bar{z}_ix_j) - \alpha(\bar{z}_ix_j)]$$

$$-\frac{1}{2}[a(\bar{z}_i\bar{z}_j) - \alpha(\bar{z}_i\bar{z}_j)] - \frac{1}{2}[a(\bar{z}_j\bar{z}_i) - \alpha(\bar{z}_j\bar{z}_i)]$$

$$= \alpha(x_ix_j) + \gamma_j^*(x_i) + \gamma_i^*(x_j),$$

where we exploited the symmetry of a.

Step 3: Show that if p is represented by (a, b) and (α, β) then there exists $(i, j, x_i) \mapsto \gamma_j(x_i)$ and $(i, x_i) \mapsto \zeta(x_i)$ such that

$$a(x_ix_j) = \alpha(x_ix_j) + \gamma_j(x_i) + \gamma_i(x_j) \text{ and } \beta(x_i) = b(x_i) + \zeta(x_i).$$

By Steps 1 and 2, we obtain the expression

$$\alpha(x_ix_j) + \frac{1}{(N-1)}[\beta(x_i) + \beta(x_j)] = a(x_ix_j) + \frac{1}{(N-1)}[b(x_i) + b(x_j)] + \gamma_j^*(x_i) + \gamma_i^*(x_j).$$

Defining $\gamma_j(x_i) := -\gamma_j^*(x_i) - \frac{1}{(N-1)}b(x_i) + \frac{1}{(N-1)}\beta(x_i)$ yields the first claim, that is, $a(x_ix_j) = \alpha(x_ix_j) + \gamma_j(x_i) + \gamma_i(x_j)$.

For the second claim, fix any $\bar{z} \in \Omega$. For any $i \in \Gamma$ fix some $j_i \in \Gamma$ distinct from i. Take any $x_i \in \Omega_i$. Letting $x_{j_i} = \bar{z}_{j_i}$ the above expression rearranges to

$$\beta(x_i) = b(x_i) + (N-1)[a(x_i\bar{z}_{j_i}) - \alpha(x_i\bar{z}_{j_i})] + [b(\bar{z}_{j_i}) - \beta(\bar{z}_{j_i})] + (N-1)[\gamma_j^*(x_i) + \gamma_i^*(\bar{z}_{j_i})].$$

Define $\zeta(x_i) = (N-1)[a(x_i\bar{z}_{j_i}) - \alpha(x_i\bar{z}_{j_i})] + [b(\bar{z}_{j_i}) - \beta(\bar{z}_{j_i})] + (N-1)[\gamma_j^*(x_i) + \gamma_i^*(\bar{z}_{j_i})]$, an expression that depends only on i and x_i. Then $\beta(x_i) = b(x_i) + \zeta(x_i)$, as desired.

Step 4 : Show that for each $i \in \Gamma$ there is k_i s.t. for each $x_i \in \Omega_i$,

$$\zeta(x_i) + \sum_{i \neq j \in \Gamma} \gamma_j(x_i) = k_i,$$

Suppose as in Step 3 that p is represented by both (a, b) and (α, β). Begin by defining, for each x and i, the quantities

$$K(x_i) := \zeta(x_i) + \sum_{i \neq j \in \Gamma} \gamma_j(x_i), \text{ and } K(x) = \sum_{i \in \Gamma} K(x_i).$$

Observe that for any $x \in \Omega$,
$\sum_{i<j} a(x_ix_j) + \sum_i b(x_i)$
$= \sum_{i<j}[\alpha(x_ix_j) + \gamma_j(x_i) + \gamma_i(x_j)] + \sum_i[\beta(x_i) + \zeta(x_i)]$
$= \sum_{i<j}[\alpha(x_ix_j)] + \sum_{i<j}[\gamma_j(x_i) + \gamma_i(x_j)] + \sum_i[\beta(x_i) + \zeta(x_i)]$
$= \sum_{i<j}[\alpha(x_ix_j)] + \sum_{i\in\Gamma}\sum_{i\neq j\in\Gamma}\gamma_j(x_i)] + \sum_i[\beta(x_i) + \zeta(x_i)]$
$= \sum_{i<j}[\alpha(x_ix_j)] + \sum_i \beta(x_i) + \sum_{i\in\Gamma}[\zeta(x_i) + \sum_{i\neq j\in\Gamma}\gamma_j(x_i)]$
$= \sum_{i<j}[\alpha(x_ix_j) + \sum_i \beta(x_i) + \sum_{i\in\Gamma} K(x_i)$, that is,

$$\sum_{i<j} a(x_ix_j) + \sum_i b(x_i) = \sum_i \beta(x_i) + \sum_{i<j}[\alpha(x_ix_j)] + K(x).$$

Given this observation, note that since p is represented by both (a, b) and (α, β) , we have that for all $x \in \Omega$,

$$\frac{exp\left[\sum_{i<j}\alpha(x_ix_j) + \sum_i \beta(x_i)\right]}{\sum_{y\in\Omega} exp\left[\sum_{i<j}\alpha(y_iy_j) + \sum_i \beta(y_i)\right]} = p(x) = \frac{exp\left[\sum_{i<j} a(x_ix_j) + \sum_i b(x_i)\right]}{\sum_{y\in\Omega} exp\left[\sum_{i<j} a(y_iy_j) + \sum_i b(y_i)\right]}$$

$$= \frac{exp\left[\sum_{i<j}\alpha(x_ix_j) + \sum_i \beta(x_i) + K(x)\right]}{\sum_{y\in\Omega} exp\left[\sum_{i<j}\alpha(y_iy_j) + \sum_i \beta(y_i) + K(y)\right]}$$

$$= exp[K(x)] \times \frac{exp\left[\sum_{i<j}\alpha(x_ix_j) + \sum_i \beta(x_i)\right]}{\sum_{y\in\Omega} exp\left[\sum_{i<j}\alpha(y_iy_j) + \sum_i \beta(y_i) + K(y)\right]}.$$

Looking at the first and last terms in this sequence of equalities, we see that for all $x \in \Omega$,

$$exp[K(x)] = \frac{\sum_{y\in\Omega} exp\left[\sum_{i<j}\alpha(y_iy_j) + \sum_i \beta(y_i) + K(y)\right]}{\sum_{y\in\Omega} exp\left[\sum_{i<j}\alpha(y_iy_j) + \sum_i \beta(y_i)\right]}.$$

Since the RHS is independent of x, it follows that $K(x)$ is independent of x. Therefore, $K(x) = K(y)$ for all x, y.

Finally, for any x and i, consider $x_i\bar{z}_{-i}$ and observe that
$K(x_i\bar{z}_{-i}) = K(\bar{z})$
$\implies K(x_i) + \sum_{i\neq j\in\Gamma} K(\bar{z}_j) = K(\bar{z}_i) + \sum_{i\neq j\in\Gamma} K(\bar{z}_j)$
$\implies K(x_i) = K(\bar{z}_i)$
$\implies \zeta(x_i) + \sum_{i\neq j\in\Gamma}\gamma_j(x_i) = \zeta(\bar{z}_i) + \sum_{i\neq j\in\Gamma}\gamma_j(\bar{z}_i)$.
Define $k_i = \zeta(\bar{z}_i) + \sum_{i\neq j\in\Gamma}\gamma_j(\bar{z}_i)$ to complete the step.

Step 5: Establish necessity.

Suppose that (a, b) represents p and consider (α, β) as in the desired statement. Observe that
$\sum_{i<j} a(x_i x_j) + \sum_i b(x_i)$
$= \sum_{i<j}[\alpha(x_i x_j) + \gamma_j(x_i) + \gamma_i(x_j)] + \sum_i [\beta(x_i) + \zeta(x_i)]$
$= \sum_{i<j} \alpha(x_i x_j) + \sum_i \beta(x_i) + \sum_{i<j}[\gamma_j(x_i)] + \sum_{i<j}[\gamma_i(x_j)] + \sum_i \zeta(x_i)$
$= \sum_{i<j} \alpha(x_i x_j) + \sum_i \beta(x_i) + \sum_{i\in\Gamma} \sum_{i\neq j\in\Gamma}[\gamma_j(x_i)] + \sum_i \zeta(x_i)$
$= \sum_{i<j} \alpha(x_i x_j) + \sum_i \beta(x_i) + \sum_{i\in\Gamma} k_i$, that is, letting $K := \sum_{i\in\Gamma} k_i$, it must be that

$$\sum_{i<j} a(x_i x_j) + \sum_i b(x_i) = [\sum_{i<j} \alpha(x_i x_j) + \sum_i \beta(x_i)] + K.$$

It is then straightforward to see that

$$p(x) = \frac{exp\left[\sum_{i<j} a(x_i x_j) + \sum_i b(x_i)\right]}{\sum_{y\in\Omega} exp\left[\sum_{i<j} a(y_i y_j) + \sum_i b(y_i)\right]}$$

$$= \frac{exp\left[\sum_{i<j} \alpha(x_i x_j) + \sum_i \beta(x_i) + K\right]}{\sum_{y\in\Omega} exp\left[\sum_{i<j} \alpha(y_i y_j) + \sum_i \beta(y_i) + K\right]}$$

$$= \frac{exp\left[\sum_{i<j} \alpha(x_i x_j) + \sum_i \beta(x_i)\right]}{\sum_{y\in\Omega} exp\left[\sum_{i<j} \alpha(y_i y_j) + \sum_i \beta(y_i)\right]},$$

thereby establishing that (α, β) represents p, as was to be shown. ■

C Appendix: Proof of Theorem 1

We explore several classes of normalized representations using Theorem 5.

C.1 Normalization 1

Lemma 7 *Fixing any $\overline{z} \in \Omega$, an Intuitive Belief representation (α, β) can be normalized by (a) setting $\beta = 0$, (b) setting $\alpha(\overline{z}_i \overline{z}_j) = 0$ for all distinct $i, j \in \Gamma$ and (c) setting $\alpha(x_i \overline{z}_j) = \alpha(x_i \overline{z}_{j'})$ for all $i \neq j, j' \in \Gamma$ and $x \in \Omega$. Such a normalized representation is unique, given $\overline{z}$. Moreover, $p(\overline{z}) \times Z = 1$ holds in the representation normalized wrt $\overline{z}$.*

Proof. Take any representation (a, b), and fix some $\overline{z} \in \Omega$. Define $\zeta(x_i)$ by

$$\zeta(x_i) = -b(x_i)$$

and

$$\gamma_j(x_i) = -\frac{1}{N-1}[\zeta(x_i) - \zeta(\overline{z}_i)] + \frac{1}{2}a(\overline{z}_i \overline{z}_j) - a(x_i \overline{z}_j) + \frac{1}{N-1}\sum_{i\neq k\in\Gamma}[a(x_i \overline{z}_k) - a(\overline{z}_i \overline{z}_k)].$$

Note that
$\zeta(x_i) + \sum_{i\neq j\in\Gamma} \gamma_j(x_i)$
$= \zeta(\overline{z}_i) + \frac{1}{2}\sum_{i\neq j\in\Gamma} a(\overline{z}_i \overline{z}_j) - \sum_{i\neq j\in\Gamma} a(x_i \overline{z}_j) + \sum_{i\neq k\in\Gamma}[a(x_i \overline{z}_k) - a(\overline{z}_i \overline{z}_k)]$
$= \zeta(\overline{z}_i) - \frac{1}{2}\sum_{i\neq j\in\Gamma} a(\overline{z}_i \overline{z}_j) = k_i$. In particular, ζ, γ along with $k_i := \zeta(\overline{z}_i) - \frac{1}{2}\sum_{i\neq j\in\Gamma} a(\overline{z}_i \overline{z}_j)$ satisfy the desired properties in Theorem 5.

Step 1: Show that p is represented by (α, β) defined by $\beta(x_i) = 0$ and

$$\alpha(x_i x_j) = [a(x_i x_j) + a(\overline{z}_i \overline{z}_j) - a(x_i \overline{z}_j) - a(x_j \overline{z}_i)]$$

$$+\frac{1}{N-1}\left[\sum_{i \neq k \in \Gamma} [a(x_i \overline{z}_k) - a(\overline{z}_i \overline{z}_k)] + \sum_{j \neq k \in \Gamma} [a(x_j \overline{z}_k) - a(\overline{z}_j \overline{z}_k)] - [\zeta(x_i) - \zeta(\overline{z}_i) + \zeta(x_j) - \zeta(\overline{z}_j)]\right].$$

By Theorem 5 we obtain a new representation (α, β) defined by:

$$\beta(x_i) = b(x_i) + \zeta(x_i) = 0$$

and
$\alpha(x_i x_j) = a(x_i x_j) + \gamma_j(x_i) + \gamma_i(x_j)$
$= a(x_i x_j) - \frac{1}{N-1}[\zeta(x_i) - \zeta(\overline{z}_i)] + \frac{1}{2} a(\overline{z}_i \overline{z}_j) - a(x_i \overline{z}_j) + \frac{1}{N-1} \sum_{i \neq k \in \Gamma} [a(x_i \overline{z}_k) - a(\overline{z}_i \overline{z}_k)]$
$-\frac{1}{N-1}[\zeta(x_j) - \zeta(\overline{z}_j)] + \frac{1}{2} a(\overline{z}_j \overline{z}_i) - a(x_j \overline{z}_i) + \frac{1}{N-1} \sum_{j \neq k \in \Gamma} [a(x_j \overline{z}_k) - a(\overline{z}_j \overline{z}_k)]$
$= a(x_i x_j) - \frac{1}{N-1}[\zeta(x_i) - \zeta(\overline{z}_i) + \zeta(x_j) - \zeta(\overline{z}_j)] + a(\overline{z}_i \overline{z}_j) - a(x_i \overline{z}_j) - a(x_j \overline{z}_i)$
$+\frac{1}{N-1} \sum_{i \neq k \in \Gamma} [a(x_i \overline{z}_k) - a(\overline{z}_i \overline{z}_k)] + \frac{1}{N-1} \sum_{j \neq k \in \Gamma} [a(x_j \overline{z}_k) - a(\overline{z}_j \overline{z}_k)]$

Step 2: Show that α satisfies for all distinct $i, j, k \in \Gamma$ and $x_i \in \Omega_i$,

$$\alpha(\overline{z}_i \overline{z}_j) = 0 \text{ and } \alpha(x_i \overline{z}_j) = \alpha(x_i \overline{z}_k).$$

Use the expression for α in Step 1 to obtain $\alpha(\overline{z}_i \overline{z}_j) = 2a(\overline{z}_i \overline{z}_j) - 2a(\overline{z}_i \overline{z}_j) + 0 = 0$, and
$\alpha(x_i \overline{z}_j) = a(x_i \overline{z}_j) - \frac{1}{N-1}[\zeta(x_i) - \zeta(\overline{z}_i) + \zeta(\overline{z}_j) - \zeta(\overline{z}_j)] + a(\overline{z}_i \overline{z}_j) - a(x_i \overline{z}_j) - a(\overline{z}_j \overline{z}_i)$
$+\frac{1}{N-1} \sum_{i \neq k \in \Gamma} [a(x_i \overline{z}_k) - a(\overline{z}_i \overline{z}_k)] + \frac{1}{N-1} \sum_{j \neq k \in \Gamma} [a(\overline{z}_j \overline{z}_k) - a(\overline{z}_j \overline{z}_k)]$
$= -\frac{1}{N-1}[\zeta(x_i) - \zeta(\overline{z}_i)] + \frac{1}{N-1} \sum_{i \neq k \in \Gamma} [a(x_i \overline{z}_k) - a(\overline{z}_i \overline{z}_k)]$, which does not depend on j.

Step 3: Conclusion.

We have thus shown that there always exists a normalized representation as stated in the proposition. Note that in a normalized representation (α, β), since $\alpha(\overline{z}_i \overline{z}_j) = \beta(\overline{z}_i) = 0$ and so

$$p(\overline{z}) = \frac{1}{Z} exp[\sum_{i<j} \alpha(\overline{z}_i \overline{z}_j) + \sum_i \beta(\overline{z}_i)] = \frac{1}{Z},$$

that is, $p(\overline{z}) = \frac{1}{Z}$.

In Lemma 8 we show that α can be written in terms of p. Therefore the uniqueness of the representation is a corollary of that result. ■

Lemma 8 *A network $(\alpha, 0)$ is a $\overline{z}$-normalized representation for p if and only if for all distinct $i, j \in \Gamma$,*

$$exp[\alpha(x_i x_j)] = p(\overline{z})^{\frac{N-3}{N-1}} \frac{p(x_i x_j \overline{z}_{-ij})}{[p(x_i \overline{z}_{-i}) p(x_j \overline{z}_{-j})]^{\frac{N-2}{N-1}}} \qquad (x_i, x_j) \in \Omega_{ij}.$$

Proof. If $(\alpha, 0)$ satisfies the noted expression then it is readily determined that $\alpha(\overline{z}_i \overline{z}_j) = 0$ for all distinct $i, j \in \Gamma$ and $\alpha(x_i \overline{z}_j) = \alpha(x_i \overline{z}_{j'})$ for all $i \neq j, j' \in \Gamma$ and $x \in \Omega$. Thus $(\alpha, 0)$ is a $\overline{z}$-normalized representation.

Conversely, take any $\overline{z}$-normalized representation $(\alpha, 0)$ for p. By Lemma 7, $Z \times p(\overline{z}) = 1$. Since $\alpha(\overline{z}_j \overline{z}_k) = 0$ for all distinct $j, k \in \Gamma$ and that $a(x_i \overline{z}_k)$ is independent of k, we see that
$p(x_i \overline{z}_{-i}) = \frac{1}{Z} \times exp\left[\sum_{i \neq k \in \Gamma} \alpha(x_i \overline{z}_k) + \sum_{i \neq j < k \neq i} \alpha(\overline{z}_j \overline{z}_k)\right]$
$= p(\overline{z}) \times exp\left[\sum_{i \neq k \in \Gamma} \alpha(x_i \overline{z}_k)\right]$

$= p(\overline{z}) \times exp\left[(N-1)\alpha(x_i\overline{z}_j)\right]$ for any $i \neq j \in \Gamma$, we obtain the property that

$$\alpha(x_i\overline{z}_j) = \frac{1}{N-1} ln[\frac{p(x_i\overline{z}_{-i})}{p(\overline{z})}],$$

for all $i \neq j \in \Gamma$. Finally, observe that, by Lemma 5, the representation must satisfy:

$$\frac{p(x_ix_j\overline{z}_{-ij})p(\overline{z})}{p(x_i\overline{z}_{-i})p(x_j\overline{z}_{-j})} = \frac{exp[\alpha(x_ix_j) + \alpha(\overline{z}_i\overline{z}_j)]}{exp[\alpha(x_i\overline{z}_j) + \alpha(\overline{z}_ix_j)]} = \frac{exp[\alpha(x_ix_j)]}{[\frac{p(x_i\overline{z}_{-i})}{p(\overline{z})}]^{\frac{1}{N-1}}[\frac{p(x_j\overline{z}_{-j})}{p(\overline{z})}]^{\frac{1}{N-1}}},$$

and so

$$\begin{aligned} exp[\alpha(x_ix_j)] &= \frac{p(x_ix_j\overline{z}_{-ij})}{p(x_i\overline{z}_{-i})p(x_j\overline{z}_{-j})} \times p(\overline{z}) \times [\frac{p(x_i\overline{z}_{-i})}{p(\overline{z})}\frac{p(x_j\overline{z}_{-j})}{p(\overline{z})}]^{\frac{1}{N-1}} \\ &= p(\overline{z})^{\frac{N-3}{N-1}} \frac{p(x_ix_j\overline{z}_{-ij})}{[p(x_i\overline{z}_{-i})p(x_j\overline{z}_{-j})]^{\frac{N-2}{N-1}}}, \end{aligned}$$

as desired. ■

C.2 Normalization 2

Lemma 9 *A belief p with full support is an Intuitive Belief if and only if, for any $\overline{z} \in \Omega$, it is represented by an associative network $(a_{\overline{z}}, b_{\overline{z}})$ with the property that for any $x_i, x_j \in \cup_{k\in\Gamma}\Omega_k$,*

$$exp[a_{\overline{z}}(x_ix_j)] = \frac{p(x_ix_j\overline{z}_{-ij})p(\overline{z})}{p(x_i\overline{z}_{-i})p(x_j\overline{z}_{-j})} \text{ and } exp[b_{\overline{z}}(x_i)] = \frac{p(x_i\overline{z}_{-i})}{p(\overline{z})}.$$

Indeed, for any $\overline{z} \in \Omega$, there exists a representation $(a_{\overline{z}}, b_{\overline{z}})$ where $a_{\overline{z}}(\overline{z}_i\overline{z}_j) = a_{\overline{z}}(x_i\overline{z}_j) = b_{\overline{z}}(\overline{z}_i) = 0$. Moreover, this representation gives rise to the reduced form

$$p(x) = p(\overline{z}) \times \left[\prod_{i<j} \frac{p(x_ix_j\overline{z}_{-ij})p(\overline{z})}{p(x_i\overline{z}_{-i})p(x_j\overline{z}_{-j})}\right] \times \prod_{i\in\Gamma} \frac{p(x_i\overline{z}_{-i})}{p(\overline{z})}, \quad x \in \Omega.$$

Proof. If there exists such a representation then p is trivially an Intuitive Belief. Conversely, suppose p is an Intuitive Belief. Consider a normalized representation for p and re-write the expression in Lemma 8 as:

$$exp[\alpha(x_ix_j)] = \frac{\frac{p(x_ix_j\overline{z}_{-ij})}{p(\overline{z})}}{\frac{p(x_i\overline{z}_{-i})}{p(\overline{z})}\frac{p(x_j\overline{z}_{-j})}{p(\overline{z})}}[\frac{p(x_i\overline{z}_{-i})}{p(\overline{z})}]^{\frac{1}{N-1}}[\frac{p(x_j\overline{z}_{-j})}{p(\overline{z})}]^{\frac{1}{N-1}}.$$

Insert this into the representation, and redefining the scaling factor as needed, observe that:

$p(x) = \frac{1}{Z} \times \prod_{i<j} exp[a(x_ix_j)]$

$= \frac{1}{Z} \times \prod_{i<j} \frac{\frac{p(x_ix_j\overline{z}_{-ij})}{p(\overline{z})}}{\frac{p(x_i\overline{z}_{-i})}{p(\overline{z})}\frac{p(x_j\overline{z}_{-j})}{p(\overline{z})}}[\frac{p(x_i\overline{z}_{-i})}{p(\overline{z})}]^{\frac{1}{N-1}}[\frac{p(x_j\overline{z}_{-j})}{p(\overline{z})}]^{\frac{1}{N-1}}$

$= \frac{1}{Z} \times \left[\prod_{i<j} \frac{p(x_ix_j\overline{z}_{-ij})p(\overline{z})}{p(x_i\overline{z}_{-i})p(x_j\overline{z}_{-j})}\right] \times \prod_{i<j}[\frac{p(x_i\overline{z}_{-i})}{p(\overline{z})}]^{\frac{1}{N-1}}[\frac{p(x_j\overline{z}_{-j})}{p(\overline{z})}]^{\frac{1}{N-1}}$

$= \frac{1}{Z} \times \left[\prod_{i<j} \frac{p(x_ix_j\overline{z}_{-ij})p(\overline{z})}{p(x_i\overline{z}_{-i})p(x_j\overline{z}_{-j})}\right] \times \prod_{i\in\Gamma}\prod_{i\neq j\in\Gamma}[\frac{p(x_i\overline{z}_{-i})}{p(\overline{z})}]^{\frac{1}{2(N-1)}}[\frac{p(x_j\overline{z}_{-j})}{p(\overline{z})}]^{\frac{1}{2(N-1)}}$

$= \frac{1}{Z} \times \left[\prod_{i<j} \frac{p(x_ix_j\overline{z}_{-ij})p(\overline{z})}{p(x_i\overline{z}_{-i})p(x_j\overline{z}_{-j})}\right] \times \prod_{i\in\Gamma}\left[[\frac{p(x_i\overline{z}_{-i})}{p(\overline{z})}]^{\frac{N-1}{2(N-1)}} \times \left[\prod_{i\neq j\in\Gamma}[\frac{p(x_j\overline{z}_{-j})}{p(\overline{z})}]^{\frac{1}{2(N-1)}}\right]\right]$

$= \frac{1}{Z} \times \left[\prod_{i<j} \frac{p(x_ix_j\overline{z}_{-ij})p(\overline{z})}{p(x_i\overline{z}_{-i})p(x_j\overline{z}_{-j})}\right] \times \prod_{i\in\Gamma}\left[[\frac{p(x_i\overline{z}_{-i})}{p(\overline{z})}]^{\frac{N-1}{2(N-1)}} \times [\frac{p(x_i\overline{z}_{-i})}{p(\overline{z})}]^{\frac{-1}{2(N-1)}}\left[\prod_{j\in\Gamma}[\frac{p(x_j\overline{z}_{-j})}{p(\overline{z})}]^{\frac{1}{2(N-1)}}\right]\right]$

$= \frac{1}{Z} \times \left[\prod_{i<j} \frac{p(x_ix_j\overline{z}_{-ij})p(\overline{z})}{p(x_i\overline{z}_{-i})p(x_j\overline{z}_{-j})}\right] \times \left[\prod_{j\in\Gamma}[\frac{p(x_j\overline{z}_{-j})}{p(\overline{z})}]^{\frac{N}{2(N-1)}}\right] \times \prod_{i\in\Gamma}\left[[\frac{p(x_i\overline{z}_{-i})}{p(\overline{z})}]^{\frac{N-2}{2(N-1)}}\right]$

$= \frac{1}{Z} \times \left[\prod_{i<j} \frac{p(x_ix_j\overline{z}_{-ij})p(\overline{z})}{p(x_i\overline{z}_{-i})p(x_j\overline{z}_{-j})}\right] \times \left[\prod_{i\in\Gamma}[\frac{p(x_i\overline{z}_{-j})}{p(\overline{z})}]^{\frac{N}{2(N-1)}+\frac{N-2}{2(N-1)}}\right]$

$= \frac{1}{Z} \times \left[\prod_{i<j} \frac{p(x_ix_j\overline{z}_{-ij})p(\overline{z})}{p(x_i\overline{z}_{-i})p(x_j\overline{z}_{-j})}\right] \times \left[\prod_{i\in\Gamma} \frac{p(x_i\overline{z}_{-j})}{p(\overline{z})}\right]$. By Lemma 7, $\frac{1}{Z} = p(\bar{z})$. This proves that if p is an Intuitive Belief, then it must admit the representation given in the statement of the Lemma. The remaining two claims are immediate corollaries. ■

C.3 Proof of Theorem 1

Proof. Collecting the $p(\bar{z})$-terms in the reduced form established in Lemma 9 and defining $\frac{1}{Z_{\bar{z}}} = p(\bar{z})^{1+\frac{N(N-1)}{2}-N} = p(\bar{z})^{\frac{(N-1)(N-2)}{2}}$, we see that p is an Intuitive Belief if and only if for each $\bar{z} \in \Omega$, it admits the reduced form

$$p(x|\Omega) = \frac{1}{Z_{\bar{z}}} \times \left[\prod_{i<j} \frac{p(x_i x_j \bar{z}_{-ij})}{p(x_i \bar{z}_{-i}) p(x_j \bar{z}_{-j})} \right] \times \prod_{i \in \Gamma} p(x_i \bar{z}_{-i}), \quad x \in \Omega.$$

Since the reduced form for each $\bar{z} \in \Omega$ expresses the same p, so will the geometric mean of all these reduced forms. Consequently, for all $x \in \Omega$,

$$p(x) = \prod_{\bar{z} \in \Omega} \left[\frac{1}{Z_{\bar{z}}} \times \left[\prod_{i<j} \frac{p(x_i x_j \bar{z}_{-ij})}{p(x_i \bar{z}_{-i}) p(x_j \bar{z}_{-j})} \right] \times \prod_{i \in \Gamma} p(x_i \bar{z}_{-i}) \right]^{\frac{1}{K}}$$

$$= \frac{1}{Z'} \times \left[\prod_{i<j} \frac{p^h(x_i x_j)}{p^h(x_i) p^h(x_j)} \right] \times \prod_{i \in \Gamma} p^h(x_i),$$

where $\frac{1}{Z'} := \frac{1}{\prod_{\bar{z} \in \Omega} [Z_{\bar{z}}^{\frac{1}{K}}]}$ and for any $I \subset \Gamma$,

$$p^h(x_I) := \prod_{\bar{z} \in \Omega} p(x_I z_{-I})^{\frac{1}{K}}, \quad x \in \Omega.$$

Given Lemma 1, dividing this by the constant $\sum_{y_I \in \Omega_I} p^h(y_I)$ converts p^h into a geo-marginal. Inserting suitable constants into the expression for $p(x)$ and defining an appropriate scaling factor $\frac{1}{Z}$, we see that p admits a representation of the desired form. ■

D Appendix: Proof of Theorem 2

Proof. By definition, if $p(.|D)$ over D is an Intuitive Belief, then there exists an extension, that is, an Intuitive Belief p over Ω, such that $p(x|D) = \frac{p(x)}{p(D)}$ for all $x \in D$. To show that this extension is unique, consider any r over Ω that is an Intuitive Belief that also satisfies $p(x|D) = \frac{r(x)}{r(D)}$ for all $x \in D$.

Richness requires that $(x_i x_j \bar{z}_{-ij}) \in D$ for any $\bar{z} \in S$, $i, j \in \Gamma$ and any $(x_i, x_j) \in \Omega_{ij}$. Then $r(x_i x_j \bar{z}_{-ij}) = r(D) \times p(x_i x_j \bar{z}_{-ij}|D) = \frac{r(D)}{p(D)} \times p(x_i x_j \bar{z}_{-ij})$. Thus, for $k = \frac{r(D)}{p(D)} > 0$,

$$r(x_i x_j \bar{z}_{-ij}) = k \times p(x_i x_j \bar{z}_{-ij}).$$

Both $r$and p must admit the normalized representation in Lemma 9 for any $\bar{z} \in S$. Exploiting the above displayed equality, we see that for any $\bar{z} \in S$ and for all $x \in \Omega$,

$$r(x) = r(\bar{z}) \times \left[\prod_{i<j} \frac{r(x_i x_j \bar{z}_{-ij}) r(\bar{z})}{r(x_i \bar{z}_{-i}) r(x_j \bar{z}_{-j})} \right] \times \prod_{i \in \Gamma} \frac{r(x_i \bar{z}_{-i})}{r(\bar{z})}$$

$$= [k \times p(\bar{z})] \times \left[\prod_{i<j} \frac{p(x_i x_j \bar{z}_{-ij}) p(\bar{z})}{p(x_i \bar{z}_{-i}) p(x_j \bar{z}_{-j})} \right] \times \prod_{i \in \Gamma} \frac{p(x_i \bar{z}_{-i})}{p(\bar{z})} = k \times p(x),$$

that is, for all $x \in \Omega$,

$$r(x) = k \times p(x).$$

But then it must also be that $k = 1$, since $1 = \sum_{x \in \Omega} r(x) = k \sum_{x \in \Omega} p(x) = k$. Conclude that $r = p$, that is, there is a unique extension, as desired.

To obtain the desired reduced form for p over Ω, we observe that, given Lemma 9 and $p(x|D) = \frac{p(x)}{p(D)}$, it must be that for any $\overline{z} \in D$,

$$p(x) = p(\overline{z}) \times \left[\prod_{i<j} \frac{p(x_i x_j \overline{z}_{-ij}) p(\overline{z})}{p(x_i \overline{z}_{-i}) p(x_j \overline{z}_{-j})} \right] \times \prod_{i \in \Gamma} \frac{p(x_i \overline{z}_{-i})}{p(\overline{z})}$$

$$= p(D) p(\overline{z}|D) \times \left[\prod_{i<j} \frac{p(x_i x_j \overline{z}_{-ij}|D) p(\overline{z}|D)}{p(x_i \overline{z}_{-i}|D) p(x_j \overline{z}_{-j}|D)} \right] \times \prod_{i \in \Gamma} \frac{p(x_i \overline{z}_{-i}|D)}{p(\overline{z}|D)}$$

$$= \frac{1}{Z_{\overline{z}}} \times \left[\prod_{i<j} \frac{p(x_i x_j \overline{z}_{-ij}|D)}{p(x_i \overline{z}_{-i}|D) p(x_j \overline{z}_{-j}|D)} \right] \times \prod_{i \in \Gamma} p(x_i \overline{z}_{-i}|D),$$

for some $Z_{\overline{z}} > 0$. Since this holds for any $\overline{z} \in D$, taking the geometric mean of these reduced forms for $p(x)$ over all $\overline{z} \in D$ yields, for some constant $Z > 0$,

$$p(x) = \frac{1}{Z} \times \left[\prod_{i<j} \frac{p^g(x_i x_j|D)}{p^g(x_i|D) p^g(x_j|D)} \right] \times \prod_{i \in \Gamma} p^g(x_i|D) \qquad x \in \Omega,$$

as desired. To complete the proof, note that since $p(\centerdot|D)$ has full support, so does the extension p. ■

E Proof of Theorem 3

E.1 Necessity

For any given $\overline{z} \in \Omega$, distinct $i, j \in \Gamma$ and $x_i, x_j \in \cup_{k \in \Gamma} \Omega_k$, consider the ratio:

$$a_{\overline{z}}(x_i x_j) = \frac{p(x_i x_j \overline{z}_{-ij}) p(\overline{z})}{p(x_i \overline{z}_{-i}) p(x_j \overline{z}_{-j})}, \tag{8}$$

We first show a key property of this ratio:

Lemma 10 *Any Intuitive Belief p on Ω with full support satisfies* Degenerate Relative Associative Separability *(DRAS): for all $\overline{z}, \overline{w} \in \Omega$ and each distinct $i, j \in \Gamma$ and $x_i \in \Omega_i$, $x_j \in \Omega_j$*

$$\frac{p(x_i x_j \overline{z}_{-ij}) p(\overline{z}_i \overline{z}_j \overline{z}_{-ij})}{p(x_i \overline{z}_j \overline{z}_{-ij}) p(\overline{z}_i x_j \overline{z}_{-ij})} = \frac{p(x_i x_j \overline{w}_{-ij}) p(\overline{z}_i \overline{z}_j \overline{w}_{-ij})}{p(x_i \overline{z}_j \overline{w}_{-ij}) p(\overline{z}_i x_j \overline{w}_{-ij})},$$

that is, $a_{\overline{z}_{ij} \overline{z}_{-ij}}(x_i x_j) = a_{\overline{z}_{ij} \overline{w}_{-ij}}(x_i x_j)$.

Proof. by Lemma 5, for any representation (a, b) it must be that

$$\frac{exp[a(x_i x_j)]}{exp[a(x_i z_j)] \times exp[a(x_j z_i)]} = \frac{\frac{p(x_i x_j z_{-ij})}{p(z)}}{\frac{p(x_i z_{-i})}{p(z)} \frac{p(x_j z_{-j})}{p(z)}}.$$

Conclude that $\frac{\frac{p(x_i x_j z_{-ij})}{p(z)}}{\frac{p(x_i z_{-i})}{p(z)} \frac{p(x_j z_{-j})}{p(z)}}$ is independent of z_{-ij}. ■

Lemma 11 *Consider any Intuitive Belief p on Ω with full support, any set S of reference states, any distinct $i,j \in \Gamma$ and the set $D_S^{ij} = \Omega_{ij} \times S_{-ij} \subset \Delta$ of simple states wrt S. There exists a scalar $\zeta_{S_{-ij}} > 0$ s.t. for all $(x_i, x_j) \in \Omega_{ij}$,*

$$\frac{p^g(x_i x_j)}{p^g(x_i)p^g(x_j)} = \zeta_{S_{-ij}} \frac{\hat{p}^{S_{-ij}}(x_i x_j)}{\hat{p}^{S_{-ij}}(x_i)\hat{p}^{S_{-ij}}(x_j)},$$

Proof. Starting with the definition of g-PMI and applying DRAS (Lemma 10) we see that for any distinct $i,j \in \Gamma$ and $x_i, x_j \in \cup_{k\in\Gamma}\Omega_k$,

$$\frac{p^g(x_i x_j)}{p^g(x_i)p^g(x_j)} = \prod_{z\in\Omega}\left[\frac{p(x_i x_j z_{-ij})}{p(x_i z_{-i})p(x_j z_{-j})}\right]^{\frac{1}{K}} \text{ by definition of } p^g$$

$$= \frac{1}{\prod_{z\in\Omega} p(z)^{\frac{1}{K}}} \times \prod_{z\in\Omega}\left[\frac{p(x_i x_j z_{-ij})p(z)}{p(x_i z_{-i})p(x_j z_{-j})}\right]^{\frac{1}{K}}$$

$$= \frac{1}{\prod_{z\in\Omega} p(z)^{\frac{1}{K}}} \times \prod_{z\in\Omega}\left[\frac{p(x_i x_j y_{-ij})p(z_i z_j y_{-ij})}{p(x_i z_j y_{-ij})p(z_i x_j y_{-ij})}\right]^{\frac{1}{K}} \text{ for any fixed } y_{-ij} \in \Omega_{-ij}$$

$$= \frac{1}{\prod_{z\in\Omega} p(z)^{\frac{1}{K}}} \times \prod_{(z_i,z_j)\in\Omega_{ij}} \prod_{z_{-ij}\in\Omega_{ij}}\left[\frac{p(x_i x_j y_{-ij})p(z_i z_j y_{-ij})}{p(x_i z_j y_{-ij})p(z_i x_j y_{-ij})}\right]^{\frac{1}{K}}$$

$$= \frac{1}{\prod_{z\in\Omega} p(z)^{\frac{1}{K}}} \times \prod_{(z_i,z_j)\in\Omega_{ij}}\left[\frac{p(x_i x_j y_{-ij})p(z_i z_j y_{-ij})}{p(x_i z_j y_{-ij})p(z_i x_j y_{-ij})}\right]^{\frac{K_{-ij}}{K}}$$

$$= \frac{1}{\prod_{z\in\Omega} p(z)^{\frac{1}{K}}} \times \prod_{(z_i,z_j)\in\Omega_{ij}}\left[\frac{p(x_i x_j y_{-ij})p(z_i z_j y_{-ij})}{p(x_i z_j y_{-ij})p(z_i x_j y_{-ij})}\right]^{\frac{1}{K_{ij}}},$$

that is, for any fixed $y_{-ij} \in \Omega_{-ij}$,

$$\frac{p^g(x_i x_j)}{p^g(x_i)p^g(x_j)} = \frac{1}{\prod_{z\in\Omega} p(z)^{\frac{1}{K}}} \times \prod_{(z_i,z_j)\in\Omega_{ij}}\left[\frac{p(x_i x_j y_{-ij})p(z_i z_j y_{-ij})}{p(x_i z_j y_{-ij})p(z_i x_j y_{-ij})}\right]^{\frac{1}{K_{ij}}}$$

But since this holds for any $y_{-ij} \in S_{-ij}$, we preserve the equality if we take a geometric mean of the RHS expression wrt $y_{-ij} \in S_{-ij}$. Therefore, writing $K_{S_{-ij}}$ for the cardinality of S_{-ij},

$$\frac{p^g(x_i x_j)}{p^g(x_i)p^g(x_j)} = \prod_{y_{-ij}\in S_{-ij}}\left[\frac{1}{\prod_{z\in\Omega} p(z)^{\frac{1}{K}}} \times \prod_{(z_i,z_j)\in\Omega_{ij}}\left[\frac{p(x_i x_j y_{-ij})p(z_i z_j y_{-ij})}{p(x_i z_j y_{-ij})p(z_i x_j y_{-ij})}\right]^{\frac{1}{K_{ij}}}\right]^{\frac{1}{K_{S_{-ij}}}}$$

$$= \frac{1}{\prod_{z\in\Omega} p(z)^{\frac{K_{S_{-ij}}}{K}}} \times \prod_{(z_i z_j y_{ij})\in\Omega_{ij}\times S_{-ij}}\left[\frac{p(x_i x_j y_{-ij})p(z_i z_j y_{-ij})}{p(x_i z_j y_{-ij})p(z_i x_j y_{-ij})}\right]^{\frac{1}{K_{ij}K_{S_{-ij}}}}$$

$$= \frac{\prod_{(z_i z_j y_{ij})\in\Omega_{ij}\times S_{-ij}} p(z_i z_j y_{-ij})}{\prod_{z\in\Omega} p(z)^{\frac{K_{S_{-ij}}}{K}}} \times \prod_{(z_i z_j y_{ij})\in\Omega_{ij}\times S_{-ij}}\left[\frac{p(x_i x_j y_{-ij})}{p(x_i z_j y_{-ij})p(z_i x_j y_{-ij})}\right]^{\frac{1}{K(\Omega_{ij}\times S_{-ij})}}$$

$$= \left[\frac{\prod_{(z_i z_j y_{ij})\in\Omega_{ij}\times S_{-ij}} p(z_i z_j y_{-ij})}{\prod_{z\in\Omega} p(z)^{\frac{K_{S_{-ij}}}{K}}}\frac{Z_{ij}}{Z_i Z_j}\right]\frac{\hat{p}^{S_{-ij}}(x_i x_j)}{\hat{p}^{S_{-ij}}(x_i)\hat{p}^{S_{-ij}}(x_j)},$$

where Z_i, Z_j, Z_{ij} are the scalars required to convert the geo-means into geometric marginals. Defining the constant in the bracket in the last expression by $\zeta_{S_{-ij}} > 0$, we obtain the desired result. ■

E.2 Sufficiency

Recall the function $exp[a_{\overline{z}}(x_i x_j)] := \frac{p(x_i x_j \overline{z}_{-ij})p(\overline{z})}{p(x_i \overline{z}_{-i})p(x_j \overline{z}_{-j})}$ defined in (8), and the DRAS property defined in Lemma 10.

Lemma 12 *If a belief p on Ω with full support satisfies RAS, then it satisfies DRAS.*

Proof. Apply RAS to the case where there is only one reference state: $S = \{\overline{z}\}$ for some $\overline{z} \in \Omega$. Then $K(\Omega_{ij} \times \{\overline{z}_{-ij}\}) = K_{ij}$. By definition,

$$\hat{p}^{S_{-ij}}(x_i x_j) := \frac{1}{Z_{ij}} \prod_{z \in \Omega_{ij} \times \{\overline{z}_{-ij}\}} p(x_i x_j z_{-ij})^{\frac{1}{K_{ij}}}$$

$$= \frac{1}{Z_{ij}} \prod_{(z_i z_j) \in \Omega_{ij}} p(x_i x_j \overline{z}_{-ij})^{\frac{1}{K_{ij}}} = \frac{1}{Z_{ij}} p(x_i x_j \overline{z}_{-ij}),$$

and similarly,

$$\hat{p}^{S_{-ij}}(x_i) := \frac{1}{Z_i} \prod_{z \in \Omega_{ij} \times \{\overline{z}_{-ij}\}} p(x_i z_j z_{-ij})^{\frac{1}{K_{ij}}}$$

$$= \frac{1}{Z_i} \prod_{z_j \in \Omega_j} \prod_{z_i \in \Omega_i} p(x_i z_j \overline{z}_{-ij})^{\frac{1}{K_{ij}}} = \frac{1}{Z_i} \prod_{z_j \in \Omega_j} p(x_i z_j \overline{z}_{-ij})^{\frac{1}{K_j}}.$$

Define $exp[\hat{a}_{\overline{z}_{-ij}}(x_i, x_j)] := \frac{p(x_i x_j \overline{z}_{-ij})}{[\prod_{z_j \in \Omega_j} p(x_i z_j \overline{z}_{-ij})^{\frac{1}{K_j}}][\prod_{z_i \in \Omega_i} p(z_i x_j \overline{z}_{-ij})^{\frac{1}{K_i}}]}$. Then,

$$\frac{exp[\hat{a}_{\overline{z}_{-ij}}(x_i, x_j)]}{exp[\hat{a}_{\overline{z}_{-ij}}(y_i, y_j)]} = \frac{\frac{p(x_i x_j \overline{z}_{-ij})}{p(y_i y_j \overline{z}_{-ij})}}{[\prod_{z_j \in \Omega_j} [\frac{p(x_i z_j \overline{z}_{-ij})}{p(y_i z_j \overline{z}_{-ij})}]^{\frac{1}{K_j}}] \times [\prod_{z_i \in \Omega_i} \frac{p(z_i x_j \overline{z}_{-ij})}{p(z_i y_j \overline{z}_{-ij})}]^{\frac{1}{K_i}}]}.$$

In particular, compute

$$\frac{exp[\hat{a}_{\overline{z}_{-ij}}(x_i, x_j)]}{exp[\hat{a}_{\overline{z}_{-ij}}(x_i, \overline{z}_j)]} = \frac{\frac{p(x_i x_j \overline{z}_{-ij})}{p(x_i \overline{z}_j \overline{z}_{-ij})}}{[\prod_{z_i \in \Omega_i} \frac{p(z_i x_j \overline{z}_{-ij})}{p(z_i \overline{z}_j \overline{z}_{-ij})}]^{\frac{1}{K_i}}]} \text{ and } \frac{exp[\hat{a}_{\overline{z}_{-ij}}(\overline{z}_i, \overline{z}_j)]}{exp[\hat{a}_{\overline{z}_{-ij}}(\overline{z}_i, x_j)]} = \frac{\frac{p(\overline{z}_i \overline{z}_j \overline{z}_{-ij})}{p(\overline{z}_i x_j \overline{z}_{-ij})}}{[\prod_{z_i \in \Omega_i} \frac{p(z_i x_j \overline{z}_{-ij})}{p(z_i \overline{z}_j \overline{z}_{-ij})}]^{\frac{1}{K_i}}]}.$$

Recalling the canonical representation (a_g, b_g) (Definition 4), by RAS, it follows that

$$\frac{exp[a_g(x_i, x_j)]}{exp[a_g(x_i, \overline{z}_j)]} \times \frac{exp[a_g(\overline{z}_i, \overline{z}_j)]}{exp[a_g(\overline{z}_i, x_j)]} = \frac{exp[\hat{a}_{\overline{z}_{-ij}}(x_i, x_j)]}{exp[\hat{a}_{\overline{z}_{-ij}}(x_i, \overline{z}_j)]} \times \frac{exp[\hat{a}_{\overline{z}_{-ij}}(\overline{z}_i, \overline{z}_j)]}{exp[\hat{a}_{\overline{z}_{-ij}}(\overline{z}_i, x_j)]} = \frac{p(x_i x_j \overline{z}_{-ij})p(\overline{z}_i \overline{z}_j \overline{z}_{-ij})}{p(x_i \overline{z}_j \overline{z}_{-ij})p(\overline{z}_i x_j \overline{z}_{-ij})}.$$

Since the left hand side expression, $\frac{exp[a_g(x_i,x_j)]}{exp[a_g(x_i,\overline{z}_j)]} \times \frac{exp[a_g(\overline{z}_i,\overline{z}_j)]}{exp[a_g(\overline{z}_i,x_j)]}$, does not depend on $\overline{z}_{-ij}$, it follows that the right hand side expression will not change if we replace $\overline{z}$ with $\overline{z}_i \overline{z}_j \overline{w}_{-ij}$ for any $\overline{w}_{-ij} \in \Omega_{-ij}$. Conclude that

$$\frac{p(x_i x_j \overline{z}_{-ij})p(\overline{z}_i \overline{z}_j \overline{z}_{-ij})}{p(x_i \overline{z}_j \overline{z}_{-ij})p(\overline{z}_i x_j \overline{z}_{-ij})} = \frac{p(x_i x_j \overline{z}_{-ij})p(\overline{z}_i \overline{z}_j \overline{w}_{-ij})}{p(x_i \overline{z}_j \overline{z}_{-ij})p(\overline{z}_i x_j \overline{w}_{-ij})},$$

which establishes DRAS. ■

Lemma 13 *If a belief p on Ω with full support satisfies DRAS then it is an Intuitive Belief.*

Proof. If $N=2$ then DRAS is satisfied vacuously. To obtain a representation in this case, simply define $exp[a(x_ix_j)]=p(x_ix_j)$. Let $Z=1$ and we obtain an Intuitive Belief representation.

Henceforth assume $N>2$. Fix some $z\in\Omega$ throughout.

Step 1: Show that, for $Z=\frac{1}{p(z)}$, and any x_i,x_j,x_k,

$$p(x_ix_jx_kz_{-ijk})=\frac{1}{Z}\prod_{l,m\in\{i,j,k\}:i<j}\frac{p(x_lx_mz_{-lm})}{p(x_lz_{-l})p(x_mz_{-m})}\times\prod_{l\in\{i,j,k\}}p(x_lz_{-l})$$

Take any x_i,x_j,x_k,z and consider $w_{-ij}=x_kz_{-ijk}$. By DRAS,

$$\frac{p(x_ix_jz_{-ij})p(z_iz_jz_{-ij})}{p(x_iz_jz_{-ij})p(z_ix_jz_{-ij})}=\frac{p(x_ix_j,x_kz_{-ijk})p(z_iz_j,x_kz_{-ijk})}{p(x_iz_j,x_kz_{-ijk})p(z_ix_j,x_kz_{-ijk})},$$

Rearranging this expression yields:

$p(x_ix_jx_kz_{-ijk})=p(z_iz_jz_{-ij})\frac{p(x_ix_jz_{-ij})p(x_iz_jx_kz_{-ijk})p(z_ix_jx_kz_{-ijk})}{p(x_iz_jz_{-ij})p(z_ix_jz_{-ij})p(z_iz_jx_kz_{-ijk})}$

$=p(z)\frac{p(x_ix_jz_{-ij})p(x_ix_kz_{-ik})p(x_jx_kz_{-jk})}{p(x_iz_{-i})p(x_jz_{-j})p(x_kz_{-k})}$.

$=p(z)\frac{p(x_ix_jz_{-ij})}{p(x_iz_{-i})p(x_jz_{-j})}\frac{p(x_ix_kz_{-ik})}{p(x_iz_{-i})p(x_kz_{-k})}\frac{p(x_jx_kz_{-jk})}{p(x_jz_{-j})p(x_kz_{-k})}p(x_iz_{-i})p(x_jz_{-j})p(x_kz_{-k})$, which yields the desired expression.

Step 2: Show that for any $x\in\Omega$,

$$p(x)=\frac{1}{Z}\left[\prod_{i,j\in\{1,..,N\},i<j}\frac{p(x_ix_jz_{-ij})}{p(x_iz_{-i})p(x_jz_{-j})}\times\prod_{i\in\{1,..,N\}}p(x_iz_{-i})\right].$$

Assume the induction hypothesis that for any n elements of Γ, which we abuse notation for and label as $1,..,n$,[13] we have the functional form

$$p(x_1..x_nz_{-1,...,n})=\frac{1}{Z}\prod_{i,j\in\{1,..,n\}:i<j}\frac{p(x_ix_jz_{-ij})}{p(x_iz_{-i})p(x_jz_{-j})}\times\prod_{i\in\{1,..,n\}}p(x_iz_{-i}),$$

where $Z=\frac{1}{p(z)}$. To prove the induction step, take any $1,..,n,n+1\in\Gamma$, any $x_{n+1}\in\Omega_{n+1}$. Define

$A=p(x_1...x_{n-1}z_nx_{n+1}z_{-1,....,n+1})$, $B=p(x_1...x_{n-2}z_{n-1}x_nx_{n+1}z_{-1,...,n+1})$, $C=p(x_1...x_{n-2}z_{n-1}z_nx_{n+1}z_{-1..n+1})$.

Adopt the simplifying notation $y_l^{l+m}:=(y_l...,y_{l+m})$ for any state y. Letting $w_{-n-1,.n}:=x_1^{n-2}x_{n+1}z_{n+2}^N$, DRAS implies

$\frac{p(x_{n-1}x_n,z_{-n-1,n})p(z_{n-1}z_n,z_{-n-1,n})}{p(x_{n-1}z_n,z_{-n-1,n})p(z_{n-1}x_n,z_{-n-1,n})}$

$=\frac{p(x_{n-1}x_n,x_1^{n-2}x_{n+1}z_{n+2}^N)p(z_{n-1}z_n,x_1^{n-2}x_{n+1}z_{n+2}^N)}{p(x_{n-1}z_n,x_1^{n-2}x_{n+1}z_{n+2}^N)p(z_{n-1}x_n,x_1^{n-2}x_{n+1}z_{n+2}^N)}$

$=\frac{p(x_1...x_{n+1}z_{-1..n+1})p(x_1...x_{n-2}z_{n-1}z_nx_{n+1}z_{-1..n+1})}{p(x_1...x_{n-1}z_nx_{n+1}z_{-1..n+1})p(x_1...x_{n-2}z_{n-1}x_nx_{n+1}z_{-1..n+1})}$

$=\frac{p(x_1...x_{n+1}z_{-1..n+1})C}{AB}$, which yields

$$p(x_1...x_{n+1}z_{-1..n+1})=\left[\frac{AB}{C}\right]\left[\frac{p(x_{n-1}x_nz_{-n-1,n})p(z)}{p(x_{n-1}z_{-n-1})p(x_nz_{-n})}\right]. \tag{9}$$

By the induction step, the terms A, B, C are given by

$$A=\frac{1}{Z}\times\prod_{i,j\in\{1,..,n-1,n+1\}:i<j}\frac{p(x_ix_jz_{-ij})}{p(x_iz_{-i})p(x_jz_{-j})}\times\prod_{i\in\{1,..,n-1,n+1\}}p(x_iz_{-i}),$$

[13] This is an abuse of notation since by deifnition, $\Gamma:=\{1,...,N\}$, whereas we will use $1,...,n$ to denote generic elements in Γ.

$$B = \frac{1}{Z} \times \prod_{i,j \in \{1,..,n-2,n,n+1\}: i<j} \frac{p(x_i x_j z_{-ij})}{p(x_i z_{-i}) p(x_j z_{-j})} \times \prod_{i \in \{1,..,n-2,n,n+1\}} p(x_i z_{-i}),$$

$$C = \frac{1}{Z} \times \prod_{i,j \in \{1,..,n-2,n+1\}: i<j} \frac{p(x_i x_j z_{-ij})}{p(x_i z_{-i}) p(x_j z_{-j})} \times \prod_{i \in \{1,..,n-2,n+1\}} p(x_i z_{-i}).$$

We see that

$$\frac{B}{C} = \left[\prod_{j \in \{1,..,n-2,n+1\}} \frac{p(x_n x_j z_{-nj})}{p(x_n z_{-n}) p(x_j z_{-j})} \right] \times p(x_n z_{-n})$$

$$= \frac{1}{\frac{p(x_{n-1} x_n z_{-n-1,n})}{p(x_{n-1} z_{-n-1}) p(x_n z_{-n})}} \left[\prod_{i \in \{1,..n-1,n+1\}} \frac{p(x_i x_n z_{-in})}{p(x_i z_{-i}) p(x_n z_{-n})} \right] \times p(x_n z_{-n})$$

and therefore

$$\frac{AB}{C} = \frac{1}{Z} \frac{p(z)}{\frac{p(x_{n-1} x_n z_{-n-1,n})}{p(x_{n-1} z_{-n-1}) p(x_n z_{-n})}} \times \prod_{i,j \in \{1,..,n+1\}: i<j} \frac{p(x_i x_j z_{-ij})}{p(x_i z_{-i}) p(x_j z_{-j})} \times \prod_{i \in \{1,..,n+1\}} p(x_i z_{-i}).$$

Inserting this into the equality (9) and using $Z = \frac{1}{p(z)}$ yields:

$$p(x_1 ... x_{n+1} z_{-1..n+1}) = \frac{1}{Z} \left[\prod_{i,j \in \{1,..,n+1\}: i<j} \frac{p(x_i x_j z_{-ij})}{p(x_i z_{-i}) p(x_j z_{-j})} \times \prod_{i \in \{1,..,n+1\}} p(x_i z_{-i}) \right],$$

completing the induction step. Conclude that $p(x|\Omega)$ can be written in the desired way.

Step 3: Conclude the proof of sufficiency.

The expression for p in Step 2 is an Intuitive Belief with $a(x_i x_j) = \frac{p(x_i x_j z_{-ij})}{p(x_i z_{-i}) p(x_j z_{-j})}$ and $b(x_i) = p(x_i z_{-i})$, where $z \in \Omega$ is fixed, as desired. ■

F Appendix: Proof of Propositions 1 and 2

While Propositions 1 and 2 derive meaning from the description of the extended environment, the formal results do not hinge on it. Below we prove several lemmas in a general setting and then specialize to the extended environment to immediately obtain the desired results.

F.1 General Training Results

Denote by $\Delta_{IB^+}(\Omega)$ the set of Intuitive Beliefs with full support. We first define a class A of normalized representations. By Lemmas 7 and 8, for any given $\overline{z}$, any belief $p \in \Delta_{IB^+}(\Omega)$ can be represented by a $\overline{z}$-normalized network $(a, 0)$ where a is defined for all $x_i, x_j \in \cup_{k \in \Gamma} \Omega_k$ by

$$exp[a(x_i x_j)] = p(\overline{z})^{\frac{N-3}{N-1}} \frac{p(x_i x_j \overline{z}_{-ij})}{[p(x_i \overline{z}_{-i}) p(x_j \overline{z}_{-j})]^{\frac{N-2}{N-1}}}.$$

Lemma 8 shows that a network $(a, 0)$ is $\overline{z}$-normalized if and only if $a(\overline{z}_i \overline{z}_j) = 0$ and $a(x_i \overline{z}_j) = a(x_i \overline{z}_k)$ for all distinct $i, j, k \in \Gamma$ and any $x \in \Omega$.

Fix any $\phi \neq D \subset \Omega$ and any $\overline{z} \in D$ and identify the set of such normalized representations with

$$A = \{a \in \mathbb{R}^{\cup_{i<j}(\Omega_i \times \Omega_j)} : a(\overline{z}_i \overline{z}_j) = 0 \text{ and } a(x_i \overline{z}_j) = a(x_i \overline{z}_k) \text{ for all distinct } i, j, k \in \Gamma \text{ and any } x \in \Omega\}.$$

This is clearly a convex and unbounded set. Moreover, for any $a_1, a_2 \in A$, if $a_1 \neq a_2$ then by the above expression of associations in terms of beliefs, the respective Intuitive Beliefs corresponding to a_1, a_2 are distinct.

Consider the subset of $\cup_{i<j}(\Omega_i \times \Omega_j)$ defined by

$$E_D = \{(x_i, x_j) : i, j \in \Gamma \text{ s.t. } i \neq j \text{ and } x \in D\},$$

consisting of all pairs of elementary states that can be constructed with the states in D. Since $\overline{z} \in D$, it follows that all pairs $(\overline{z}_i \overline{z}_j)$ are in E_D. Define A_D as the projection of A to $\mathbb{R}^{E_D}$. Any $a \in A_D$ contains all terms $a(\overline{z}_i \overline{z}_j) = 0$. It may not contain all (or any) terms of the type $a(x_i \overline{z}_j)$, but it remains consistent with the condition "$a(x_i \overline{z}_j) = a(x_i \overline{z}_k)$" whenever $(x_i \overline{z}_j), (x_i \overline{z}_k) \in E_D$. As the projection of a convex set, A_D is a convex set, though in contrast to A it may be bounded (for instance, it is a singleton if $D = \{\overline{z}\}$).

Say that $p(\cdot|D)$ on D is an Intuitive Belief if there exists $a \in A_D$ such that

$$p(x|D) = \frac{1}{Z_a} exp[\sum_{i<j} a(x_i x_j)], \qquad x \in D,$$

where $Z_a := \sum_y exp[\sum_{i<j} a(y_i y_j)]$. (In Lemma 16 we will confirm that definition is equivalent to the definition given in Section 3.2). Since a is real-valued, $p(\cdot|D)$ has full support. Denote by $\Delta_{IB^+}(D)$ the set of Intuitive Beliefs on D with full support.

Given $\phi \neq D \subset \Omega$ and any full support probability measure $q(\cdot|D)$ over D (which we will often write as q to ease notation) we will study the following minimization problem

$$min\ KL(q||p) \text{ s.t. } p \in \Delta_{IB}(D), \tag{10}$$

where KL-divergence is defined for probability measures over D.

For any $a \in A_D$, define

$$KL(q||a) := KL(q||p_a),$$

where $p_a \in \Delta_{IB}(D)$ is the Intuitive Belief on D defined by a. The training problem (10) can therefore be viewed as one where $KL(q||a)$ is minimized over $a \in A_D$.

KL divergence is well-known to be strictly convex, and indeed $p \mapsto KL(q||p)$ is strictly convex in p.[14] We determine the shape of $a \mapsto KL(q||a)$ in the next lemma. Define the set

$$D_{\overline{z}} := \{(x_i x_j \overline{z}_{-ij}) : x_i, x_j \in \cup_{k \in \Gamma} \Omega_k\},$$

that consists of states that are "simple" wrt the $\overline{z} \in D$ we fixed in the preceding.

Lemma 14 *The function $a \mapsto KL(q||a)$ is convex and satisfies*

$$KL(q||a) = lnZ_a + \sum_{x \in D} q(x) \left[lnq(x) - \sum_{i<j} a(x_i x_j) \right]. \tag{11}$$

It is strictly convex if $D_{\overline{z}} \subset D$, that is, if the set of "simple" states are contained in D.

[14] Since logs are a strictly concave function, it must be that for any $p \neq r$ and $\alpha \in (0,1)$,

$$KL(q, \alpha p + (1-\alpha)r) := \sum_x q(x)[lnq(x) - ln[\alpha p(x) + (1-\alpha)r(x)]]$$

$$< \sum_x q(x)[ln\, q(x) - \alpha ln\, p(x) - (1-\alpha) ln\, r(x)] = \alpha KL(q,p) + (1-\alpha)KL(q,r).$$

Proof. We prove the result in steps.

Step 1: Show that $KL(q||a)$ satisfies (11) for any $a \in A_D$.

Compute that for any $a \in A_D$,

$KL(q||a) = KL(q||p_a) = \sum_{x\in D} q(x)[lnq(x) - lnp_a(x)]$

$= \sum_{x\in D} q(x)[lnq(x) - ln\frac{1}{Z_a}exp[\sum_{i<j} a(x_i x_j)]]$

$= \left[\sum_{x\in D} q(x)[lnq(x) - \sum_{i<j} a(x_i x_j)]\right] + [\sum_{x\in D} q(x) \times lnZ_a]$

$= \left[\sum_{x\in D} q(x)[lnq(x) - \sum_{i<j} a(x_i x_j)]\right] + [lnZ_a] \times [\sum_{x\in D} q(x)]$, which yields the desired expression since $\sum_{x\in D} q(x) = 1$.

Step 2: Show that for any $a_1, a_2 \in A_S$ with respective Intuitive Beliefs on D given by $p_1, p_2 \in \Delta_{IB^+}(D)$, and any $\theta \in [0,1]$, the network $a = \theta a_1 + (1-\theta)a_2$ represents Intuitive Beliefs given by the following normalized geometric mixture: for all $x \in D$,

$$p(x|D) = \frac{p_1^\theta(x)p_2^{1-\theta}(x)}{\sum_y p_1(y)^\theta p_2(y)^{1-\theta}}.$$

Denote the normalizing constants for $p_1, p_2 \in \Delta_{IB^+}(S)$ by Z_1, Z_2 respectively. The normalizing constant Z_a for $a = \theta a_1 + (1-\theta)a_2$ must satisfy

$Z_a = \sum_{y\in D} exp[\sum_{i<j}[a(y_i y_j)]$

$= \sum_{y\in D} exp[\sum_{i<j}[\theta a_1(y_i y_j) + (1-\theta)a_2(y_i y_j)]$

$= \sum_{y\in D} \left[exp[\sum_{i<j} a_1(y_i y_j)]\right]^\theta \left[exp[\sum_{i<j}[a_2(y_i y_j)]\right]^{1-\theta}$, and so

$$Z_a = Z_1^\theta Z_2^{1-\theta} \times \sum_{y\in D} p_1(y)^\theta p_2(y)^{1-\theta}.$$

Therefore $(a, 0)$ represents an Intuitive Belief $p(\mathbf{.}|D)$ given by: for all $x \in D$,

$p(x|D) = \frac{1}{Z_a} exp[\sum_{i<j}[a(x_i x_j)]$

$= \frac{1}{Z_1^\theta Z_2^{1-\theta}\sum_y p_1(y)^\theta p_2(y)^{1-\theta}} exp[\sum_{i<j}(\theta a_1(x_i x_j) + (1-\theta)a_2(x_i x_j))]$

$= \frac{1}{\sum_y p_1(y)^\theta p_2(y)^{1-\theta}} \frac{exp[\theta\sum_{i<j} a_1(x_i x_j)]}{Z_1^\theta} \times \frac{exp[(1-\theta)\sum_{i<j} a_2(x_i x_j)]}{Z_2^{1-\theta}}$

$= \frac{p_1^\theta(x)p_2^{1-\theta}(x)}{\sum_y p_1(y)^\theta p_2(y)^{1-\theta}}$, therefore establishing the step.

Step 3: Define $M_\theta := \sum_y p_1(y)^\theta p_2(y)^{1-\theta}$. Show that $M_\theta \leq 1$ and, moreover, when $\theta \in (0,1)$,

$$M_\theta < 1 \iff p_1 \neq p_2.$$

Consider networks $a_1, a_2 \in A_D$ with respective normalizing constants Z_1, Z_2 and Intuitive Beliefs $p_1, p_2 \in \Delta_{IB^+}(D)$. Trivially, $M_\theta = 1$ if $\theta \in \{0,1\}$ or if $\theta \in (0,1)$ and $p_1 = p_2$. If $p_1 \neq p_2$ then we show that $M_\theta < 1$ by Jensen's inequality (for strictly concave functions):

$M_\theta = \sum_y p_1(y)^\theta p_2(y)^{1-\theta} = \sum_y [\frac{p_1(y)}{p_2(y)}]^\theta p_2(y)$

$< [\sum_y \frac{p_1(y)}{p_2(y)} p_2(y)]^\theta = [\sum_y p_1(y)]^\theta = 1$. This completes the step.

Step 4: Prove the result.

Consider networks $a_1, a_2 \in A_D$ with respective Intuitive Beliefs $p_1, p_2 \in \Delta_{IB^+}(D)$. Given Step 2, we see that

$KL(q||\theta a_1 + (1-\theta)a_2) = KL(q||p_1^\theta p_2^{1-\theta})$

$= \sum_{x\in D} q(x)[lnq(x) - ln[\frac{1}{M_\theta} p_1^\theta(x)p_2^{1-\theta}(x)]]$

$= \sum_{x\in D} q(x)[lnq(x) - \theta lnp_1(x) - (1-\theta)lnp_2(x) + lnM_\theta]$

$= \theta\sum_{x\in D} q(x)[lnq(x) - lnp_1(x)] + (1-\theta)\sum_{x\in D} q(x)[lnq(x) - lnp_2(x)] + lnM_\theta]$

$= \theta KL(q||p_1) + (1-\theta)KL(q||p_2) + lnM_\theta$

$= \theta KL(q||a_1) + (1-\theta)KL(q||a_2) + lnM_\theta$, that is,

$$KL(q||\theta a_1 + (1-\theta)a_2) = \theta KL(q||a_1) + (1-\theta)KL(q||a_2) + lnM_\theta.$$

Since $M_\theta \leq 1$ (Step 3) we see that $a \mapsto KL(q||a)$ is convex, as desired.

Next, suppose $D_{\overline{z}} \subset D$. Consider networks $a_1, a_2 \in A_D$ s.t. $a_1 \neq a_2$ and denote the respective Intuitive Beliefs by p_1, p_2. Take $\theta \in (0,1)$. We need to show that $KL(q||\theta a_1 + (1-\theta)a_2) < \theta KL(q||a_1) + (1-\theta)KL(q||a_2)$. The preceding argument yields this conclusion if $M_\theta < 1$. By Step 3, $M_\theta < 1$ holds if $p_1 \neq p_2$. Therefore we must establish that $p_1 \neq p_2$. More specifically, we show that $a_1 \neq a_2$ implies $p_1 \neq p_2$.

Under the condition $D_{\overline{z}} \subset D$, the set E_D contains all possible pairs of elementary events, and therefore $A_D = A$. If we suppose by way of contradiction that $p_1 = p_2$ then it must be, by Lemma 8, that for all $x_i, x_j \in \cup_{k\in\Gamma}\Omega_k$,

$$exp[a_1(x_ix_j)] = p_1(\overline{z})^{\frac{N-3}{N-1}} \frac{p_1(x_ix_j\overline{z}_{-ij})}{[p_1(x_i\overline{z}_{-i})p_1(x_j\overline{z}_{-j})]^{\frac{N-2}{N-1}}}$$

$$= p_2(\overline{z})^{\frac{N-3}{N-1}} \frac{p_2(x_ix_j\overline{z}_{-ij})}{[p_2(x_i\overline{z}_{-i})p_2(x_j\overline{z}_{-j})]^{\frac{N-2}{N-1}}} = exp[a_2(x_ix_j)].$$

which contradicts the hypothesis that $a_1 \neq a_2$. Therefore $p_1 \neq p_2$, as desired. ■

Lemma 15 *Intuitive Belief $p(.|D)$ is a solution to (10) if and only if $p(x_ix_j|D) = q(x_ix_j|D)$ for all distinct $i, j \in \Gamma$ and $(x_i, x_j) \in E_S$.*

Proof. Consider the training problem (10). Since q over D has full support, and since $KL(q||p) = \infty$ if $p(x) = 0$ for some $x \in D$, the training problem cannot be solved by Intuitive Beliefs outside $\Delta_{IB^+}(D)$. Consequently, if a solution to the training problem exists, it must exist in $\Delta_{IB^+}(D)$. The training problem is therefore equivalent to seeking to a solution to

$$KL(q||a) \text{ over } a \in A_D.$$

Since the optimization problem is unconstrained, the Lagrangian is

$$\mathcal{L}(a) = KL(q||a) = lnZ_a + \sum_{x\in D} q(x|D)\left[lnq(x|D) - \sum_{i<j} a(x_ix_j)\right],$$

where the expression for $KL(q||a)$ was established in Lemma 14. As noted earlier, A_D is a convex set, and the function $a \mapsto KL(q||a)$ is convex (by Lemma 14). Consequently, the first order conditions are both necessary and sufficient for optimality. It follows that a minimizer exists if and only if it satisfies the first order conditions, which we derive below.

For any $x_i, x_j \in \cup_{k\in\Gamma}\Omega_k$ denote by $D_{x_ix_j}$ all the states $z \in D$ that take on values $z_i = x_i$ and $z_j = x_j$ on dimension i and j respectively. The first order condition wrt any $a(x_ix_j)$, where $i, j \in \Gamma$ are distinct and $x_i, x_j \in \cup_{k\in\Gamma}\Omega_k$, yields:

$\frac{\partial\mathcal{L}}{\partial a(x_ix_j)} = 0$

$\iff \left[\sum_{z\in D_{x_ix_j}} q(z|D)[0-1]\right] + \frac{\sum_{w\in D_{x_ix_j}} exp[\sum_{k<l} a(w_kw_l)]}{\sum_{y\in D} exp[\sum_{k<l} a(y_ky_l)]} = 0$

$\iff \frac{\sum_{w\in D_{x_ix_j}} exp[\sum_{k<l} a(w_kw_l)]}{\sum_{y\in D} exp[\sum_{k<l} a(y_ky_l)]} = \sum_{z\in D_{x_ix_j}} q(z|D)$

$\iff \sum_{w\in D_{x_ix_j}} p(z|D) = \sum_{z\in D_{x_ix_j}} q(z|D)$

$\iff p(x_ix_j|D) = q(x_ix_j|D)$, which yields the conclusion that the 2-dimensional marginals must match, if a solution exists. This completes the proof ■

Lemma 16 *Suppose* $D_{\overline{z}} \subset D$*. If* $p(.|D)$ *is a solution to (10) then it is the unique solution. Moreover, it extends uniquely to an full-support Intuitive Belief* p *over* Ω *that satisfies* $p(.|D) = \frac{p(.)}{p(D)}$ *on* D*.*

Proof. Under the condition $D_{\overline{z}} \subset D$, the set E_D contains all possible pairs of elementary events, and therefore

$$A_D = A.$$

Since (i) the problem (10) can be viewed as minimizing $KL(q||a)$ over $a \in A_D = A$, (ii) A is convex, and (ii), the objective function $a \mapsto KL(q||a)$ is strictly convex (by Lemma 14), it follows that if a solution to (10) exists, then it must be unique. Furthermore if $p(.|D)$ is a solution to (10), then by virtue of it being an Intuitive Belief over D, it is represented by some network $a \in A_D = A$. But a represents some Intuitive Belief p over Ω. Observe that $p(x|D) = \frac{exp[\sum_{i<j} a(x_i x_j)]}{\sum_{y \in D} exp[\sum_{i<j} a(y_i y_j)]} = \frac{p(x)}{p(D)}$ for all $x \in D$. Thus, p over Ω is an extension in the sense defined in Section 3.2. Given $D_{\overline{z}} \subset D$, Theorem 2 implies that this extension is unique. ■

F.2 Proofs of Propositions 1 and 2

These propositions respectively correspond to Lemmas 15 and 16 specialized (without any consequence except the "*" notation) to the extended environment.

G Proof of Theorem 4

Proof. Consider the minimization problem (4).

Step 1: Show that the solution is given by $p(.|D^*) = q(.|D^*)$.

As is evident from the reduced form in Lemma 9, Intuitive Beliefs over Ω^* are unrestricted on the subdomain D_{z^*}. Consequently, given $D^* = D_{z^*}$, we can set $p(x_i x_j z^*_{-ij}|D^*) = q(x_i x_j z^*_{-ij}|D^*)$ for all $x_i x_j z^*_{-ij} \in D^* = D_{z^*}$ and solve the minimization problem (4) with $KL(q||p) = 0$.

Step 2: Show that $p(.|D^*)$ uniquely extends to some $p(.|\Omega^*)$ that satisfies, for some $k > 0$,

$$p(x|\Omega^*) = kq(z^*|D^*) \times \left[\prod_{i<j} \frac{q(x_i x_j z^*_{-ij}|D^*)q(z^*|D^*)}{q(x_i z^*_{-i}|D^*)q(x_j z^*_{-j}|D^*)}\right] \times \prod_{i \in \Gamma} \frac{q(x_i z^*_{-i}|D^*)}{q(z^*|D^*)}, \quad x \in \Omega^*.$$

By Theorem 2, there exists a unique extension to some $p(.|\Omega^*)$. However, by Lemma 9, $p(.|\Omega^*)$ can be described by the representation

$$p(x|\Omega^*) = p(z^*|\Omega^*) \times \left[\prod_{i<j} \frac{p(x_i x_j z^*_{-ij}|\Omega^*)p(z^*|\Omega^*)}{p(x_i z^*_{-i}|\Omega^*)p(x_j z^*_{-j}|\Omega^*)}\right] \times \prod_{i \in \Gamma} \frac{p(x_i z^*_{-i}|\Omega^*)}{p(z^*|\Omega^*)}, \quad x \in \Omega^*.$$

Since $D^* = D_{z^*}$, the set D^* contains all $x_i x_j z^*_{-ij}$ such that $x_i, x_j \in \cup_{k \in \Gamma} \Omega^*_k$. By definition of an extension, $\frac{p(x_i x_j z^*_{-ij}|D^*)}{p(y_i y_j z^*_{-ij}|D^*)} = \frac{p(x_i x_j z^*_{-ij}|\Omega^*)}{p(y_i y_j z^*_{-ij}|\Omega^*)}$ for all $x_i x_j z^*_{-ij}, y_i y_j z^*_{-ij} \in D^*$. Fix $y_i y_j z^*_{-ij}$ and let $k = \frac{p(y_i y_j z^*_{-ij}|\Omega^*)}{p(y_i y_j z^*_{-ij}|D^*)} > 0$. Then by the preceding (including Step 1),

$$p(x_i x_j z^*_{-ij}|\Omega^*) = k \times p(x_i x_j z^*_{-ij}|D^*) = k \times q(x_i x_j z^*_{-ij}|D^*)$$

for all $x_i x_j z^*_{-ij} \in D^*$, and we can rewrite the above representation for $p(.|\Omega^*)$ as desired.

Step 3: Show that there is some $W > 0$ s.t. for any $x \in \Omega^*$,

$$p(x|\Omega^*) = \frac{1}{W} \left[\prod_{i<j} \frac{q^m(x_i x_j)}{q^m(x_i)q^m(x_j)}\right] \times \left[\prod_{i \in I} q^m(x_i)\right].$$

Given Step 2, and using the expression for the objective distribution over signals, $q(x_I z^*_{-I}|D^*) = \sigma(I) \times q^m(x_I)$ (given in (3)), we see that for all $x \in \Omega^*$,

$$p(x|\Omega^*) = kq(z^*|D^*) \times \left[\prod_{i<j} \frac{q(x_i x_j z^*_{-ij}|D^*) q(z^*|D^*)}{q(x_i z^*_{-i}|D^*) q(x_j z^*_{-j}|D^*)}\right] \times \prod_{i \in \Gamma} \frac{q(x_i z^*_{-i}|D^*)}{q(z^*|D^*)}$$

$$= kq(z^*|D^*) \times \left[\prod_{i<j} \frac{\sigma(i,j) q^m(x_i x_j) \times \sigma(\phi)}{\sigma(i) q^m(x_i) \times \sigma(j) q^m(x_j)}\right] \times \left[\prod_{i \in I} \frac{\sigma(i) q^m(x_i|\Omega)}{\sigma(\phi)}\right].$$

The desired claim is established by appropriately defining $W > 0$ after collecting the σ terms (which are all strictly positive by definition of D^*).

Step 4: Complete the proof.

The belief p over Ω induced by $p(\cdot|\Omega^*)$ is given by: for all $x \in \Omega$,

$$p(x) := \frac{p(x|\Omega^*)}{\Sigma_{y \in \Omega} p(y|\Omega^*)} = \frac{1}{Z} \left[\prod_{i<j} \frac{q^m(x_i x_j)}{q^m(x_i) q^m(x_j)}\right] \times \left[\prod_{i \in I} q^m(x_i)\right],$$

for some constant $Z > 0$, as desired. This also justifies the associative network defined in the statement of the Theorem. ■

H Appendix: Proof of Proposition 4

Proof. The objective distribution is given by

$$q(h_i h_j h_k) = q(l_i l_j l_k) = \frac{1}{2}\alpha + \frac{1}{8}(1-\alpha) = \frac{1}{8}(1+3\alpha)$$

$$q(h_i h_j l_k) = q(h_i l_j l_k) = \frac{1}{8}(1-\alpha),$$

with marginals

$$q(h_i) = q(l_i) = \frac{1}{2}.$$

Using the expression for p in terms of q given by Theorem 4, compute that

$$p(h_1 h_2 h_3) = p(l_i l_j l_k) = \frac{1}{Z}\left[\frac{[\alpha + \frac{1}{2}(1-\alpha)]^3}{[\alpha + (1-\alpha)]^3}\right] = \frac{1}{Z}\frac{1}{8}(1+\alpha)^3$$

$$p(h_i h_j l_k) = p(h_i l_j l_k) = \frac{1}{Z}\left[\frac{[\frac{1}{2}\alpha + \frac{1}{4}(1-\alpha)][\frac{1}{4}(1-\alpha)]^2}{[\frac{1}{2}\alpha + \frac{1}{2}(1-\alpha)]^3}\right] = \frac{1}{Z}\frac{1}{8}(1+\alpha)(1-\alpha)^2$$

and $Z = Z\sum_{(x_i x_j x_k)} p(x_i x_j x_k) = \frac{2}{8}(1+\alpha)^3 + \frac{6}{8}(1+\alpha)(1-\alpha)^2 = 1 + \alpha^3$. Therefore,

$$p(h_1 h_2 h_3) = p(l_i l_j l_k) = \frac{1}{8}\frac{(1+\alpha)^3}{1+\alpha^3} \text{ and } p(h_i h_j l_k) = p(h_i l_j l_k) = \frac{1}{8}\frac{(1+\alpha)(1-\alpha)^2}{1+\alpha^3}.$$

The result follows from these expressions. ■

I Appendix: Proof of Proposition 5

Proof. The objective distribution is given by

$$q(h_ih_jl_k) = q(l_il_jl_k) = \frac{1}{4}\alpha + \frac{1}{8}(1-\alpha)$$

$$q(h_ih_jh_k) = q(h_il_jl_k) = 0\alpha + \frac{1}{8}(1-\alpha) = \frac{1}{8}(1-\alpha)$$

with marginals for all x_ix_j

$$q^m(x_ix_j) = \frac{1}{8}(1-\alpha) + \frac{1}{4}\alpha + \frac{1}{8}(1-\alpha) = \frac{1}{4} \text{ and } q^m(x_i) = \frac{1}{2}.$$

Then, $\frac{q^m(x_ix_j)}{q^m(x_i)q^m(x_j)} = 1$ for all x_ix_j and so $p(x) = \frac{1}{Z}q^m(x_i)q^m(x_j)q^m(x_k) = \frac{1}{8}$. ■

J Appendix: Proof of Proposition 6

Proof. We prove the result for the more general case where the base-line distribution is $q_b(x_1x_2x_3) = \xi(x_1)\xi(x_2)\xi(x_3)$ where $\xi(x_i)$ is a full support probability on $\Omega_i = \{h, l\}$. Recall that

$$q(x_1x_2x_3) = \alpha r(x_1x_2x_3) + (1-\alpha)\xi(x_1)\xi(x_2)\xi(x_3), \qquad (x_ix_jx_k) \in \Omega,$$

and write $r(h_1h_2h_3) = \theta$. Consider the good market, where $\theta > 0.5$. Compute that

$$q(h_1h_2h_3) = \alpha\theta + (1-\alpha)\xi(h_1)\xi(h_2)\xi(h_3),\ q(l_1l_2l_3) = \alpha(1-\theta) + (1-\alpha)\xi(l_1)\xi(l_2)\xi(l_3),$$

$$q(x_1x_2x_3) = (1-\alpha)\xi(x_1)\xi(x_2)\xi(x_3), \qquad (x_1x_2x_3) \neq (l_1l_2l_3), (h_1h_2h_3).$$

Then for any $x_i \in \{s, f\}$, the two dimensional marginals are

$$q^m(h_ih_j) = \alpha\theta + (1-\alpha)\xi(h_i)\xi(h_j), \quad q^m(l_il_j) = \alpha(1-\theta) + (1-\alpha)\xi(l_i)\xi(l_j)$$

$$\text{and } q^m(h_il_j) = (1-\alpha)\xi(h_i)\xi(l_j)$$

and the one-dimensional marginals are

$$q^m(h_i) = \alpha\theta + (1-\alpha)\xi(h_i) \text{ and } q^m(l_i) = \alpha(1-\theta) + (1-\alpha)\xi(l_i).$$

Given $Z = \sum_{(y_1y_2y_3)\in\Omega} \frac{q^m(y_1y_2)q^m(y_1y_3)q^m(y_2y_3)}{q^m(y_1)q^m(y_2)q^m(y_3)}$, consider the term for $(h_1h_2h_3)$ given by

$$\frac{q^m(h_ih_j)q^m(h_ih_k)q^m(h_jh_k)}{q^m(h_i)q^m(h_j)q^m(h_k)} = \frac{[\alpha\theta + (1-\alpha)\xi(h_i)\xi(h_j)][\alpha\theta + (1-\alpha)\xi(h_i)\xi(h_k)][\alpha\theta + (1-\alpha)\xi(h_j)\xi(h_k)]}{[\alpha\theta + (1-\alpha)\xi(h_i)][\alpha\theta + (1-\alpha)\xi(h_j)][\alpha\theta + (1-\alpha)\xi(h_k)]}$$

and observe that this term approaches 1 as $\alpha \to 1$. Similarly the term for $(l_1l_2l_3)$ also approaches 1. The other terms have a 0 limit: the terms for $(h_ih_jl_k)$ and $(h_il_jl_k)$ are respectively:

$$\frac{[\alpha\theta + (1-\alpha)\xi(h_i)\xi(h_j)][(1-\alpha)\xi(h_i)\xi(l_k)][(1-\alpha)\xi(h_j)\xi(l_k)]}{[\alpha\theta + (1-\alpha)\xi(h_i)][\alpha\theta + (1-\alpha)\xi(h_j)][\alpha(1-\theta) + (1-\alpha)\xi(l_k)]}$$

$$\text{and } \frac{[(1-\alpha)\xi(h_i)\xi(l_j)][(1-\alpha)\xi(h_i)\xi(l_k)][\alpha(1-\theta) + (1-\alpha)\xi(l_j)\xi(l_k)]}{[\alpha\theta + (1-\alpha)\xi(h_i)][\alpha\theta + (1-\alpha)\xi(h_j)][\alpha(1-\theta) + (1-\alpha)\xi(l_k)]}.$$

Therefore we find that

$$lim_{\alpha\to 1}Z = 2.$$

Now establish the result. Given that $p(x_1x_2x_3) = \frac{1}{Z}\left[\frac{q^m(x_1x_2)q^m(x_1x_3)q^m(x_2x_3)}{q^m(x_1)q^m(x_2)q^m(x_3)}\right]$ we see that the beliefs given by

$$p(h_1h_2h_3) = \frac{1}{Z}\left[\frac{[\alpha\theta + (1-\alpha)\xi(h_i)\xi(h_j)][\alpha\theta + (1-\alpha)\xi(h_i)\xi(h_k)][\alpha\theta + (1-\alpha)\xi(h_j)\xi(h_k)]}{[\alpha\theta + (1-\alpha)\xi(h_i)][\alpha\theta + (1-\alpha)\xi(h_j)][\alpha\theta + (1-\alpha)\xi(h_k)]}\right]$$

exhibit

$$lim_{\alpha\to1}p(h_1h_2h_3) = lim_{\alpha\to1}\frac{1}{Z} = \frac{1}{2} < r(h_1h_2h_3) = lim_{\alpha\to1}q(h_ih_jh_k),$$

establishing that the agent underestimates the true probability of the state $(h_1h_2h_3)$. To establish that the true probability of $(l_1l_2l_3)$ is underestimated, observe first that

$$lim_{\alpha\to1}p(h_ih_jl_k) = lim_{\alpha\to1}\frac{1}{Z}\left[\frac{[\alpha\theta + (1-\alpha)\xi(h_i)\xi(h_j)][(1-\alpha)\xi(h_i)\xi(l_k)][(1-\alpha)\xi(h_j)\xi(l_k)]}{[\alpha\theta + (1-\alpha)\xi(h_i)][\alpha\theta + (1-\alpha)\xi(h_j)][\alpha(1-\theta) + (1-\alpha)\xi(l_k)]}\right]$$
$$= 0 = r(h_ih_jl_k) = lim_{\alpha\to1}q(h_ih_jl_k),$$

$$lim_{\alpha\to1}p(h_il_jl_k) = lim_{\alpha\to1}\frac{1}{Z}\left[\frac{[(1-\alpha)\xi(h_i)\xi(l_j)][(1-\alpha)\xi(h_i)\xi(l_k)][\alpha(1-\theta) + (1-\alpha)\xi(l_j)\xi(l_k)]}{[\alpha\theta + (1-\alpha)\xi(h_i)][\alpha(1-\theta) + (1-\alpha)\xi(l_j)][\alpha(1-\theta) + (1-\alpha)\xi(l_k)]}\right]$$
$$= 0 = r(h_il_jl_k) = lim_{\alpha\to1}q(h_ih_jl_k).$$

It follows that $lim_{\alpha\to1}p(l_1l_2l_3) = lim_{\alpha\to1}[1 - p(l_1l_2h_3) - .... - p(h_1h_2h_3)] = 1 - lim_{\alpha\to1}p(h_1h_2h_3)$. We saw already that $lim_{\alpha\to1}p(h_1h_2h_3) < lim_{\alpha\to1}q(h_1h_2h_3) = r(h_1h_2h_3)$. Conclude that

$lim_{\alpha\to1}p(l_1l_2l_3)$
$= 1 - lim_{\alpha\to1}p(h_1h_2h_3)$
$> 1 - r(h_1h_2h_3)$
$= r(l_1l_2l_3) = lim_{\alpha\to1}q(h_ih_jl_k)$, as desired.

The result for the case of the bad market ($\theta < 0.5$) is obtained by switchng the labels h and l. ■

K Appendix: Proof of Proposition 7

Proof. Compute that $q(l_1l_2l_3) = (1-\alpha)\frac{1}{8}$ and

$$q(x_1x_2x_3) = \alpha\frac{1}{7} + (1-\alpha)\frac{1}{8}, \qquad (x_1x_2x_3) \neq (l_1l_2l_3)$$

Then for any $x_i \in \{s, f\}$, the two dimensional marginals are

$$q^m(h_ix_j) = \alpha\frac{2}{7} + (1-\alpha)\frac{1}{4} \text{ and } q^m(l_il_j) = \alpha\frac{1}{7} + (1-\alpha)\frac{1}{4}$$

and the one-dimensional marginals are

$$q^m(h_i) = \alpha\frac{4}{7} + (1-\alpha)\frac{1}{2},\ q^m(l_i) = \alpha\frac{3}{7} + (1-\alpha)\frac{1}{2}$$

Using the formula for trained Intuitive Beliefs we see that

$$p(h_1h_2h_3) = \frac{1}{Z}\frac{q^m(h_ih_j)q^m(h_ih_k)q^m(h_jh_k)}{q^m(h_i)q^m(h_j)q^m(h_k)} = \frac{1}{Z}\frac{[\alpha\frac{2}{7} + (1-\alpha)\frac{1}{4}]^3}{[\alpha\frac{4}{7} + (1-\alpha)\frac{1}{2}]^3}$$

$$p(h_ih_jl_k) = \frac{1}{Z}\frac{q^m(h_ih_j)q^m(h_il_k)q^m(h_jl_k)}{q^m(h_i)q^m(h_j)q^m(l_k)} = \frac{1}{Z}\frac{[\alpha\frac{2}{7} + (1-\alpha)\frac{1}{4}]^3}{[\alpha\frac{4}{7} + (1-\alpha)\frac{1}{2}]^2[\alpha\frac{3}{7} + (1-\alpha)\frac{1}{2}]}$$

$$p(h_il_jl_k) = \frac{1}{Z}\frac{q^m(h_il_j)q^m(h_il_k)q^m(l_jl_k)}{q^m(h_i)q^m(l_j)q^m(l_k)} = \frac{1}{Z}\frac{[\alpha\frac{2}{7}+(1-\alpha)\frac{1}{4}]^2[\alpha\frac{1}{7}+(1-\alpha)\frac{1}{4}]}{[\alpha\frac{4}{7}+(1-\alpha)\frac{1}{2}][\alpha\frac{3}{7}+(1-\alpha)\frac{1}{2}]^2}$$

$$p(l_1l_2l_3) = \frac{1}{Z}\frac{q^m(l_il_j)q^m(l_il_k)q^m(l_jl_k)}{q^m(l_i)q^m(l_j)q^m(l_k)} = \frac{1}{Z}\frac{[\alpha\frac{1}{7}+(1-\alpha)\frac{1}{4}]^3}{[\alpha\frac{3}{7}+(1-\alpha)\frac{1}{2}]^3}.$$

Compute that $lim_{\alpha\to1}Z = \frac{1}{8}+\frac{1}{2}+\frac{1}{3}+\frac{1}{27} \approx 0.995$. Therefore

$$lim_{\alpha\to1}p(l_1l_2l_3) \approx 0.037.$$

Moreover $lim_{\alpha\to1}q(l_1l_2l_3) = 0$ since $q(l_1l_2l_3) = (1-\alpha)\frac{1}{8}$. ■

References

[1] Ackley, D., G. Hinton and T. Sejnowski (1985): "A Learning Algorithm for Boltzmann Machines", *Cognitive Science* 9(1), pp. 147–169.

[2] Anderson, J. R. (1983): "A Spreading Activation Theory of Memory", *Journal of Verbal Learning and Verbal Behavior* 22(3), pp. 261-295.

[3] Barro R. (2006): "Rare Disasters and Asset Markets in the Twentieth Century", *Quarterly Journal of Economics* 121, pp. 823-866.

[4] Benjamin, D., A. Bodoh-Creed, and M. Rabin (2019): "Base-Rate Neglect: Foundations and Implications", working paper.

[5] Betch, T. (2007): "The Nature of Intuition and Its Neglect in Research on Judgment and Decision Making", in E. Plessner, C. Betsch and T. Betsch (Eds.), *Intuition in Judgement and Decision Making*, Psychology Press; 1st edition.

[6] Bodoh-Creed, A. (2020): "Mood, Memory and the Evaluation of Asset Prices", *Review of Finance* 24(1), pp. 227–262.

[7] Bordalo, P., N. Gennaioli, and A. Shleifer (2019): "Memory, Attention and Choice", working paper.

[8] Collins, A. and E. Loftus (1975). "A Spreading-Activation Theory of Semantic Processing", *Psychological Review* 82(6), pp. 407–428.

[9] Gennaioli, N. and A. Shleifer (2010): "What Comes to Mind", *Quarterly Journal of Economics* 125 (4), pp. 1399-1433.

[10] Gilboa, I. and D. Schmeidler (2001): *A Theory of Case-Based Decisions*, Cambridge University Press.

[11] Gilboa, I. and D. Schmeidler (2003): "Inductive Inference: An Axiomatic Approach", *Econometrica* 71(1), pp. 1-26.

[12] Hensman, A. and E. Sadler-Smith (2011): "Intuitive Decision-Making in Banking and Finance", *European Management Journal* 29, pp. 51–66.

[13] Hinton, G. and T. Sejnowski (1983): "Optimal Perceptual Inference", *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, Washington DC, pp. 448-453.

[14] Hopfield, J. J. (1982): "Neural Networks and Physical Systems with Emergent Collective Computational Abilities", *Proceedings of the National Academy of Sciences* 79 (8), pp. 2554–2558.

[15] Huang, L. and J. Pearce (2015): “Managing the Unknowable: The Effectiveness of Early-stage Investor Gut Feel in Entrepreneurial Investment Decisions”, *Administrative Science Quarterly* 60(4), pp. 634-670.

[16] Huang, L. (2018): “The Role of Investor Gut Feel in Managing Complexity and Extreme Risk”, *Academy of Management Journal* 61(5), pp. 1821–1847.

[17] Hume, D. (1748). *An Inquiry Concerning Human Understanding.* Oxford Clarendon Press, second edition 1966.

[18] Kahana, M. (2020): “Computational Models of Memory Search”, *Annual Review of Psychology* 71, pp. 107–38.

[19] Keynes, J. (1936). *The General Theory of Employment, Interest and Money.* London. Macmillan.

[20] Moore, D. and P. Healy (2008). “The Trouble with Overconfidence”, *Psychological Review* 115(2), pp. 502–517.

[21] Morewedge, C. and D. Kahneman (2010): “Associative Processes in Intuitive Judgment”, *Trends in Cognitive Science* 14(10), pp. 435–440.

[22] Mullainathan, S. (2002): “A Memory-Based Model of Bounded Rationality”, *Quarterly Journal of Economics*, 2002, vol. 117, issue 3, 735-774.

[23] Noor, J. (2020a): “Non-Probabilistic Intuitive Beliefs”, mimeo.

[24] Noor, J. (2020b): “Intuitive Learning”, mimeo.

[25] Rabin, M. (2002), “Inference by Believers in the Law of Small Numbers”, *Quarterly Journal of Economics* 117, pp. 775–816.

[26] Rabin, M. and D. Vayanos (2010): “The Gambler’s and Hot-Hand Fallacies: Theory and Applications”, *Review of Economic Studies* 77, pp. 730–778.

[27] Rietz T. (1988): “The Equity Premium: a Solution”, *Journal of Monetary Economics* 22, pp. 117-131.

[28] Salas, E., M. Rosen, D. Diaz-Granados (2010): “Expertise-Based Intuition and Decision Making in Organizations”, *Journal of Management* 36 (4), pp. 941–973.

[29] Savage, L. (1954): *The Foundations of Statistics.* John Wiley & Sons, Inc.

[30] Schotter, A. and I. Trevino (2014): “Belief Elicitation in the Laboratory”, *Annual Review of Economics* 6, pp. 103-128.

[31] Shefrin, H., & Statman, M. (1985). “The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence”, *The Journal of Finance* 40(3), pp. 777-790.

[32] Spiegler, R. (2016): “Bayesian Networks and Boundedly Rational Expectations", *Quarterly Journal of Economics* 131, pp. 1243-1290.

[33] Tversky, A. and D. Kahneman (1974), “Judgement under Uncertainty: Heuristics and biases”, *Science* 185, pp. 1124–1131.

[34] Vazsonyi, A. (1999) “Which Door Has the Cadillac?”, *Decision Line* (December/January issue), pp 17–19.

[35] Wasserman, E. and R. Miller (1997): “What’s Elementary About Associative Learning?” *Annual Review of Psychology* 48, pp. 573-607.