Identification of Beliefs in Decision Making

Tiago Caruso* Greg Kubitz†

October 28, 2017

Abstract

This paper studies a binary choice model where an agent makes a decision that is informed by his beliefs after observing a public signal. This model generalizes to a wide range of economic environments which require econometricians to estimate the beliefs of agents. With minimal structure imposed on the agent's utility function, we characterize the structure of information needed to identify the beliefs of the agent after observing both signals and decisions. We find that the information must be sufficiently convincing and dense for the agent's beliefs to be point identified. When the full range of information is relaxed, we show how the agents beliefs can be partially identified. Additionally, we explicitly show how the econometrician can construct the sharpest boundaries around the agents beliefs, as she observes signals and decisions.

JEL classification codes: C44, D81, D83

1 Introduction

The choice behavior of an agent in an uncertain environment is of central importance in many economic settings. This behavior depends on the agent's subjective beliefs about how the uncertainty may be resolved. Observed choices therefore have information about the agent's beliefs, and this same information can be used to predict the agent's future choice behavior. Previous literature has examined how to identify agent's beliefs from a rich choice set[1] or, as in Lu (2016), from discrete choice where one action gives constant payoff regardless of the realized state. This paper examines the ability of an econometrician to identify an agent's beliefs solely from binary choice data and the public information used by the agent to make his decision.

We first show that when the size of the state space is two, binary choice data and public information signals are sufficient to identify the agent's prior as long as

*UCLA

†NYU Stern, corresponding author

[1] In Arieli and Mueller-Frank (2017) agents have an uncountable action space and in Dillenberger et al (2014) agents choose between menus of acts.

the information set contains signals that are sufficiently dense and convincing. This condition is formalized in Assumption 2. We next show a sufficient condition, stated in Assumption 5, for identification when there are more than two states. In the setting of a binary choice, the requirements on the information set to guarantee identification of the agents beliefs are strong, and while Theorem 2.4 shows this is only a sufficient condition, we show that this condition cannot be meaningfully relaxed. Lastly, we return to a state space with two states to show how sharp bounds can be placed on an agent's prior (and similarly, posterior) belief distribution from dynamic choice data. The bounds require the econometrician to use all past observations of the agent, not just the most recent.

Our setting has the standard components of a static decision problem under uncertainty: a set of states which contain all payoff relevant information and a set of intermediate actions that yield a payoff to the agent that depends on the realized state. Payoffs in each given state are assumed to be known and the agent is risk-neutral. The agent's subjective probability distribution over the states is unknown and will be of primary interest to the econometrician. Prior to making a decision, the agent observes a public signal that allows him to revise his beliefs via Bayesian updating. Therefore, the agent will choose the action that maximizes expected payoffs given his posterior subjective probability distribution. The econometrician observes both the signal and the agent's choice and uses this information to make inferences on the agent's beliefs.

We first consider a state space with two states that yield payoffs for the agent so that his decision problem is non-trivial. If the decision of the agent is to invest or not invest in a project, then we assume that investing will lead to a higher payoff in one state, while not investing gives higher payoffs in the other state. We show that each distinct prior belief held by the agent will lead to a different choice rule. Specifically the agent's investment decision is a cut-off rule in terms of the likelihood ratio of the signal he receives. Additionally, the cut-off point is strictly decreasing in the agent's initial optimism about the investment. The uniqueness of this choice rule allows the econometrician to place sharp and non-trivial bounds on the agent's beliefs after observing his signal and subsequent choice. Moreover, if the information set is rich enough, namely if the range of the likelihood ratio of signals in the set is dense on the positive real numbers, then the econometrician can point identify the agent's prior after observing repeated pairs of signals and decisions.[2]

We extend our analysis to allow for any finite number of states. Additional states allow the agent to be optimistic (or pessimistic) about an action choice for more than one reason. For example, the agent may choose an investment either because they believe that significant returns is a likely outcome or because the outside option of not investing is likely to have very low returns. This qualitative difference increases the dimensionality of the problem of inference for the econometrician. The agents choice rule is a cut-off whose boundary is a hyperplane in terms of the signal's likelihood ratios of all other states relative to a single state, say state one.

Because the agent is still making a binary decision, a single observation by the econometrician of a signal and choice will reveal how optimistic the agent is about the investment but not why he is optimistic. Therefore, in order for the econometrician to place a sharp bound on the agent's belief about the likelihood the investment yields high returns she must observe the agent choosing to invest after receiving a signal that not only increases the likelihood that the returns on investment are high, but also says nothing about the likelihood of low returns of the outside option. A sufficient condition for the econometrician to point identify the agent's beliefs from his choice rule, the range of the likelihood ratio for each pair of states where one in state the agent would prefer to invest and the other would prefer not to invest must be dense on the positive reals with all other likelihood ratios being arbitrarily small. This sufficient condition is

[2] We focus our paper in the individual interpretation of the decision making process, when we observe an individual making the same decision repeatedly. However, the results of the paper, can be understood in the context of a population of identical individuals receiving conditionally independent signals and the variable of interest would be the average belief in the population.

formalized in Assumption 5. While this condition seems unnecessarily strong, we show that it cannot be significantly weakened.

Lastly, we return to the setting of two states to analyze inferences that can be made on an agent's subjective belief distribution when the econometrician observes a series of signals and decisions between each signal. We call this process the identification of the agent's priors. For a given sequence of signals, if we know the conditional distribution of signals, non-degenerate priors hold a one-to-one relation with posteriors. Using this relationship, we show that with any finite number of observations, we can construct sharp (non-trivial) bounds on the agent's priors by looking at the decisions made when the agent is least sure about the optimal decision for their current state. Since we are looking at these marginal decisions, the sharpest bounds will depend on the entire history of signals and decisions. As such, to best infer the agent's priors and current posterior beliefs, the econometrician should be concerned with all past observations.

This paper follows the tradition founded by Ramsey (1931) and Savage (1972), who describe probability as a subjective feature based on the actions they could generate. According to Savage's view, if we could offer the agent a menu of lotteries, we could find the one which makes him indifferent between actions. From this it would be possible to recover the probability he assigned to each state of the world. However, when inferring subjective probabilities from given choice data, an econometrician cannot arbitrarily pick the choice set of the agent. Therefore recent literature has examined what can be learned about these probabilities for a given choice environment. We only offer the agent binary choices but infer these probabilities from how the choice changes as the agent makes choices given signals from a rich information set.

McFadden et al (1973) describes how probabilistic choice models can be estimated from discrete choice data of a heterogeneous group of people which can be used to predict future choice behavior. More recently, Arieli and Mueller-Frank (2017) shows that if the action set is uncountable, then almost any continuous utility function will allow an econometrician to infer beliefs from actions. Other studies use discrete choice models to identify specific characteristics of choice behavior rather than attempting to identify the beliefs themselves. Dillenberger et al (2014) shows that an agent's choice of menu and then subsequent action from the menu can reveal characteristics about the type of information he expects to receive between the two choices. Specifically, an analyst can identify the fineness of information an agent expects to receive from the menu choice. Lu (2016) also analyzes a discrete choice model where one action is an outside option that offers the agent a fixed payoff in any state. When the value of this outside option changes, changes from choosing a state dependent action and the outside option. He uses this value to compute a test function of the agent and directly uses the properties of the test function to analyze information that the agent may have. Therefore without directly inferring the agents beliefs, an analyst can identify the value of a specific menu to an agent, or detect if that agent is biased. In our model, the choice space is restricted to the simplest example of a binary decision and we do not require the outside option to be state independent. Because of this, the agent is deciding whether to invest of not based on public information that could effect both the value of investing or not investing, rather than just changed in value to not investing.

An alternative approach to inferring beliefs is to directly elicit them from agents. Savage (1971) formalized the discussion on how to elicit probabilities using scoring rules. Additional methods of direct elicitation are discussed in De Finetti (1974), Kadane and Winkler (1988), and Karni (2009). [3] Manski (2004) offers a comprehensive survey of this literature, defending the survey based approach for determining beliefs. However, this approach can suffer from misreporting due to inability or unwillingness to give correct answers. Misreporting can lead to predictions that are not consistent with actual choice behavior. For example, Nyarko and Schotter (2002) finds that the difference

[3] Direct elicitation is used in Delavande (2008) to estimate the expectations of outcomes of using alternative contraceptive methods and Lochner (2003) use directly elicited expectations of being arrested to construct a model of utility for criminal behavior. Delavande et al (2011) reviews the methods of eliciting beliefs in developing countries and argue that probabilistic questions can be accurately answered even by individuals with very low education levels.

between true beliefs and beliefs elicited from a fictitious scenario can vary substantially. Therefore, results that rely on the on beliefs that are directly elicited from agents should be consistent with inference of beliefs from choice data.

The rest of the paper is organized as follows. Section 2 describes the general model and derive the results for the identification of beliefs. Section 3 uses a modified model to address the issue of identification of priors. Section 4 concludes.

2 Identification of Beliefs

The decision making problem There is one agent making a binary decision in an uncertain world with two possible states.[4] The agent does not know which world he is in, but has a belief about the likelihood of each state.[5] Let the decision of the agent be whether or not to invest ($d = i$ or ni) in a world that that is either good or bad for investment ($s = h$ or ℓ)[6]. Before making his decision, the agent receives a signal $x \in X$, say available economic indicators, which allows him to update his beliefs from known conditional distributions, $f(x|h)$ and $f(x|\ell)$.

We will incorporate the uncertainty in our model in two different ways. In the first one, which we consider throughout this section, there is an independent identically distributed move of nature in each period. Therefore a new state is drawn before each signal is received and decision is made. In this case the belief of the agent, denoted by $p = (p(h), 1 - p(h))$, can be interpreted as the agent's belief about the distribution of this move of nature, and it remains fixed for each decision. Therefore the agent only considers the signal realized in the current decision process when deciding which action to take. In the next section we will allow the agents beliefs to be evolve as signals are aggregated.

[4]In a later subsection we will discuss our results for a finite number of states S.

[5]We treat subjective beliefs as defined in Anscombe and Aumann (1963). It is shown in Gul (1992) that these beliefs exist in a setting with a finite number of states.

[6]In this sense a binary decision is the hardest case since working with more decisions would make the identification strategies easier since every decision would provide more information.

We restrict the signal space to be countable but potentially infinite in every state of the world. Although informative about the state, the signal does not directly influence the agent's payoff. After observing it, the agent updates his beliefs according to Bayes' rule and his posterior probability is given by:

$$p(h|x) = \frac{f(x|h)p(h)}{f(x)} = \frac{f(x|h)p(h)}{f(x|h)p(h) + f(x|\ell)(1 - p(h))} \tag{1}$$

His payoffs are characterized by the values, $v(i, h)$, $v(ni, h)$, $v(i, \ell)$, and $v(ni, \ell)$, the payoff of each decision in each state. We define $\tilde{v}(s) \equiv v(i, s) - v(ni, s)$, the difference between investing and not investing in each state. In order to keep the decision problem from becoming trivial, we make an assumption on these payoffs.

Assumption 1. $\tilde{v}(h) > 0 > \tilde{v}(\ell)$

This condition implies that the agent wants to invest in the high state and does not want to invest in the low state. Since he does not know in which state he is, he will invest exactly when

$$E_{p(s|x)}\left[\tilde{v}(s)\right] \geq 0. \tag{2}$$

The econometrician problem The econometrician sees the same signals as the agent and she observes his decisions. She also knows the conditional distribution of signals, $f(x|s)$, so every observation can be summarized as a pair (d_i, x_i). Her main objective is to identify the agent's initial beliefs p. We follow the notation and definition of identification in Manski (2003).

Definition 1. *The set of priors that is observationally equivalent to* p *is given by* $H[p] \equiv \left\{p' \in \Delta\mathbb{R}^S : \{d(p, x_i)\}_{i=1}^{\infty} = \{d(p', x_i)\}_{i=1}^{\infty}\right\}.$

Each $p' \in H[p]$ is consistent with the decisions of an agent with prior p for any information that could be received. If $H[p]$ is a strict subset of $\Delta\mathbb{R}^S$, p is partially identified, and if $H[p]$ is a singleton, then we say that p is point identified. If $H[p]$ is a singleton for each p, then we say the beliefs of the agent are point identified.

Identification result We can manipulate equation (2) to characterize when the agent will invest given his posteriors and payoffs.

$$p(h|x) \geq \frac{\tilde{v}(h)}{\tilde{v}(h) - \tilde{v}(\ell)} \equiv \overline{d} \tag{3}$$

Hereafter $\overline{d}$ can be interpreted as the threshold posterior of the decision rule, where the agent is indifferent between investing and not investing.[7] Combining equations (1) and (3), we can write the decision rule in terms of the agents beliefs.

$$\frac{f(x|h)p(h)}{f(x|h)p(h) + f(x|\ell)(1 - p(h))} \geq \overline{d} \tag{4}$$

We will use the likelihood ratio as a summary statistic for the information that is observed by the agent and the econometrician. This likelihood ratio is defined by

$$\gamma(x) \equiv \frac{f(x|h)}{f(x|\ell)} \text{ for each } x \in X.$$

We can now write the decision rule for the agent in terms of his prior belief about the distribution of states, our variable of interest. The agent chooses to invest given signal x exactly when

$$p(h) \geq \frac{\overline{d}}{\overline{d} + \gamma(x)\left(1 - \overline{d}\right)}. \tag{5}$$

In order for the econometrician to point identify the agent's beliefs there must be enough information available. Not only must there be information that can sway both an optimistic and pessimistic agent, this information must also be dense enough to differentiate beliefs that are relatively similar. These requirements are formalized in the following full range assumption on the likelihood ratio of the signals in the information

[7] Even if the econometrician does not know agent's utilities, assumption 1.1 is sufficient to guarantee that $\overline{d} \in (0, 1)$.

set.

Assumption 2. *For each $y \in (0, \infty)$ and $\epsilon > 0$ there exists a signal x such that $\gamma(x) \in B_\epsilon(y)$.*[8]

An information set which satisfies the full range assumption is sufficient for the econometrician to identify the agent's beliefs. This is shown in the following result.

Proposition 2.1. *If the state is relevant and likelihood ratio of the signals has full range then the initial beliefs of the decision maker are point identified.*

Proof. Let $p = (p(h), 1-p(h))$ and $p' = (p'(h), 1-p'(h))$ be two distinct beliefs. Without loss of generality assume that $p(h) > p'(h)$. We define $\hat{p}$ and $\hat{y}$ such that

$$\hat{p} = \frac{p(h) + p'(h)}{2} \text{ and } \hat{y} = \frac{(1-\hat{p})\overline{d}}{\hat{p}(1-\overline{d})}$$

Since $\hat{p} \in (0, 1)$ and $\overline{d} \in (0, 1)$ we have that $\hat{y} \in (0, \infty)$. By the full range assumption, there exists an $\hat{x}(\varepsilon)$ such that $\gamma(\hat{x}(\varepsilon)) \in B_\varepsilon(\hat{y})$, for all $\varepsilon > 0$. For small enough $\varepsilon > 0$, we have that

$$p(h) > \frac{\overline{d}}{\overline{d} + \gamma(\hat{x}(\epsilon))\left(1-\overline{d}\right)} \approx \hat{p} > p'(h).$$

With the signal of $\hat{x}(\varepsilon)$, $d(p, \hat{x}(\varepsilon)) = i$ and $d(p', \hat{x}(\varepsilon)) = ni$. This shows that the two beliefs p and p' are not observational equivalent. Since p and p' are arbitrary beliefs of the agent, all beliefs are identified. □ □

This result closely resembles the seminal result of Savage (1972). Here, instead of offering the agent a pool of lotteries between outcomes, the agent may observe a range of signals that allows the econometrician to pin down his initial belief over the distribution of the states.

Additionally, the full range assumption is necessary for point identification of the agent's beliefs. If there is an interval on the positive real line for which there is no

[8]Since the rationals are countable and dense on the real line, any information set X which $\gamma(\cdot)$ maps onto the rational numbers would satisfy this property.

signal x whose likelihood ratio $\gamma(x)$ lies within the interval (the range of $\gamma(x)$ is not dense on the positive real line), then there will be two different beliefs p and p' which will be observationally equivalent. For any such interval, there is a pair or priors that can only be differentiated by a signal whose likelihood ratio falls within that interval.

In the following sections, we show that when the assumption of full range is relaxed the econometrician can still partially identify the priors of the agent.

2.1 Partial identification without full range

We now weaken the assumption of full range of information and consider partial identification when the range of the likelihood ratio is limited. Intuitively, this may happen when there is no information that is strong enough to either convince very pessimistic agents to invest or keep very optimistic agents from investing. Specifically, for the given information set, the range of the likelihood function is bounded, either above, away from zero, or both.

$$\text{Range}(\gamma(x)) \subset (\underline{\alpha}, \overline{\alpha}) \subsetneq \mathbb{R}_+$$

Within these bounds, we still assume that information is sufficiently dense as in the full range assumption from the previous section. In order for the econometrician to know the range of prior beliefs that can be identified, she must have information about the agent's utility function. It is sufficient for her to know the value of $\overline{d}$, which determines the decision rule of the agent.

Assumption 3. *The econometrician knows* $\overline{d}$.

With this assumption, the econometrician can point identify prior beliefs of an agent who can be persuaded by available information. The beliefs of agents who cannot be persuaded by any of the available information will be partially identified.

Proposition 2.2. *All beliefs of the agent for which $p(h)$ belongs to the interval*

$$\left(\frac{\overline{d}}{\overline{d}+\overline{\alpha}\left(1-\overline{d}\right)}, \frac{\overline{d}}{\overline{d}+\underline{\alpha}\left(1-\overline{d}\right)}\right)$$

are point identified. If the belief is such that $p(h)$ is the right of the interval, then

$$H\left[p\right]=\left[\frac{\overline{d}}{\overline{d}+\underline{\alpha}\left(1-\overline{d}\right)}, 1\right],$$

and if it is to the left of the interval

$$H\left[p\right]=\left[0, \frac{\overline{d}}{\overline{d}+\overline{\alpha}\left(1-\overline{d}\right)}\right].$$

This proof follows the steps of Proposition 1 for the beliefs contained within the point identified interval. If the agent's beliefs are such that his decision does not change, his beliefs can still be partially identified. For example, if the agent always decides to invest, then his beliefs are such that $p(h)$ is observationally equivalent to 1.

Note that without Assumption 3, or any knowledge about the agent's utility function, $\overline{d}$ can take on any value between 0 and 1. In this case, the econometrician would not know which beliefs are point identified and which are just partially identified. The region of priors that are point identified is determined by $\overline{d}$.

2.2 Partial identification with finite information

Point identification, even over a small region of the agent's priors, requires an information set that is infinite. Nevertheless, with assumptions 2.1 and 2.3, we can partially identify the prior beliefs of the agent with a finite information set. The tightness of these bounds will depend on how dense and how persuasive these observations are.

We now define $\gamma^{*}(p)$ as the likelihood ratio that would make the agent indifferent between investing and not investing given beliefs p.

$$p(h) = \frac{\overline{d}}{\overline{d} + \gamma^*(p)\left(1 - \overline{d}\right)} \Rightarrow \gamma^*(p) = \frac{\overline{d}(1 - p(h))}{1 - \overline{d}}$$

Note that the agent will decide to invest after receiving signal x if and only if $\gamma(x) > \gamma^*(p)$. Using this, we can impose bounds on $p(h)$.

Assume we have M possible signals given the information set X, $|X| = M$. We can order the M potential observations according to the size of their likelihood ratio, i.e. $\gamma(x_{m-1}) \leq \gamma(x_m)$, for all $1 \leq m \leq M$. Then let $\tilde{m}$ be the smallest m for which the agent decides to invest after receiving signal $x_{\tilde{m}}$. This would imply (for $\tilde{m} > 1$) that $\gamma(x_{\tilde{m}}) > \gamma^*(p) > \gamma(x_{\tilde{m}-1})$ and the identification region is given by

$$H_M[p(h)] = \left[\frac{\overline{d}}{\overline{d} + \gamma(x_{\tilde{m}})(1 - \overline{d})}, \frac{\overline{d}}{\overline{d} + \gamma(x_{\tilde{m}-1})(1 - \overline{d})}\right]. \tag{6}$$

If $\tilde{m} = 1$, then the agent always invests, and the identification region is

$$H_M[p_h] = \left[\frac{\overline{d}}{\overline{d} + \gamma(x_1)(1 - \overline{d})}, 1\right].$$

On the other hand, if the agent never decides to invest, then the identification region is

$$H_M[p_h] = \left[0, \frac{\overline{d}}{\overline{d} + \gamma(x_M)(1 - \overline{d})}\right].$$

2.3 Multiple States

We now consider a situation where the agent is making a binary decision, but there are more than two possible states of the world. This allows us to capture a situation where the agent has multiple reasons for why he would, or would not, want to invest. The decision rule is now a function of the agents posterior beliefs, $p(s|x)$, and gain from investing, $\tilde{v}(s)$, for each state $s \in \{1, \ldots, S\}$. From Bayes' rule, these posterior beliefs are given by

$$p(s|x) = \frac{f(x|s)p(s)}{\sum_{s' \in S} f(x|s')p(s')} \text{ for all } s \in S. \tag{7}$$

Without loss, we order the states according to the gain in deciding to invest such that: $\tilde{v}(S) > \tilde{v}(S-1) > \cdots > \tilde{v}(1)$. As in the two state model, we require the agent to prefer investing in at least one state and to prefer not investing in at least one different state.

Assumption 4. *There exists a state s^* such that $\tilde{v}(s^*) > 0 > \tilde{v}(s^* - 1)$.*

For a given signal x, the agent will decide to invest exactly when the expected gain from investing is positive. Thus the agent invests when

$$E_{p(s|x)}[\tilde{v}(s)|x] = \sum_{s \in S} \tilde{v}(s)p(s|x) = \sum_{s \in S} \tilde{v}(s)p(s)f(x|s) \geq 0. \tag{8}$$

In order for the decision rule to be non-trivial, we assume that the agent is uncertain about whether the state of the world is good for investment. In this case, we say the agent has a prior belief that may be identified by the econometrician for a given information set. An agent who is certain that the state is good, or certain that it is bad, will always make the same decision, and the econometrician will not be able to determine why this decision is being made.

Definition 2. *A prior $p \in \Delta(\mathbb{R}^S)$ is relevant if $p(r) > 0$ and $p(t) > 0$ for some states $r < s^* \leq t$. Let the set of relevant priors be denoted by $P \subset \Delta(\mathbb{R}^S)$.*

For the econometrician to identify the beliefs of the agent in this broader setting, there must be information that not only separates beliefs by how optimistic they are, but also tell us why the agent is optimistic or pessimistic. The full range assumption which gives the econometrician sufficient information for identification is similar to the corresponding assumption when there are only two states. For each pair of states where one state is good for investing, and the other bad, we must have information that can sway both optimistic and pessimistic individuals, and be dense enough to separate beliefs that are similar, *in this particular dimension.* This assumption is formalized below.

Assumption 5. *For each $y \in \mathbb{R}_+$, $\epsilon > 0$, $r < s^*$ and $t \geq s^*$ there is an $x \in X$, such that $f(x|s) = 0$ for all $s \neq r, t$ and $\frac{f(x|t)}{f(x|r)} = \gamma_{tr}(x) \in B_\epsilon(y)$.*

The following lemma shows that the econometrician can point identify the priors of the agent when information satisfies full range and we can find two states, one good for investing, where the agents have different beliefs. The argument is similar to the proof of identification when there are two states. The proof is relegated to the appendix.

Lemma 2.3. *For any $p, p' \in P$, if there are two states $r < s^* \leq t$ such that $p'(t)$ and $p'(r)$ are non zero and $\frac{p(t)}{p'(t)} \neq \frac{p(r)}{p(r)}$, then under the assumption of full range in multiple states, $p' \notin H[p]$.*

To show identification of the set of relevant priors, it is now enough to show that any two distinct, relevant priors have the stated property of Lemma 1. Again the proof is in the appendix.

Theorem 2.4. *The set of relevant priors is point identified if X satisfies full range in multiple states.*

While the full range in multiple states assumption appears strong, it cannot be weakened significantly. To see this, consider two priors that are indexed by 3 parameters, $k_1 > 1$, $k_2 > 1$, $n \geq 3$. Let $p(s; k_1, k_2, n) = 1 - \frac{1}{(k_1+1)n} - \frac{1}{k_2 n}$, $p(r; k_1, k_2, n) = \frac{1}{(k_1+1)n}$, $p(t; k_1, k_2, n) = \frac{1}{k_2 n}$, and $p(s'; k_1, k_2, n) = 0$ for all $s' \neq r, s, t$, where $r < s^*$ and $t \geq s^*$. Let $p'(s; k_1, k_2, n) = 1 - \frac{1}{k_1 n} - \frac{1}{(k_2+1)n}$, $p'(r; k_1, k_2, n) = \frac{1}{k_1 n}$, $p'(t; k_1, k_2, n) = \frac{1}{(k_2+1)n}$, and $p'(s'; k_1, k_2, n) = 0$ for all $s' \neq r, s, t$, where $r < s^*$ and $t \geq s^*$. Note that a decision maker with a prior of p is more optimistic than one with a prior of p' for all $k_1 > 1, k_2 > 1$, and $n \geq 3$. In order for $p'(k_1, k_2, n) \notin H[p(k_1, k_2, n)]$, there must be an

$x \in X$ such that the following inequalities hold.

$$0 \leq f(x|s)\left[\frac{(k_1+1)k_2n-(k_1+1)-k_2}{(k_1+1)k_2n}\right]\tilde{v}(s) + f(x|r)\frac{1}{(k_1+1)n}\tilde{v}(r)+f(x|t)\frac{1}{k_2n}\tilde{v}(t) \tag{9}$$

$$0 > f(x|s)\left[\frac{k_1(k_2+1)n-k_1-(k_2+1)}{k_1(k_2+1)n}\right]\tilde{v}(s) + f(x|r)\frac{1}{k_1n}\tilde{v}(r)+f(x|t)\frac{1}{(k_2+1)n}\tilde{v}(t) \tag{10}$$

If $s \geq s^*$, then (10) implies that

$$\gamma_{sr}(x) < \left[\frac{k_2+1}{k_1(k_2+1)n-k_1-k_2-1}\right]\frac{-\tilde{v}(r)}{\tilde{v}(s)}.$$

For large n, both decision makers are almost certain that they are in a state where investing is the optimal choice. In order to differentiate between the two, they must receive a piece of information that almost never occurs in this state. Condition (10) implies that the more pessimistic player must decide not to invest. Therefore $\gamma_{sr}(x) \to 0$ as $n \to \infty$.

Also, if $\gamma_{sr}(x) = 0$. Then condition (9) and (10) combine to require a piece of information for which the optimistic player invests, while the pessimistic player does not

$$\frac{-\tilde{v}(r)}{\tilde{v}(t)}\frac{k_2}{(k_1+1)} < \gamma_{tr}(x) < \frac{-\tilde{v}(r)}{\tilde{v}(t)}\frac{k_2+1}{k_1}.$$

Therefore $\gamma_{sr}(x)$ must have full range while $\gamma_{sr}(x) = 0$. This argument can be made for each triple of states, (s, r, t), where $r < s^*$ and $t \geq s^*$.

3 Identification of priors

In this section, we depart from the idea of identifying fixed, steady state beliefs and focus on the situation where there is only one move of nature that describes the state of the world and a series of signals about this underlying state. Provided that individuals have strictly positive priors and that we understand how they update their beliefs, there is a one to one mapping between priors and posteriors. Thus although we will focus on estimating priors, both objects are equivalent for predicting the agent's behavior.

We return to the binary world and binary decision framework. Now the agent observes a sequence of signals, $X_M = \{x\}_{m=1}^M$, and after each signal he updates his posteriors and makes a decision, but he does not observe payoffs. In our investment problem, this assumption would correspond to the idea that the agent observes M signals and makes M subsequent decisions, before the state of the economy is revealed. We assume that the signals are independent, conditional on the state of the world.

The econometrician observes the sequence of signals X_M and the sequence of decisions $D_M = \{d\}_{m=1}^M$. From these observations, she creates the sharpest bounds on the agent's prior. The sharpest bounds constitute the smallest identification region for the given sequence of observations. After each signal, the agent will invest exactly when the expected value of investment is positive, given his current posterior beliefs: $E_{p(s|X_m)}[\tilde{v}(s)|X_m] \geq 0$. This holds whenever $p(h|X_m) \geq \overline{d}$.

In order to highlight the one-to-one relationship between posteriors and priors, we take the log of the ratio of the posterior. Additionally, we can separate the effect that each signal has on the posterior of the agent. The assumption that the signals are conditionally independent given the state allows us to describe this relationship in the following equation:

$$\ln\left(\frac{p(h|X_M)}{p(\ell|X_M)}\right) = \ln\left(\frac{f(x_1|h)\cdots f(x_M|h)p(h)}{f(x_1|\ell)\cdots f(x_M|\ell)p(\ell)}\right) = \ln\left(\prod_{m=1}^{M}\gamma(x_m)\left(\frac{p(h)}{p(\ell)}\right)\right)$$
$$= \sum_{m=1}^{M}\ln\left(\gamma(x_m)\right) + \ln\left(\frac{p(h)}{p(\ell)}\right) \tag{11}$$

The following proposition provides the sharpest bounds the econometrician can impose on the agent for a given sequence of signals and decisions.

Proposition 3.1. *The sharpest bounds in the region of identification of priors are given by*

$$H\left[\ln\left(\frac{p(h)}{p(\ell)}\right)\right] =$$

$$\left[\max_{\substack{1\le m\le M\\ d_m=i}} \ln\left(\frac{\overline{d}}{1-\overline{d}}\right) - \sum_{m=1}^{M}\ln\left(\gamma(x_m)\right), \min_{\substack{1\le m\le M\\ d_m=ni}} \ln\left(\frac{\overline{d}}{1-\overline{d}}\right) - \sum_{m=1}^{M}\ln\left(\gamma(x_m)\right)\right].$$

Proof. : In every period j, the agent will decide to invest if and only if

$$\sum_{m=1}^{j}\ln\left(\gamma(x_m)\right) + \ln\left(\frac{p(h)}{p(\ell)}\right) \ge \tilde{d}.$$

Thus whenever the agent chooses to invest, given observation pair $(\{x\}_{m=1}^{j}, d_j)$, we can impose a lower bound for $\ln(p(h)/p(\ell))$. The sharpest lower bound is derived from the period where the agent has the lowest posterior but still decides to invest. This is exactly

$$\max_{\substack{1\le m\le M\\ d_m=i}} \ln\left(\frac{\overline{d}}{1-\overline{d}}\right) - \sum_{m=1}^{M}\ln\left(\gamma(x_m)\right).$$

Similarly the sharpest upper bound is formed from the period where the agent has the highest posteriors given that he decides not to invest. □ □

Figure 1 depicts an example of the construction of these bounds. The econometrician knows $\overline{d}$, which is denoted as the horizontal line, and she also knows the summation of the log likelihood ratio of the sequence of signals. On the figure, this is the vertical distance between the log of the initial beliefs and the log of the current beliefs. In this example, the upper bound is derived from observation 4, which is the highest posterior for which the agent chooses not to invest. The lower bound is constructed from observation 7. We construct each bound by subtracting this vertical distance from the $\overline{d}$ in each key observation. After finding the bounds of $\ln(p(h)/p(\ell))$ we can use them to derive bounds for $p(h)$ and $p(\ell)$.

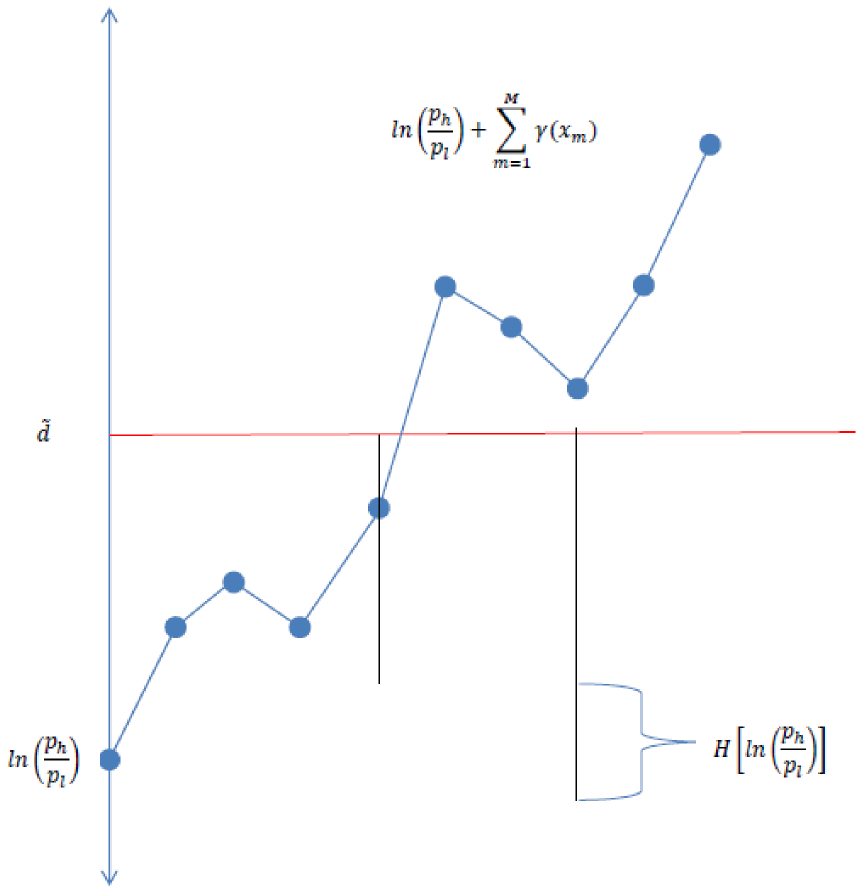

Figure 1: Constructing the bounds on priors.

There is no reason to believe that the inference about the prior will necessarily improve asymptotically. However, since we are working with conditionally independent signals, the posterior of the agent approaches the true state of the world, so identification becomes less important.

From this example it is clear that all observations can be important in forming bounds on the priors of the agent. The econometrician should be concerned with all the past decisions, not because the present world mechanically depends on the past, but because past observations tighten the bounds on the agent's beliefs.

4 Concluding remarks

In this paper we direct our attention to a classical question in economics: what can we infer about an agent's beliefs about the state of the world if we observe his decisions? We believe that there is a growing class of problems where economists are interested in an agent's beliefs and are comfortable making assumptions about his utility function, or at least about his decision rule. We show that even in binary decision making with multiple states, agents beliefs can be precisely identified if we observe a range of signals that are sufficiently rich. Moreover, multiple observations of an agent's choice behavior provide sharp bounds on an agent's belief distribution which can be used for predicting future choices.

Appendix

Lemma 1

For any $p, p' \in P$, if there are two states $r < s^* \leq t$ such that $p'(t)$ and $p'(r)$ are non zero and $\frac{p(t)}{p'(t)} \neq \frac{p(r)}{p(r)}$, then under the assumption of full range in multiple states, $p' \notin H[p]$.

Proof. We take $\frac{p(r)}{p'(r)} > \frac{p(t)}{p'(t)}$ (the argument is symmetric for the case where $\frac{p(r)}{p'(r)} < \frac{p(t)}{p'(t)}$).

Then,

$$\frac{p(r)}{p'(r)} > \frac{p(t)}{p'(t)} \Rightarrow \frac{\tilde{v}(r)p(r)}{\tilde{v}(r)p'(r)} > \frac{\tilde{v}(t)p(t)}{\tilde{v}(t)p'(t)}$$

$$\Rightarrow \frac{\tilde{v}(r)p(r)}{\tilde{v}(t)p(t)} < \frac{\tilde{v}(r)p'(r)}{\tilde{v}(t)p'(t)} \Rightarrow \frac{-\tilde{v}(r)p(r)}{\tilde{v}(t)p(t)} > \frac{-\tilde{v}(r)p'(r)}{\tilde{v}(t)p'(t)}$$

For $p' \notin H[p]$, we must have that $d(p,x) \neq d(p',x)$ for some $x \in X$. Consider $X_{rt} = \{x \in X : f(x|s) = 0 \text{ for } s \neq r, t\}$. Then for $x \in X_{rt}$, $d(p,x) \neq d(p',x)$ exactly when

$$\tilde{v}(t)p(t)f(x|t) + \tilde{v}(r)p(r)f(x|r) < 0 \text{ and } \tilde{v}(t)p'(t)f(x|t) + \tilde{v}(r)p'(r)f(x|r) \geq 0.$$

Dividing by $f(x|r)$

$$\tilde{v}(t)p(t)\gamma_{tr}(x) + \tilde{v}(r)p(r) < 0 \text{ and } 0 \leq \tilde{v}(t)p'(t)\gamma_{tr}(x) + \tilde{v}(r)p'(r)$$

Combining the inequalities,

$$\frac{-\tilde{v}(r)p'(r)}{\tilde{v}(t)p'(t)} \leq \gamma_{tr}(x) < \frac{-\tilde{v}(r)p(r)}{\tilde{v}(t)p(t)}$$

Now let

$$y = \frac{1}{2}\left(\frac{-\tilde{v}(r)p'(r)}{\tilde{v}(t)p'(t)} + \frac{-\tilde{v}(r)p(r)}{\tilde{v}(t)p(t)}\right), \text{ and } \varepsilon = \frac{1}{3}\left(\frac{-\tilde{v}(r)p(r)}{\tilde{v}(t)p(t)} - \frac{-\tilde{v}(r)p'(r)}{\tilde{v}(t)p'(t)}\right)$$

Then there is an $x \in X_{rt} \subset X$ such that $\gamma_{tr}(x) \in B_\varepsilon(y)$ and for this x, $d(p,x) \neq d(p',x)$ and therefore $p' \notin H[p]$. □

Theorem 1

The set of relevant priors is point identified if X satisfies full range in multiple states.

Proof. Let $p, p' \in P$ where $p \neq p'$. Then there is an $s \in S$, where $p(s) \neq p'(s)$. Without loss of generality, we take $0 \leq p(s) < p'(s)$.

If $s < s^*$ then there is a state $t \geq s^*$ where $p'(t) > 0$, which implies $\frac{p(s)}{p'(s)}$ and $\frac{p(t)}{p'(t)}$

are non-negative numbers. If $\frac{p(s)}{p'(s)} \neq \frac{p(t)}{p'(t)}$, then by Lemma 2.2 we know that $p' \notin H[p]$. If $\frac{p(s)}{p'(s)} = \frac{p(t)}{p'(t)}$, then since $\sum_{i=1}^{S} p(i) = \sum_{i=1}^{S} p'(i) = 1$, there must be an s' where $0 \leq p'(s') < p(s')$. If $s' < s^*$ then $\frac{p(s')}{p'(s')} \neq \frac{p(t)}{p'(t)}$, and we can use Lemma 2.2 to show that $p' \notin H[p]$. If $s' \geq s^*$ then $\frac{p(s)}{p'(s)} \neq \frac{p(s')}{p'(s')}$, and we can again use Lemma 2.2.

Now if $s \geq s^*$, then there is a state $r < s^*$ where $p'(r) > 0$. We can now repeat the argument above to show that $p' \notin H[p]$.

Since p and p' were arbitrary, then $H[p]$ is a singleton for all $p \in P$, and P is point identified.

□

References

Anscombe FJ, Aumann RJ (1963) A definition of subjective probability. Annals of mathematical statistics pp 199–205

Arieli I, Mueller-Frank M (2017) Inferring beliefs from actions. Games and Economic Behavior 102:455–461

De Finetti B (1974) Theory of Probability: A Critical Introductory Treatment. Transl. by Antonio Machi and Adrian Smith. J. Wiley

Delavande A (2008) Pill, patch, or shot? subjective expectations and birth control choice*. International Economic Review 49(3):999–1042

Delavande A, Giné X, McKenzie D (2011) Measuring subjective expectations in developing countries: A critical review and new evidence. Journal of Development Economics 94(2):151–163

Dillenberger D, Lleras JS, Sadowski P, Takeoka N (2014) A theory of subjective learning. Journal of Economic Theory 153:287–312

Gul F (1992) Savages theorem with a finite number of states. Journal of Economic Theory 57(1):99–110

Kadane JB, Winkler RL (1988) Separating probability elicitation from utilities. Journal of the American Statistical Association 83(402):357–363

Karni E (2009) A mechanism for eliciting probabilities. Econometrica 77(2):603–606

Lochner L (2003) Individual perceptions of the criminal justice system. Tech. rep., National Bureau of Economic Research

Lu J (2016) Random choice and private information. Econometrica 84(6):1983–2027

Manski CF (2003) Partial identification of probability distributions. Springer Science & Business Media

Manski CF (2004) Measuring expectations. Econometrica 72(5):1329–1376

McFadden D, et al (1973) Conditional logit analysis of qualitative choice behavior

Nyarko Y, Schotter A (2002) An experimental study of belief learning using elicited beliefs. Econometrica 70(3):971–1005

Ramsey FP (1931) Truth and probability. The foundations of mathematics and other logical essays pp 156–198

Savage LJ (1971) Elicitation of personal probabilities and expectations. Journal of the American Statistical Association 66(336):783–801

Savage LJ (1972) The foundations of statistics. Courier Corporation