Causality and Causal Misperception in Dynamic Games

Sungmin Park*

October 31, 2024

Click here for the latest version
https://ssrn.com/abstract=4852603

Abstract

I develop a game-theoretic solution concept in which agents have misperceptions about causality. In an observation-consistent equilibrium (OE), each player best responds to a belief about Nature or other players' actions consistent with observed outcomes. A maximum-entropy observation-consistent equilibrium (MOE) is an OE in which each player's belief maximizes the Shannon entropy among all observation-consistent beliefs. Every finite extensive-form game with perfect recall has a MOE. The equilibrium captures common causal misperceptions of correlation neglect, omitted-variable bias, and simultaneity bias. When players observe outcomes perfectly, OE and MOE are equivalent to self-confirming equilibrium and perfect Bayesian equilibrium.

Keywords: Causality, Causal misperception, Bounded rationality, Dynamic games, Extensive-form games, Principle of maximum entropy, Self-confirming equilibrium, Analogy-based expectation equilibrium, Markov perfect equilibrium, Game theory

JEL Classification: D83, C73

*Department of Economics, The Ohio State University. 1945 North High Street, Columbus, Ohio, United States 43210. Email: park.2881@buckeyemail.osu.edu.

I am grateful to Yaron Azrieli, PJ Healy, and Meta Brown for their guidance and frequent feedback throughout this paper's development. The work benefited from conversations with Alistair Barton, Raphael Boleslavsky, Davide Bordoli, Katie Coffman, Robert Daniel, Amanda Friedenberg, Wade Hann-Caruthers, Changkuk Im, Irfan Khan, Suha Kim, Hyeounggyun Ko, Giacomo Lanzani, Changyeop Lee, Adnan Mahmood, Kirby Nielson, Marieke Pahlke, Hyoeun Park, Jim Peck, Asha Sadanand, Peter Schwardmann, Ran Spiegler, Jason Tayawa, Daisy Wang, Han Wang, Matthew Weinberg, Michael Weisbach, Huanxing Yang, and Zexin Ye. I am thankful for the comments by the participants of the OSU Theory/Experimental Reading Group, the Midwest Economic Theory Conference of Spring 2024, the Conference on Mechanism and Institution Design 2024, and the 35th Stony Brook Game Theory Conference.

1 Introduction

Because people have limited observation of reality, they develop varying perceptions of causality—how one's actions affect outcomes. For example, until the mid-twentieth century, many doubted that smoking tobacco caused lung cancer, as there had not been sufficient clinical observation and animal experiments to support the link (Proctor, 2012). For less dated examples, college students tend to have rosy beliefs about how their degrees will secure high-paying jobs (Wiswall and Zafar, 2021). During the 2020 George Floyd protests, some suggested that increased police presence contributed to a rise in violent crimes, whereas others believed the opposite (Bump, 2020). In today's households, parents and children have varying perceptions about the effects of social media on mental health. In lab experiments, subjects interpret the same data differently depending on causal narratives, which leads them to recommend different policies (Kendall and Charles, 2022).

Motivated by the variety of causal perceptions that arise from limited observation, I develop a unified game-theoretic framework to represent causal misperception. My solution concept is *observation-consistent equilibrium* (OE), in which agents best respond to possibly incorrect beliefs about Nature or others' actions while those beliefs are observationally indistinguishable from the actual probabilities. In other words, each player sustains a belief that generates the same observable outcomes as the actual probability distribution. Because there are typically many observation-consistent beliefs, however, I further require that each player's belief maximizes the Shannon entropy among all observation-consistent beliefs: An OE satisfying this additional condition is called a *maximum-entropy observation-consistent equilibrium* (MOE). In essence, when faced with multiple beliefs that are consistent with observation, agents choose beliefs that assume the least information beyond what they observe. For example, when people see no evidence of a causal link between smoking and cancer, they assume none. Given a profile of strategies, this requirement provides a specific point-prediction on agents' beliefs, following the Principle of maximum entropy originally developed in statistical physics (Jaynes, 1957a,b) and adopted in decision theory for agents with limited data access (Spiegler, 2017, 2021).

I show that this framework is widely applicable in strategic situations involving multiple agents over time: Every finite extensive-form game with perfect recall has a MOE (Theorem 1). In other words, there always exists a predicted pattern of how agents play a dynamic game, such that everyone is doing their best given their specific, yet potentially incorrect, beliefs about how others play. Because players' predicted beliefs are generally different from the actual play, MOE typically predicts a different outcome from standard solution concepts in which players are assumed to have accurate beliefs. For instance, when a worker receives a smaller bonus than expected, he may be uncertain whether the smaller bonus is due to the manager's decision or an external factor. In standard predictions, such as in perfect Bayesian equilibrium, the worker knows the manager's strategy and responds accordingly. In contrast, in a MOE prediction, without knowing manager's strategy, the worker assumes that both the manager and an external factor are equally likely causes, and chooses his behavior accordingly. Therefore, MOE serves as an alternative solution concept that accommodates causal misperception while still offering precise predictions.

In addition to providing narrow predictions of strategic behavior by misperceiving agents, maximum-entropy observation-consistent equilibrium captures common causal misperceptions such as (a) correlation neglect, (b) omitted-variable bias and (c) simultaneity bias. Correlation neglect, the most basic feature of the equilibrium, occurs when agents observe only the marginal probabiliy distributions of two variables and not their joint distribution. In this case, players mistakenly assume that the two variables are independent (Proposition 1). The intuition behind the result is straightforward: when people see no evidence of correlation, they assume none. That is the belief presuming the least information beyond what they observe.

A more intriguing form of causal misperception in MOE is omitted-variable bias. Consider a three-stage game where Nature assigns different states, players take actions, and Nature assigns consequences. For example, players are born with varying natural abilities (states), make different educational investments (actions), and face different employment outcomes (consequences). Suppose players observe correlations between states and actions, and between actions and consequences, but not the full relationship among all three. In a perfect Bayesian equilibrium, players accurately perceive the effects of their abilities and education on employment. In contrast, in a MOE, players infer the effects of their actions on outcomes without accounting for the effects of unobserved factors, leading to an omitted-variable bias (Proposition 2). Consequently, they underestimate or overestimate the true effect of education on employment and make suboptimal educational choices.

The last form of causal misperception I examine is simultaneity bias, also known as reverse causality bias. Consider a different three-stage game where Nature initially assigns a binary state that determines the order of play between players and Nature. In one state, players take actions and Nature assigns outcomes afterward. In the other state, Nature assigns outcomes and players take actions afterward. For example, local governments (players) may decide to increase police presence or not, which in turn affects violent crime rates (Nature). In other instances, the order is reversed, with violent crime rates prompting local governments to adjust police presence. Suppose players observe others' actions and outcomes, but not their states. In a perfect Bayesian equilibrium, players accurately perceive the effects of increased police presence on violent crime rates, understanding the simultaneous causal relationship. In contrast, in a MOE, they misinterpret the observed correlation between police and crime as the effect of police on crime.

The previous example constructs a finite-horizon game with simultaneity between Nature and players' actions. However, simultaneity emerges more naturally in infinite-horizon settings. In stochastic games, in particular, players' actions affect future states; those states, in turn, influence subsequent actions. For example, there may be a simultaneous causal relationship over multiple periods between one's depression (state) and social media use (action): one's depression can lead one to use social media more often, and using social media often may exacerbate one's depression. I extend MOE to stochastic games such as this example and demonstrate how the predictions of my solution concept differ from those of Markov perfect equilibrium. In a stochastic game between a parent and a child, I show that the child in a Markov perfect equilibrium correctly perceives the adverse effects of social media and limits his use voluntarily. In contrast, a child in a MOE indulges in social media every period, prompting his parent to put a strict screen time policy on the child.

By incorporating causal misperception into finite- and infinite-horizon dynamic games, my work bridges the gap between the behavioral theory of causal misperception and conventional game theory. The existing literature on causal misperceptions focuses on single-person decisions. In Spiegler (2016, 2020, 2022, 2023), a decision maker's beliefs are distorted by a directed acyclic graph (DAG) that represents an exogenously specified set of conditional independence assumptions between variables. Building on this framework, Eliaz and Spiegler (2020) and Eliaz, Galperti, and Spiegler (2022) have DAGs represent narratives about how the world works, and let agents choose a DAG to maximize their subjective expected utility. Spiegler (2017) and Spiegler (2021) show that when a decision maker has limited access to data, the causal misperception represented by a DAG coincides with the maximum-entropy belief about the underlying data-generating process. My work extends their maximum-entropy approach to settings where multiple strategic agents interact over time.

By adopting the maximum-entropy approach on observation-consistent beliefs in games, my solution concepts incorporate self-confirming equilibrium (Fudenberg and Levine, 1993; Fudenberg and Kreps, 1995), perfect Bayesian equilibrium (Kreps and Wilson, 1982; Fudenberg and Tirole, 1991), and Markov perfect equilibrium (Maskin and Tirole, 1988a, 2001) as special cases. Specifically, when agents have perfect observation over a game's terminal histories, OE coincides with self-confirming equilibrium, and MOE coincides with perfect Bayesian equilibrium (Proposition 3). For infinite-horizon games such as stochastic games, MOE coincides with Markov perfect equilibrium (Proposition 5).

Relatedly, when agents have imperfect observation over terminal histories, OE closely resembles conjectural equilibrium (Battigalli and Guaitoli, 1988, 1997; Rubinstein and Wolinsky, 1994; Azrieli, 2009). As noted by Fudenberg and Levine (1993), however, the broad generality of the conjectural or observation-consistent equilibrium makes these concepts difficult to characterize. Addressing this critique, MOE offers a sharp refinement on OE by selecting the belief that assumes the least information, in the spirit of Occam's razor. This refinement is similar in motivation to analogy-based expectation equilibrium (Jehiel, 2005; Jehiel and Koessler, 2008; Jehiel, 2022; Jehiel and Weber, 2024), which assumes that players believe others play the same strategies in analogous situations.

All in all, maximum-entropy observation-consistent equilibrium (MOE) is a viable alternative to a dominant paradigm in applications of dynamic games. In dynamic structural modeling and estimation, it is "ubiquitous" to assume *rational expectations*—that agents fully understand the data-generating processes of their environments—even though "this is a very strong assumption in many applications" (Aguirregabiria and Mira, 2010). Empirical applications commonly use perfect Bayesian equilibrium (PBE, for finite horizons) or Markov perfect equilibrium (MPE, for infinite horizons; Maskin and Tirole, 1988a,b, 2001), both of which assume that agents correctly perceive the world and respond accordingly. Relaxing this assumption is challenging because it requires estimating both the model environment and players' beliefs. In recent work, Aguirregabiria and Magesan (2020) address this challenge by assuming that each player's payoff depends on one's state and others' actions but not others' states, allowing them to identify players' beliefs. In contrast, MOE offers a different solution by providing a point-prediction on each player's belief based on the Principle of maximum entropy.

I organize the remaining sections as follows. Section 2 illustrates my solution concepts in a simple example of a smoker's decision problem. Section 3 presents my main result that every finite extensive-form game has a MOE. Section 4 shows that the solution concept captures common misperceptions of correlation neglect, omitted-variable bias, and simultaneity bias. Section 5 shows that OE and MOE coincide with self-confirming and perfect Bayesian equilibria, respectively, under perfect observation of a game's terminal histories. It also makes a comparison between MOE and analogy-based expectation equilibrium. Section 6 applies the solution concept to infinite-horizon games, contrasting it with Markov perfect equilibrium. Section 7 concludes. I include all omitted proofs in the Appendix.

2 Simplest example

To see my ideas in the simplest possible scenario, consider a problem involving one strategic player (Smoker) and Nature. First, the smoker chooses to smoke tobacco ($s = 1$) or not ($s = 0$). Conditional on the smoker's choice, Nature gives him cancer ($y = 1$) or not ($y = 0$). The objective probabilities of Nature's moves are exogenously given as $(\pi_0, \pi_1) \in (0, 1)^2$ where π_0 is the probability of cancer given $s = 0$ and π_1 is the probability of cancer given

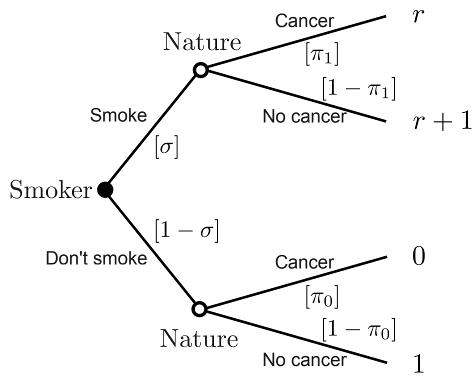


Figure 1: A smoker's decision problem

Note: Numbers in square brackets are the probabilities of the smoker's and Nature's moves. Numbers on terminal nodes are the smoker's payoffs.

$s = 0$. Let $\pi_1 > \pi_0$, meaning that smoking increases the probability of cancer, or smoking *causes* cancer. I use this working definition of causality, while deferring the precise definition for extensive-form games to Section 4.4. The smoker's payoff is $u(s, y) = 1 - y + rs$, where $r > 0$ represents the reward from smoking. Suppose $\pi_1 - \pi_0 > r$; that is, the negative impact of smoking on health outweighs the reward from smoking. Figure 1 illustrates this decision problem.

A *strategy* of the smoker is the probability $\sigma \in [0, 1]$ of smoking. A smoker's *belief* is $\beta = (\beta_0, \beta_1)$ where β_s is a subjective probability that Nature gives cancer conditional on s.

Definition 1. Given strategy $\sigma \in [0, 1]$, a belief $\beta \in [0, 1]^2$ is *observation-consistent* if

$$(1 - \sigma)\beta_0 + \sigma\beta_1 = (1 - \sigma)\pi_0 + \sigma\pi_1.$$

This condition requires that the Smoker's ex-ante subjective probability of getting cancer (left-hand side) is equal to the ex-ante objective probability (right-hand side). In other words, the smoker's belief is observation-consistent if it results in the same overall probability of cancer as would occur according to Nature's actual probabilities. Alternatively, one could view this definition as consisting of two separate conditions:

1. matching the observed cancer frequency: $(1 - \sigma)\beta_0 + \sigma\beta_1 = (1 - \bar{\sigma})\pi_0 + \bar{\sigma}\pi_1$, and

2. consistency of smoking rates: $\bar{\sigma} = \sigma$,

where $\bar{\sigma}$ is the frequency of cancer observed in a population of smokers facing the same decision problem. The first condition states that the subjective belief about the cancer rate must match the population cancer rate. The second condition requires that the individual's smoking rate is consistent with the population smoking rate.

There are infinitely many beliefs that are observation-consistent, as they consist of two unknowns that satisfy the above single equation. For this reason, I focus on the observation-consistent belief that maximizes the Shannon entropy: the belief with the least information content among observation-consistent ones. Given strategy σ and belief β, let $\mathbf{p}(\sigma, \beta)$ denote the vector of subjective probabilities over this decision problem's terminal histories, which are (smoke, cancer), (smoke, no cancer), (don't smoke, cancer), and (don't smoke, no cancer):

$$\mathbf{p}(\sigma, \beta) = [\sigma\beta_1,\ \sigma(1-\beta_1),\ (1-\sigma)\beta_0,\ (1-\sigma)(1-\beta_0)]^T.$$

Let $G(\cdot)$ be the Shannon entropy function: for any probability vector p with elements p_1, p_2, and so on,

$$G(p) = \sum_{p_k > 0} -p_k \log p_k.$$

Definition 2. Given a strategy $\sigma \in (0,1)$, an observation-consistent belief $\beta \in [0,1]^2$ *maximizes the entropy* if

$$\beta \in \operatorname*{argmax}_{\beta' \in [0,1]^2} G(\mathbf{p}(\sigma, \beta')) \quad \text{subject to} \quad (1-\sigma)\beta_0 + \sigma\beta_1 = (1-\sigma)\pi_0 + \sigma\pi_1.$$

In this case, β is the *maximum-entropy belief* given the strategy σ.

Figure 2 illustrates this definition. The four corners of the 3-simplex are the terminal histories of the decision problem, and each point in the simplex represents a probability distribution over the terminal histories. The smoker's own strategy σ as well as observational consistency places restrictions on the beliefs. Among the observation-consistent beliefs represented by the line segment, the maximum-entropy belief is the one closest to the center of the simplex. Note that the maximum-entropy belief is defined for a totally mixed strategy $\sigma \in (0,1)$ to guarantee a unique solution. If $\sigma = 0$ or $\sigma = 1$, different values of β_1 or β_0 result in the same entropy.

The unique observation-consistent belief maximizing the entropy is that smoking is unrelated with cancer.

Claim 1. *For every $\sigma \in (0,1)$, the maximum-entropy belief β satisfies*

$$\beta_0 = \beta_1 = (1-\sigma)\pi_0 + \sigma\pi_1.$$

That is, when a smoker observes a population of smokers like himself smoking with frequency σ and getting cancer with average frequency $(1-\sigma)\pi_0 + \sigma\pi_1$, his belief with minimal assumptions consistent with this observation is that smoking has no effect on cancer.

Because the set of observation-consistent beliefs and the maximum-entropy beliefs depend on the smoking frequency σ, it is natural to ask which smoking frequency σ and belief

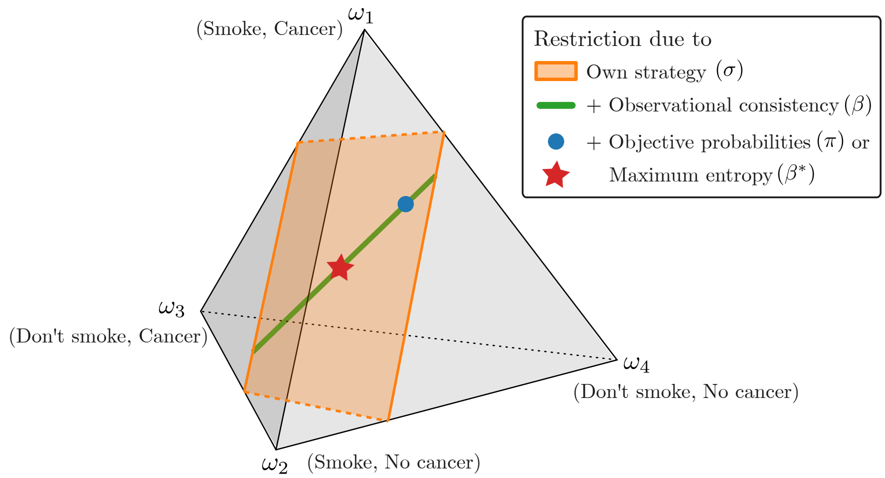


Figure 2: Observational consistency and maximum entropy

β form an equilibrium. Let $U(\sigma, \beta)$ denote the smoker's subjective expected payoff given the smoker's strategy σ and belief β.

Definition 3. An *observation-consistent equilibrium* (OE) of the smoker's problem is a strategy-belief pair (σ, β) such that

1. Given the belief β, the smoker's strategy σ is (subjectively) rational:

$$\sigma \in \operatorname*{argmax}_{\sigma' \in [0,1]} U(\sigma', \beta), \quad \text{and}$$

2. Given his strategy σ, the smoker's belief β is observation-consistent.

Given that there are many observation-consistent beliefs for any strategy σ, it is not surprising that there are many observation-consistent equilibria. In fact, for every strategy $\sigma \in [0, 1]$, there is at least one observation-consistent belief that rationalize that strategy.

Claim 2. *A strategy-belief pair (σ, β) is an OE if and only if one of the following is true:*

(a) $\sigma = 0$, $\beta_0 = \pi_0$, *and* $\beta_1 \geq \pi_0 + r$,

(b) $\sigma = 1$, $\beta_1 = \pi_1$, *and* $\beta_0 \geq \pi_1 - r$, *or*

(c) $\sigma \in (0, 1)$, $\beta_0 = (1 - \sigma)\pi_0 + \sigma(\pi_1 - r)$, *and* $\beta_1 = (1 - \sigma)(\pi_0 + r) + \sigma\pi_1$.

The intuition behind this large set of equilibria is the following. When the agent never smokes ($\sigma = 0$), he correctly figures that the probability of cancer conditional on not smoking is π_0. However, because he never smokes, any belief β_1 on cancer conditional on smoking is observationally consistent with the actual conditional probability π_1. Therefore, as long as he has a sufficiently high belief β_1, the decision to never smoke is a best response. Similarly, when the agent always smokes ($\sigma = 1$), any belief β_0 on cancer conditional on not smoking is observationally consistent. Therefore, as long as he has a sufficiently high belief β_0, the decision to always smoke is a best response. Finally, when the agent mixes between the two actions, the belief β is chosen to satisfy $\beta_1 - \beta_0 = r$, the condition that makes the smoker is exactly indifferent between smoking and not smoking.

Because OE is too permissive, I focus on the following refinement.

Definition 4 (MOE)**.** An observation-consistent equilibrium (σ, β) is a *maximum-entropy observation-consistent equilibrium (MOE)* if there exists a sequence $\{\sigma^k, \beta^k\}_{k=1}^{\infty}$ converging to (σ, β) such that each σ^k is totally mixed and β^k maximizes the entropy given σ^k.

Importantly, MOE is defined as an OE whose belief is a limit point of maximum-entropy beliefs, as the latter is defined only for totally mixed strategies.

Compared to OE, MOE provides a much sharper prediction.

Claim 3. *A strategy-belief pair (σ, β) is a MOE if and only if $\sigma = 1$ and $\beta_0 = \beta_1 = \pi_1$.*

That is, when the agent smokes with 100% frequency, he correctly believes that the probability of getting cancer conditional on smoking is π_1, but he incorrectly believes that the probability of getting cancer conditional on not smoking is also π_1. Such is the belief that assumes the least information beyond what he observes. He believes that smoking is unrelated with cancer ($\beta_0 = \beta_1$, Claim 1), and thus continues to smoke. No other strategy can be supported as a MOE, as smoking with less than 100% frequency requires that the effect of smoking on cancer is no smaller than the reward ($\beta_1 - \beta_0 \geq r$) for the strategy to be rational.

3 Equilibria in finite extensive-form games

Having seen the simplest example of the Smoker's problem, I now generalize the observation-consistent equilibrium (OE) and its refinement, maximum-entropy observation-consistent equilibrium (MOE) to a general class of games that involve multiple players' actions over time under imperfect information. With multiple players, each player's observation-consistent belief is a belief about other players' (including Nature's) play based on his observation. MOE is a strategy profile and a belief profile such that each player best responds to his belief maximizing the entropy. I show that every finite extensive-form game with perfect recall and observational constraint has a MOE.

3.1 Main result: Existence of maximum-entropy observation-consistent equilibrium (MOE)

Consider a finite extensive-form game with perfect recall $\Gamma = (N, H, \iota, \pi, \mathcal{I}, u)$ as in Osborne and Rubinstein (1994), which consists of the following elements:

1. N, the set of players,
2. H, the set of histories of actions,
3. ι, the mapping of non-terminal histories to players or Nature,
4. π, a probability distribution with full support over Nature's moves,
5. $\mathcal{I}$, the collection of information sets, and
6. u, the payoff function.

Let $\Omega \subset H$ denote the set of terminal histories. A *finite extensive-form game with perfect recall and observational constraint* includes all of the above and the following additional component:

7. C, an *observational structure*, a matrix with $|\Omega|$ columns.

Defined as a matrix, the observational structure C is a linear mapping that assigns each probability distribution $q \in \Delta(\Omega)$ over terminal histories to the product Cq, a real vector. The vector Cq represents the outcomes observed by players. For instance, suppose there are four terminal histories: $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$. Examples of observational structures are

$$C = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & \cdot & 1 & \cdot \\ \cdot & 1 & \cdot & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & 1 \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix},$$

where dots represent zeros for convenience. The first structure means that nothing is observable to players besides the fact that all probabilities sum to 1. The second structure is the observational structure in Smoker's example (Section 2). It means that players only observe the marginal probabilities of cancer. The third structure means that players perfectly observe the probability distribution over terminal histories. The fourth structure means that players observe only the expected value of the random variable that assigns each terminal history ω_k to the number k.

Let a finite extensive-form game with perfect recall and observational constraint be given as (Γ, C). For every non-terminal history $h \in H \backslash \Omega$, let $\Omega(h)$ denote the set of terminal histories that are successors of h. For every information set I, let $A(I)$ denote the set of actions availabe at the information set I. Let $\mathcal{I}_i$ denote the collection of information sets of player i. Let $\mathcal{I}_{-i}$ denote the collection of information sets of all players other than i and Nature.

A *strategy* (or *behavioral strategy*) of player i is σ_i that specifies the probability $\sigma_i(a|I)$ of choosing action $a \in A(I)$ at every information set $I \in \mathcal{I}_i$ of player i. Let $\mathcal{S}_i$ denote the set of player i's possible strategies. A strategy σ_i is *totally mixed* if it specifies a positive probability on all actions available at every information set $I \in \mathcal{I}_i$. A *strategy profile* is $\sigma \in \mathcal{S} = \prod_i \mathcal{S}_i$. A profile of strategies of all players except i is $\sigma_{-i} \in \mathcal{S}_{-i} = \prod_{j \neq i} \mathcal{S}_j$.

A *belief* of player i is a function β_i that specifies player i's subjective probability $\beta_i(a|I)$ of action $a \in A(I)$ given $I \in \mathcal{I}_{-i}$. The set of player i's feasible beliefs is $\mathcal{B}_i = \mathcal{S}_{-i} \times \mathcal{P}$, where $\mathcal{P}$ is the set of possible probability distributions of Nature's moves with full support. A *belief profile* is an element $\beta \in \mathcal{B} = \prod_i \mathcal{B}_i$.

A *posterior function* of player i is a function μ_i that specifies player i's subjective probability $\mu_i(h|I)$ of history $h \in I$ conditional on reaching information set $I \in \mathcal{I}_i$. The set of player i's possible posterior functions is $\mathcal{M}_i$. A *posterior function profile* is an element $\mu \in \mathcal{M} = \prod_i \mathcal{M}_i$. An *assessment* is a triple (σ, β, μ) of a strategy profile, a belief profile, and a posterior function profile.

For every strategy-belief pair (σ_i, β_i), let $p(\sigma_i, \beta_i)(h_L|h_K)$ denote player i's subjective probability of reaching a history $h_L = (a_1, a_2, \ldots, a_L)$ conditional on reaching a preceding history $h_K = (a_1, a_2, \ldots, a_K)$. That is,

$$p(\sigma_i, \beta_i)(h_L|h_K) = \prod_{L \leq k \leq K-1,\ \iota(h_k)=i} \sigma_i(a_{k+1}|h_k) \prod_{L \leq k \leq K-1,\ \iota(h_k)\neq i} \beta_i(a_{k+1}|h_k).$$

Similarly, let $p(\sigma_i, \beta_i)(h) = p(\sigma_i, \beta_i)(h|h_0)$ denote player i's subjective unconditional probability on the history h, given the root h_0 of all histories. Let $\mathbf{p}(\sigma_i, \beta_i)$ denote the vector of unconditional probabilities over the set of terminal histories Ω. That is, each element of the vector $\mathbf{p}(\sigma_i, \beta_i)$ is $p(\sigma_i, \beta_i)(\omega)$ where $\omega \in \Omega$.

Given a strategy-belief pair (σ_i, β_i), a posterior function μ_i is *Bayes-consistent* if for every $I \in \mathcal{I}_i$ with $\sum_{h \in I} p(\sigma_i, \beta_i)(h) > 0$,

$$\mu_i(h|I) = \frac{p(\sigma_i, \beta_i)(h)}{\sum_{h \in I} p(\sigma_i, \beta_i)(h)}.$$

In other words, the posterior function is Bayes-consistent if it satisfes the Bayes rule on the equilibrium paths of play.

For every information set $I \in \mathcal{I}_i$ of player i, let $U_i(\sigma_i, \beta_i, \mu_i|I)$ denote player i's (subjective) expected utility given his strategy σ_i, belief β_i, and posterior function μ_i. That is,

$$U_i(\sigma_i, \beta_i, \mu_i|I) = \sum_{h \in I} \sum_{\omega \in \Omega(h)} u_i(\omega) \cdot p(\sigma_i, \beta_i)(\omega|h) \cdot \mu_i(h|I).$$

Given a belief β_i and a posterior function μ_i, a strategy σ_i of player i is (subjectively) *sequentially rational* if for every information set $I \in \mathcal{I}_i$,

$$\sigma_i \in \underset{\sigma'_i \in \mathcal{S}_i}{\operatorname{argmax}} U_i(\sigma'_i, \beta_i, \mu_i|I).$$

Definition 5. Let a strategy profile σ be given. A belief β_i of player i is *observation-consistent* if

$$C\mathbf{p}(\sigma_i, \beta_i) = C\mathbf{p}(\sigma_i, (\sigma_{-i}, \pi)).$$

That is, player i's belief is observation-consistent if the subjective probability distribution generated by how he plays and what he thinks how others play is consistent with the objective probability distribution generated by how everyone (including Nature) actually plays, for observable outcomes defined by the linear map C.

Definition 6. Let a totally mixed strategy profile σ be given. An observation-consistent belief *maximizes the entropy* if

$$\begin{aligned} \beta_i \in \underset{\beta'_i \in \mathcal{B}_i}{\operatorname{argmax}} & \, G(\mathbf{p}(\sigma_i, \beta'_i)) \\ \text{subject to} & \quad C\mathbf{p}(\sigma_i, \beta'_i) = C\mathbf{p}(\sigma_i, (\sigma_{-i}, \pi)), \end{aligned}$$

where $G(\cdot)$ is the Shannon entropy function. In this case, the β_i is player i's *maximum-entropy belief* given the strategy profile σ.

We are ready to define observation-consistent equilibrium (OE) and maximum-entropy observation-consistent equilibrium (MOE).

Definition 7 (Equilibrium)**.** An assessment (σ, β, μ) is an *observation-consistent equilibrium (OE)* if for every player i,

1. the strategy σ_i is sequentially rational given (β_i, μ_i),

2. the belief β_i is observation-consistent given the strategy profile σ, and

3. the posterior function μ_i is Bayes-consistent given (σ_i, β_i).

An OE (σ, β, μ) is a *maximum-entropy observation-consistent equilibrium (MOE)* if there exists a sequence $\{\sigma^k, \beta^k\}_{k=1}^{\infty}$ converging to (σ, β) where each σ^k is a totally mixed strategy profile and each β_i^k maximizes the entropy given σ^k.

In other words, in any OE, every player best responds to one's belief about Nature's and other players' play at every point in the game, and that belief is observationally equivalent to how Nature and other players actually play. In a MOE, this belief is the minimally informative one based on the partially observed probability distribution over the game's terminal histories. Thus, MOE places much more stringent requirement on the players' beliefs than an OE. A special case is when the set of observation-consistent belief is a singleton, such as when the observational structure C is the identity matrix. In this case, OE coincides with self-confirming equilibrium and MOE coincides with perfect Bayesian equilibrium. I defer the detailed discussion to Section 5. However, if there are multiple observation-consistent beliefs, the belief that maximizes the entropy generally differs from the actual play. Since players' beliefs in a MOE generally diverges from the actual play, their best response strategy profiles are in general different from those predicted by perfect Bayesian equilibrium.

We know that an OE exists because a sequential equilibrium exists (Kreps and Wilson, 1982). Every sequential equilibrium is an OE in which every player best responds to the correct belief about how others play. However, it is not immediately obvious that a MOE exists, because of the sharp requirement on everyone's belief. We can rest assured that it does.

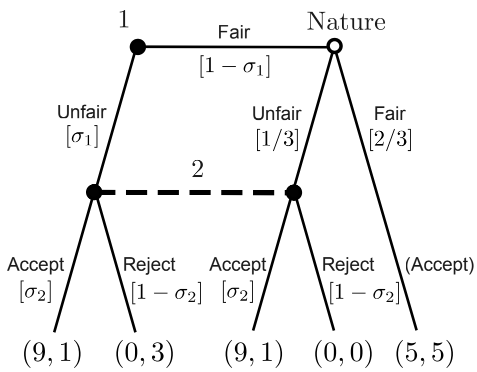
Theorem 1. *Every finite extensive-form game with perfect recall and observational constraint has a maximum-entropy observation-consistent equilibrium (MOE).*

Put differently, we always have an equilibrium in which everyone is best responding to what they think how others play, based on the least informed belief that is observationally equivalent to how others really play.

The proof (in Appendix A) utilizes ϵ-constrained strategy profiles: totally mixed strategies profiles in which each player's available action at every information set is played with at least ϵ probabilities. Because the mappings from ϵ-constrained strategy profiles to the maximum-entropy belief profiles and to the posterior function profiles are well-defined and continuous, the best response correspondence from the set of ϵ-constrained strategy profiles to itself is well-behaved: nonempty-valued, convex-valued, and upper-hemicontinuous. This implies that the correspondence has a fixed point, which approaches a MOE as ϵ approaches zero. This proof structure mirrors the extensive-form-based proof of the existence of sequential equilibrium by Chakrabarti and Topolyan (2016).

3.2 Example: Manager-Worker game

To see how a maximum-entropy observation-consistent equilibrium (MOE) works compared to perfect Bayesian equilibrium, consider an ultimatum-game-like situation depicted in



$$C = \begin{bmatrix} 1 & \cdot & 1 & \cdot & \cdot \\ \cdot & 1 & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 \end{bmatrix}.$$

Figure 3: Manager-Worker game

Note: Numbers in square brackets are the probabilities of actions. Numbers in parentheses are the payoffs of Player 1 (Manager) and Player 2 (Worker). The matrix C is the observational structure whose five columns correspond to the five outcomes of the game.

Figure 3. There are Manager (Player 1, *she*), Worker (Player 2, *he*), and Nature. Manager moves first, deciding between offering a fair or unfair bonus to Worker. If Manager chooses a fair offer, Nature keeps the offer fair with probability $2/3$ or changes the offer to an unfair one with probability $1/3$. If Nature keeps the offer fair, Manager and Worker receive payoffs of 5 each. If either Manager chooses an unfair offer or Nature makes the offer unfair, Worker may either accept or reject the unfair offer. If Worker accepts the unfair offer, Manager gets away with a payoff of 9 while Worker gets 1. If Worker rejects the unfair offer, both receive zero monetary payoff. However, if Worker rejects the unfair offer chosen by Manager, he gets a nonpecuniary payoff of 3—interpreted as a thrill he receives by taking a revenge on Manager. Although the explicit inclusion of this nonpecuniary payoff is unconventional, it makes the game interesting by making Worker's subjectively rational behavior to depend on his belief about the cause of an unfair offer.

The game's observational structure, represented by matrix C in Figure 3, implies that the Worker only observes the overall probability distribution of fair and unfair offers, without knowing the specific probabilities of the Manager's and Nature's individual moves. The first row of matrix C ensures that the sum of the subjective and objective probabilities of reaching the first and third terminal histories are equal. Similarly, the second row equates the subjective and objective probabilities of reaching the second and fourth terminal histories. The final row implies that all players know the probability of reaching the last terminal history. The interpretation is that each Manager-Worker pair sees a population of similar Manager-Worker relationships and only the coarse outcomes. As a result, any belief about the Manager's and Nature's moves that leads to the same overall distribution of fair and unfair offers is observationally equivalent to the true distribution of their moves.

Let us examine the standard prediction in this game. Let σ_1 denote Manager's probability of choosing an unfair offer. Let σ_2 denote Worker's probability of accepting when given an unfair offer. Let μ denote Worker's posterior that Manager chose the unfair offer conditional on receiving an unfair offer. A *perfect Bayesian equilibrium (PBE)* of the Manager-Worker game is a tuple $((\sigma_1, \sigma_2), \mu)$ such that σ_1 is a best response to σ_2, σ_2 is a best response to σ_1 given μ, and μ satisfies Bayes rule:

$$\mu = \frac{\sigma_1}{\sigma_1 + \frac{1}{3}(1 - \sigma_1)}.$$

In a PBE, Manager treats Worker unfairly with significant frequency.

Claim 4. *The unique perfect Bayesian equilibrium of the Manager-Worker game is*

$$((\sigma_1, \sigma_2), \mu) = ((1/7,\ 5/9)\,,\ 1/3)\,.$$

That is, Manager chooses an unfair offer one out of seven times, while Worker—believing ex-post that an unfair offer is due to Manager's choice with probability 1/3—accepts an unfair offer five out of nine times. Mixing between fair and unfair offers is a best response for Manager because she is indifferent between the two: Given that Worker accepts the unfair offer with probability 5/9, both options give her expected payoffs of 5. Similarly, mixing between accepting and rejecting an unfair offer is a best response for Worker because he is indifferent between the two: Given that Manager chooses an unfair offer one out of seven times, Worker knows that Manager is the cause of the unfair offer one out of three times, making both accepting and rejecting result in expected payoffs of 1.

There is no pure strategy PBE, however. If Manager always chooses an unfair offer ($\sigma_1 = 1$), Worker knows that any unfair offer is entirely due to Manager's choice ($\mu = 1$), and thus chooooses to reject an unfair offer. Given this strategy by Worker, choosing an unfair offer is not a best response for Manager. If Manager always chooses a fair offer ($\sigma_1 = 0$), Worker knows that any unfair offer is never Manger's fault ($\mu = 0$), and thus chooses to accept any unfair offer. Given this strategy by Worker, Manager, choosing a fair offer is not a best response for Manager.

A key aspect we see in the standard solution concept is that players have the correct causal perceptions in equilibrium. In this Manager-Worker game, Worker knows the objective posterior probability of the cause behind an unfair offer. If Manager always treats him unfairly, Worker knows that Manager is the cause of the unfair treatment rather than Nature. If Manager always treats him fairly, Worker knows that any unfair treatment is caused by Nature rather than Manager. If Manager mixes, Worker forms the correct posterior proability of her being the cause.

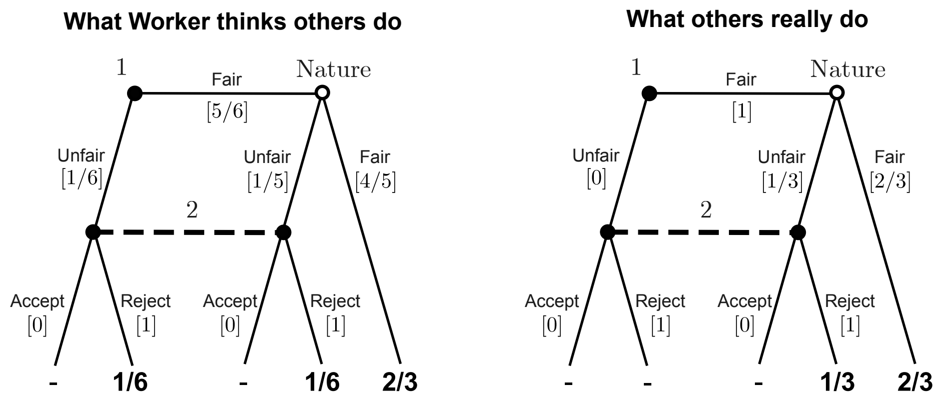


Figure 4: Maximum-entropy observation-consistent equilibrium (MOE) of Manager-Worker game

Note: Numbers in square brackets are the subjective (left) and objective (right) probabilities of actions. Numbers in bold at each terminal node are subjective (left) and objective (right) probabilities of reaching that node.

Let us see how this standard prediction is different from that with MOE. Recall that σ_1 is Manager's probability of choosing an unfair offer, σ_2 is Worker's probability of accepting an unfair offer, and μ is Worker's posterior on Manager being the cause of an unfair offer. Let $\beta_1 = (\beta_1^N, \beta_1^W)$ denote Manager's belief, where β_1^N is her belief on Nature choosing an unfair offer and β_1^W is her belief on Worker accepting an unfair offer. Let $\beta_2 = (\beta_2^N, \beta_2^M)$ denote Worker's belief, where β_2^N is his belief on Nature choosing the unfair offer and β_2^M is his belief on Manager choosing the unfair offer.

The MOE prediction is markedly different from the standard prediction.

Claim 5. *The unique MOE of the Manager-Worker game is* $((\sigma_1, \sigma_2), (\beta_1, \beta_2), \mu)$ *where*

$$(\sigma_1, \sigma_2) = (0, 0), \quad (\beta_1^N, \beta_1^W) = (1/3, 0), \quad (\beta_2^N, \beta_2^M) = (1/5, 1/6), \quad \textit{and} \quad \mu = 0.5.$$

That is, Manager has the correct belief about Nature's and Worker's play, and always chooses the fair offer. Meanwhile, upon seeing an unfair offer, Worker believes that it is equally likely caused by Manager or Nature, and always rejects an unfair offer. Such posterior belief is based on his observationally consistent yet incorrect belief that Manager chooses an unfair offer one out of six times while Nature chooses an unfair offer one out of five times. Choosing the fair offer is Manager's best response given that Worker always rejects an unfair offer. Rejecting any unfair offer is Worker's best response given his incorrect belief. Figure 4 illustrates what Worker thinks others (Manager and Nature) do and what they really do in this MOE.

Although I include the proof of this claim in Appendix A, it is not difficult to reason through why this profile of strategies and beliefs is the only MOE. First, any maximum-entropy belief (β_1^N, β_1^W) of Manager is a correct belief about Nature's and Worker's plays, as the observational constraint is sufficient to pin down Worker's and Nature's strategies precisely. Specifically, because Manager knows her own strategy σ_1 and observes the probability of the Fair-Fair outcome, he correctly deduces the probability of Nature choosing Unfair. Second, any maximum-entropy belief (β_2^N, β_2^M) of Worker results in the uniform probability distribution over the two histories in his information set, whether Manager or Nature is behind an unfair offer. That is, $\beta_2^M = (1 - \beta_2^M)\beta_2^N$. The reason is that conditional on reaching his information set, he does not observe anything beyond his own strategy. He observes the sum of the probabilities over the terminal histories Unfair-Reject and Fair-Unfair-Reject but cannot distinguish between the two. Similarly, he observes the sum of the probabilities over the terminal histories Unfair-Accept and Fair-Unfair-Accept but cannot distinguish between the two. Therefore, he simply believes that Manager and Nature are equally likely causes behind an unfair offer. Lastly, because Worker receiving an unfair offer thinks that Manager and Nature are equally likely causes, his best response is to reject any unfair offer. Knowing that Worker would reject any unfair offer, Manager's best response is to always chooses a fair offer. These steps show that the aforementioned profile of strategies and beliefs is the only MOE.

The crucial reason behind the sharp differences in the PBE and MOE predictions is that the latter allows causal misperception. In the unique MOE, even as Manager always chooses a fair offer, there is a chance that Nature changes it to an unfair offer. Not having any ex-ante evidence about Manager's or Nature's behavior, Worker assumes that any unfair offer could have come from Manager or Nature with equal likelihood, driving him to reject unfair offers. Such causal misperception is impossible with PBE. Worker knows how Manager plays, so his posterior belief on the cause of the unfair offer is always correct. Always choosing the fair offer cannot be a PBE strategy for Manager, because Worker correctly perceives that any unfair offer if entirely due to Nature and thus will always accept.

This example is interesting because the observational constraint and the resulting misperception leads to more cooperative outcomes. This result aligns with the experimental findings that adding noise by Nature and enforcing an observational constraint about opponents' actions in repeated Prisoner's Dilemma games increases cooperation (Aoyagi, Bhaskar, and Fréchette, 2019). In terms of objective ex-ante expected payoffs, however, both of them are worse off due to the observational constraint. In the PBE, the Worker's expected payoff is around 3.86, while the Manager's is 5. In the MOE, both Worker and Manager earn an objective expected payoff of 3.33, as Worker always rejects an unfair offer.

3.3 Discussion: Interpretation of equilibrium

As in the smoker's decision problem (Section 2), we may interpret the MOE of the Manager-Worker game as a symmetric equilibrium for a population of players playing the same extensive-form game. Specifically, many manager-worker pairs in the corporate world may face the same strategic situation of splitting a profit between a manager and a worker. Given the observational constraint, workers do not know the relative frequency of managers intentionally choosing an unfair bonus. Without this knowledge, they best respond to beliefs consistent with the observed behavior in the population. This interpretation takes observational consistency literally.

This literal interpretation of MOE has testable implications for lab experiments. For example, consider the following experiment. Initially, subjects are randomly assigned into four groups: treated managers, treated workers, control managers, and control workers. This assignment is permanent throughout the experiment. In each round, each treated manager is randomly paired with one treated worker, and each control manager is randomly paired with one control worker. Each pair plays the Manager-Worker game, knowing the whole extensive form except Nature's probabilities of moves (π). At the end of each round, treated pairs learn the coarse outcome (Fair-Accept, Unfair-Accept, and Unfair-Reject) and the distribution of those coarse outcomes among all treated pairs. In contrast, control pairs learn the precise outcome and the distribution of the exact outcomes among all control pairs. In other words, the treated pairs are given the coarse observational structure C as in the Manager-Worker game, whereas the control pairs are given the identity matrix as the observational structure. Every subject is paid according to the random problem selection (RPS) mechanism (Azrieli, Chambers, and Healy, 2018). My solution concept predicts that the treated pairs would behave as the MOE described in Claim 5, whereas the control pairs would behave as the PBE described in Claim 4. Thus, this experiment would test people's behavior facing uncertainty about others' intentions in one-shot interactions, compared to Aoyagi, Bhaskar, and Fréchette (2019).

Alternatively, rather than literally interpreting the observational constraint, one could interpret it as a modeling device to represent strategic uncertainty. When the observational structure C is the identity matrix, it means that there is no uncertainty about opponents' strategies. When the observational constraint is the one considered in the Manager-Worker game, the worker has little idea about the manager's strategy. I explore this interpretation further in Section 5 in the context of centipede games.

4 Maximum entropy and causal misperception

As demonstrated in the Smoker's example from Section 2, observation-consistent equilibrium (OE) tends to be quite permissive, admitting a wide range of predictions due to its flexibility in accommodating different beliefs. To address this issue, I focus on maximum-entropy observation-consistent equilibrium (MOE), which refines OE by selecting the belief that maximizes the Shannon entropy among all observation-consistent ones, following the Principle of maximum entropy. In the smoker's example and the Manager-Worker game of Section 3, this refinement significantly narrows down the range of predicted outcomes. This principle ensures that players hold the beliefs that assume the least information consistent with their observations, resulting in sharper predictions about strategic behavior.

The Principle of maximum entropy has been widely adopted in fields such as statistical physics, information theory, and machine learning, where it is justified by strong axiomatic foundations (Shore and Johnson, 1980; Csiszar, 1991). In the decision theory of causal misperception by Spiegler (2017, 2020, 2021), the principle aligns well with the use of Directed Acyclic Graphs (DAGs) to represent agents' beliefs about causal relationships. In my framework, I extend this justification by highlighting the practical value of MOE in capturing common real-world misperceptions. Specifically, it provides a structured way to model biases such as correlation neglect, where individuals fail to account for correlation among variables, omitted-variable bias, where individuals fail to account for unobserved factors, and simultaneity bias, where individuals fail to account for the fact that there are simultaneous causal relationships.

4.1 Correlation neglect

Consider the following *correlated consequences game*, a two-stage extensive-form game with a set $N = \{1, 2, \dots, n\}$ of players. In the first stage, each player $i \in N$ simultaneously decides an action $x_i \in \mathcal{X}_i$, resulting in an action vector $x = (x_1, x_2, \dots, x_n) \in \mathcal{X} = \prod \mathcal{X}_i$. In the second stage, Nature chooses a consequence $y = (y_1, y_2) \in \mathcal{Y} = \mathcal{Y}_1 \times \mathcal{Y}_2$ with probability $\pi(y|x) > 0$ for all $(x, y) \in \mathcal{X} \times \mathcal{Y}$. The sets $\mathcal{X}$ and $\mathcal{Y}$ are finite. Let $u_i(x, y)$ be the payoff for player i. A *(behavioral) strategy* σ_i is a probability distribution over the set $\mathcal{X}_i$ of player i's actions. A *strategy profile* is a collection $\{\sigma_i\}_{i \in N}$.

Let the observational structure C be such that the players observe only the marginal probabilities of pairs (x, y_1) and (x, y_2), respectively. That is, a belief β_i is observation-consistent given a strategy profile σ if for all x, y_1, y_2,

$$\sum_{y_2'} p(\sigma_i, \beta_i)(x, y_1, y_2') = \sum_{y_2'} p(\sigma_i, (\sigma_{-i}, \pi))(x, y_1, y_2'), \text{ and}$$
$$\sum_{y_1'} p(\sigma_i, \beta_i)(x, y_1', y_2) = \sum_{y_1'} p(\sigma_i, (\sigma_{-i}, \pi))(x, y_1', y_2).$$

In other words, a belief β_i is observation-consistent given σ if for all x, y_1, y_2,

$$\beta_i(y_1|x)\sigma_i(x_i)\beta_i(x_{-i}) = \pi(y_1|x)\sigma_i(x_i)\sigma_{-i}(x_{-i}), \text{ and}$$
$$\beta_i(y_2|x)\sigma_i(x_i)\beta_i(x_{-i}) = \pi(y_2|x)\sigma_i(x_i)\sigma_{-i}(x_{-i}),$$

where $\beta_i(y_1|x) = \sum_{y_2} \beta_i(y_1, y_2|x)$ and $\beta_i(y_2|x) = \sum_{y_1} \beta_i(y_1, y_2|x)$ are the marginal subjective probabilities of y_1 and y_2, respectively, given an action profile x. The marginal objective probabilties $\pi(y_1|x)$ and $\pi(y_2|x)$ are defined analogously.

In a maximum-entropy observation-consistent equilibrium (MOE), all players of this game neglect the correlation between the two types of consequences.

Proposition 1. *An observation-consistent equilibrium (σ, β, μ) of the correlated consequences game is a MOE if and only if for every player i,*

$$\beta_i(x_{-i}) = \sigma_{-i}(x_{-i}) \quad \textit{for all } x_{-i} \in \mathcal{X}_{-i}, \textit{ and}$$
$$\beta_i(y_1, y_2|x) = \pi(y_1|x)\pi(y_2|x) \quad \textit{for all } x \in \mathcal{X} \textit{ and } (y_1, y_2) \in \mathcal{Y}.$$

This result characterizes MOE in terms of players' beliefs. The first equation says that the players have correct beliefs about other players' strategies. This result is not surprising, given that the observational structure fully reveals the probability distribution over all players' actions. The second equation is more interesting: it says that the players are correct about the marginal conditional probabilities over the two sets of consequences, but they wrongly believe that the two consequences are independent, conditional on players actions. The intuition behind this result is that the maximum-entropy criterion chooses the belief that assumes the least information beyond observed outcomes. When agents are given no evidence of correlation between two variables, they believe in no correlation.

As an illustrative example, consider the following decision problem by an investment bank. In the first stage, the investment bank decides whether or not they will issue mortgage-backed securities (MBS), by pooling together risky mortgage loans to households to transform them into financial assets that appear less risky. In the second stage, Nature decides whether (a) two households default on their mortgage loans, (b) only household A defaults, (c) only household B defaults, or (d) neither defaults. If the investment bank

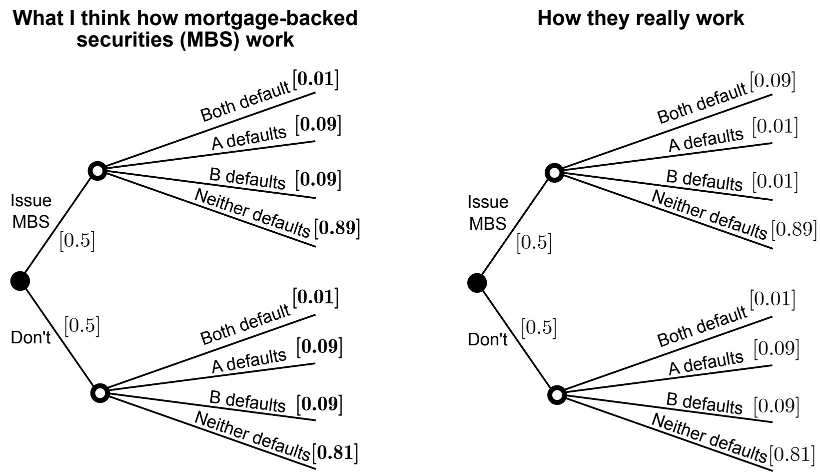


Figure 5: Correlation neglect on household probability of default

Note: Empty circles are Nature's decision nodes. Solid circles is Player's decision node. *High* and *Low* indicate Player's ability. Bold numbers in brackets are the player's subjective probabilities of Nature's moves. Other numbers in brackets are the objective probabilities of the player's and Nature's moves.

does not issue MBS, each household defaults independently and identically with 10 percent probability. Thus, the probability of both defaulting is 1 percent. If the investment bank issues MBS, the households' defaults are positively correlated: both default with 9 percent probability, only household A defaults with with 1 percent probability, only household B defaults with 1 percent probability, and neither defaults with 89 percent probability. That is, conditional on one household's default, the probability of the other household's default is 90 percent, although the unconditional probability of default for each household remains the same at 10 percent.

Without specifying the utility function for the investment bank, suppose an OE is given by the strategy and belief shown in Figure 5. Regardless of issuing MBS or not, the investment bank believes that the two households' probabilities of default are independently and identically 10 percent. With this belief, the investment bank decides to issue MBS with one half probability. By Proposition 1, this belief maximizes the entropy given his strategy, hence the OE is a MOE. As a result, the investment bank not only has correlation neglect, but also does not realize that his action causes correlation. Issuing complex financial securities causes household defaults to be correlated, yet he does not believe that such is the case. Although it is a stylized example, this misperception captures how investment banks in the period leading up to the 2007–2008 financial crisis issued mortgage-backed securities that pooled risky assets yet received the highest credit ratings, a major cause behind the crisis (Acharya and Richardson, 2009).

The literature on correlation neglect takes the correlation between variables in the data as exogenously given and analyze agents' perception of the data, in models of political behavior (Ortoleva and Snowberg, 2015; Levy and Razin, 2015) or in lab experiments (Eyster and Weizsacker, 2016; Enke and Zimmermann, 2019; Laudenbach, Ungeheuer, and Weber, 2023). The modeling framework presented here allows for joint determination of players' actions, beliefs, and correlations in the data. In this framework, correlated outcomes arise as a result of the interaction between strategic agents, and the misperceptions about the correlations in turn affect agents' behavior in equilibrium.

4.2 Omitted-variable bias

Consider the following *omitted-variable game*, a 3-stage finite extensive-form game. There is a set $N = \{1, 2, \dots, n\}$ of players. In the first stage, Nature assigns the game's *state* $t \in \mathcal{T}$ with probability $\pi(t)$. In the second stage, each player i simultaneously choose an action $x_i \in \mathcal{X}_i$, resulting in an action profile $x = (x_1, x_2, \dots, x_n) \in \mathcal{X} = \prod_i \mathcal{X}_i$. In the third stage, Nature assigns an outcome $y \in \mathcal{Y}$, whose conditional probability $\pi(y|t, x)$ depends on the history (t, x). The sets $\mathcal{T}$, $\mathcal{X}$, $\mathcal{Y}$ are finite sets of states, action profiles, and consequences. For example, in a single person's three-stage decision problem, $\mathcal{T}$, $\mathcal{X}$, and $\mathcal{Y}$ may be the sets of one's ability states (high or low), educational choices (go to college or not), and labor market outcomes (employed or not). A *terminal history* is a triple $(t, x, y) \in \Omega = \mathcal{T} \times \mathcal{X} \times \mathcal{Y}$. The payoff for each player is $u_i(t, x, y)$. Each player's *strategy* is a function $\sigma_i : \mathcal{T} \to \Delta(\mathcal{X}_i)$. A *strategy profile* σ is the collection $\{\sigma_i\}_{i \in N}$.

Let the observational constraint of this game be that players only observe the marginal probability distributions of state-action pairs (t, x) and action-outcome pairs (x, y). That is, a belief β_i is observation-consistent if

$$\sum_{y \in \mathcal{Y}} p(\sigma_i, \beta_i)(t, x, y) = \sum_{y \in \mathcal{Y}} p(\sigma_i, (\sigma_{-i}, \pi))(t, x, y), \text{ and}$$
$$\sum_{t \in \mathcal{T}} p(\sigma_i, \beta_i)(t, x, y) = \sum_{t \in \mathcal{T}} p(\sigma_i, (\sigma_{-i}, \pi))(t, x, y).$$

Equivalently, a belief β_i is observation-consistent if

$$\sigma_i(x_i|t)\beta_i(x_{-i}|t)\beta_i(t) = \sigma(x|t)\pi(t) \quad \text{for all } (t, x) \in \mathcal{T} \times \mathcal{X}, \text{ and}$$
$$\sum_{t \in \mathcal{T}} \beta_i(y|t, x)\beta_i(x_{-i}|t)\sigma_i(x_i|t)\beta_i(t) = \sum_{t \in \mathcal{T}} \pi(y|t, x)\sigma(x|t)\pi(t) \quad \text{for all } (x, y) \in \mathcal{X} \times \mathcal{Y}.$$

The left-hand and right-hand sides of the first equation are repectively the subjective and objective marginal probability of the pair (t, x). Similarly, the left-hand and right-hand sides of the second equation are respectively the subjective and objective marginal probability of the pair (x, y).

Players of this game have an omitted-variable bias in equilibrium.

Proposition 2. *An observation-consistent equilibrium (σ, β, μ) of the omitted-variable game is a MOE if and only if every player's belief β_i satisfies, for all (t, x, y),*

$$\begin{aligned} \beta_i(t) &= \pi(t), \\ \beta_i(x_{-i}|t) &= \sigma_{-i}(x_{-i}|t), \ \textit{and} \\ \beta_i(y|t,x) &= \sum_{t' \in \mathcal{T}} \pi(y|t',x) w(t',x), \end{aligned}$$

where $w(\cdot)$ is a weight function such that

$$w(t', x) = \lim_{k \to \infty} \frac{\sigma^k(x|t')\pi(t')}{\sum_{t'' \in \mathcal{T}} \sigma^k(x|t'')\pi(t'')},$$

for some sequence $\{\sigma^k\}_{k=1}^{\infty}$ of totally mixed strategy profiles converging to σ.

Put differently, players have correct beliefs about how Nature assigns states and how other players choose their actions, but have biased beliefs about the effects of their actions on consequences. Specifically, they believe that the effects of their actions are a weighted average of the objective effects $\pi(y|t, x)$ across different game states t, and do not believe that the effect depends on t. Hence, the state t is the omitted variable that the players of this game fail to account for when perceiving the effect of their actions on consequences. The intuition is that players observe no evidence that consequences y are correlated with state t, conditional on actions x. Therefore, the simplest explanation—maximizing the entropy—is to believe that, given actions x, consequences y indeed do not depend on states t. In other words, $\beta(y|t, x) = \beta(y|t', x)$ for any two states t and t'. Along with observational consistency, this condition implies that the belief $\beta(y|t, x)$ is a weighted average of the actual probabilities $\pi(y|t, x)$. Importantly, however, the belief β_i is uniquely determined only when the equilibrium strategy profile σ is totally mixed. Otherwise, the belief depends on the sequence $\{\sigma^k\}$.

To see the omitted-variable bias in a stylized setting, consider the following single-person decision problem. At birth, everyone is given *high* or *low* ability with equal probabilities. Once grown, one can attend college or not. Afterward, 80 percent of high-ability graduates find jobs, 40 percent of high-abity non-graduates find jobs, 60 percent of low-ability graduates find jobs, and 40 percent of low-ability non-graduates find jobs. Without specifying the utility function, suppose the strategy and belief shown in Figure 6 is an OE. The left panel shows the player's strategy σ and his belief β about Nature ("how I think college education

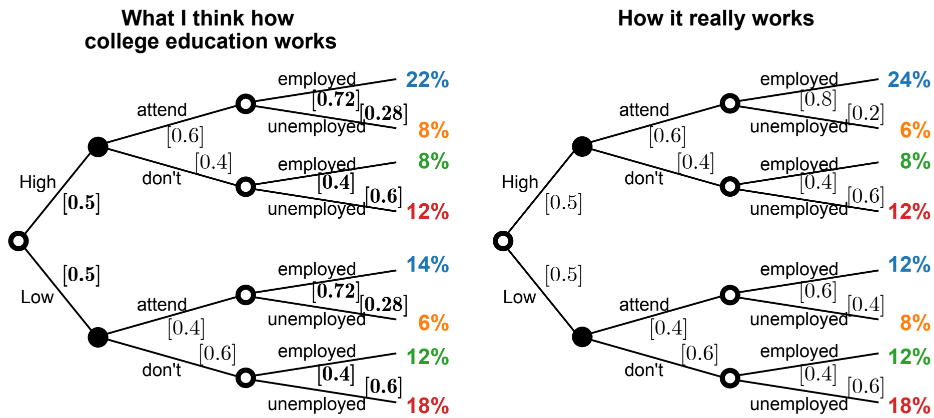


Figure 6: Omitted-variable bias about the effects of college education

Note: Empty circles are Nature's decision nodes. Solid circles are Player's decision nodes. *High* and *Low* indicate Player's ability. Bold numbers in brackets are the player's subjective probabilities of Nature's moves. Other numbers in brackets are the objective probabilities of the player's and Nature's moves. Numbers on terminal nodes are subjective (left) and objective (right) probabilities of reaching those nodes.

works"). The right panel shows the player's strategy σ and the objective probability π of Nature's moves.

We can verify the equilibrium is in fact a MOE. By Proposition 2, the subjective probability of employment after attending college is, for both states,

$$\beta(employed \,|\, \cdot, attend) = \sum_t \frac{\sigma(attend \,|\, t)\pi(t)}{\sum_{t'} \sigma(attend \,|\, t')\pi(t')} \pi(employed \,|\, t, attend)$$
$$= \frac{0.6 \cdot 0.5}{0.6 \cdot 0.5 + 0.4 \cdot 0.5} \cdot 0.8 + \frac{0.4 \cdot 0.5}{0.6 \cdot 0.5 + 0.4 \cdot 0.5} \cdot 0.6 = 0.72,$$

as indicated in the figure. From the same formula, the subjective probability of employment after not attending college is, for both states,

$$\beta(employed \,|\, \cdot, don't) = \sum_t \frac{\sigma(don't \,|\, t)\pi(t)}{\sum_{t'} \sigma(don't \,|\, t')\pi(t')} \pi(employed \,|\, t, don't)$$
$$= \frac{0.4 \cdot 0.5}{0.4 \cdot 0.5 + 0.6 \cdot 0.5} \cdot 0.4 + \frac{0.6 \cdot 0.5}{0.4 \cdot 0.5 + 0.6 \cdot 0.5} \cdot 0.4 = 0.4.$$

These numbers tell us the degrees of underestimation and overestimation of causal effects. Both high-ability and low-ability students believe that the value of attending college is a boost of $0.72 - 0.4 = 0.32$ in the probability of employment. In comparison, the objective effects of college on employment are $0.8 - 0.4 = 0.4$ for high-ability students and $0.6 - 0.4 = 0.2$ for low-ability students. On average, it is $0.5(0.4) + 0.5(0.2) = 0.3$. Consequently, high-ability students underestimate the value of education by 8 percentage points, whereas low-ability students overestimate it by 12 percentage points. On average, students overestimate the value of college education by 2 percentage points. Depending on the payoffs of the game, this misperception could result in their objectively suboptimal behavior.

This numerical example also illustrates the role of the players' strategies in forming their beliefs. The reason why the subjective effect of college education (32 percent) is larger than the average objective effect (30 percent) is that high-ability students are more likely to attend college. Because there are more high-ability students attending college, the perceived effect of college is computed as a weighted average of the objective effects while placing a greater weight on the objective effect for high-ability students, which is larger than that for low-ability students. In other words, because high-ability students are more likely to select themselves into college, people who neglect this selection effect tend to overestimate the effect of college education. In this sense, omitted-variable bias is similar to selection neglect (Koehler and Mercer, 2009; Jehiel, 2018).

In labor economics and applied econometrics, much of the focus is on reducing the omitted-variable bias from the researcher's perspective. In the meantime, papers in this literature often maintain the assumption that agents in their models have correct perceptions about the world. For example, in Keane-Wolpin-Ecstein models (Keane and Wolpin, 1994, 1997) of dynamic discrete choice, agents have abilities that are private information and are unobservable to the researcher. They subsequently make educational decisions and labor market choices over their life cycles, knowing the full probability distributions of their outcomes conditional on their choices. Thus, the solution of this model may be viewed as an extreme case when the agents know everything whereas the econometrician has an observational constraint. The MOE offers an alternative on the other extreme where the agents know only as much as the econometrician knows based on the observational constraint.

4.3 Simultaneity bias (reverse causality bias)

Next, consider the following *simultaneity game*, a 3-stage finite extensive-form game. There is a set $N = \{1, 2, \ldots, n\}$ of players. In the first stage, Nature assigns the game's binary *state* $t \in \mathcal{T} = \{t_F, t_R\}$ with probability $\pi(t)$, where t_F denotes Forward and t_R denotes Reverse. If the state is Forward, each player simultaneously chooses an action $x_i \in \mathcal{X}_i$ in the second stage, resulting in an action profile $x = (x_1, x_2, \ldots, x_n) \in \mathcal{X} = \prod_{i \in N} \mathcal{X}_i$. In the third stage, Nature chooses a consequence $y \in \mathcal{Y}$ with probability $\pi(y|t_F, x)$. If the state is Reverse, the order of moves is reversed: Nature chooses a consequence y in the second stage with probability $\pi(y|t_R)$. In the third stage, each player simultaneously choosing an action x_i, resulting in an action profile x. The sets $\mathcal{X}$ and $\mathcal{Y}$ are finite. All information sets are singletons. A *terminal history* is a triple (t, x, y). Each player's *strategy* is a function σ_i

such that the probabilty of action x_i is $\sigma_i(x_i|t_F)$ given t_F and $\sigma_i(x_i|t_R, y)$ given (t_R, y). A *strategy profile* is a collection $\{\sigma_i\}_{i\in N}$.

Let a strategy profile σ be given. A player's belief β_i is observation-consistent if for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$,

$$\sum_{t\in\mathcal{T}} p(\sigma_i, \beta_i)(t, x, y) = \sum_{t\in\mathcal{T}} p(\sigma_i, (\sigma_{-i}, \pi))(t, x, y)$$

That is, each player knows the marginal probabilities of the action-consequence pairs (x, y) but do not know the joint probability distribution of (t, x, y). Equivalently, a belief β_i is observation-consistent if, for all (x, y),

$$\begin{aligned} &\beta_i(y|t_F, x)\sigma_i(x_i|t_F)\beta_i(x_{-i}|t_F)\beta_i(t_F) + \sigma_i(x_i|t_R, y)\beta_i(x_{-i}|t_R, y)\beta_i(y|t_R)\beta_i(t_R) \\ = &\pi(y|t_F, x)\sigma_i(x_i|t_F)\sigma_{-i}(x_{-i}|t_F)\sigma_i(t_F) + \sigma_i(x_i|t_R, y)\sigma_{-i}(x_{-i}|t_R, y)\pi(y|t_R)\sigma_i(t_R). \end{aligned}$$

Unlike the simple correlation game of Section 4.1 and the omitted-variable game of Section 4.2, it is difficult to analytically characterize the beliefs in MOE of the simultaneity game. This is because, unlike in those games, the observational structure of the simultaneity game does not immediately pin down any of Nature's and other players' strategies. Numerically, however, the maximum-entropy belief is straightforward to compute by directly implementing the constrained optimization problem

$$\begin{aligned} &\beta_i^* \in \underset{\beta_i}{\operatorname{argmax}}\ G(p(\sigma_i, \beta_i)) \\ &\text{subject to } \sum_{t\in\mathcal{T}} p(\sigma_i, \beta_i)(t, x, y) = \sum_{t\in\mathcal{T}} p(\sigma_i, (\sigma_{-i}, \pi))(t, x, y) \text{ for all } x, y. \end{aligned}$$

To see a numerical solution of maximum-entropy belief in a stylized example, consider an single-player example with binary actions and binary consequences. In the first stage, Nature decides that Police Chief moves first ("Forward") with 20 percent probability and that Nature moves first ("Reverse") with 80 percent probability. If the police chief moves first, he decides whether to have large or small police presence. If he decides to put on large police presence, Nature increases violent crimes rates ("Up") with 40 percent probability and decreases them ("Down") with 60 percent probability. If he decides to have small police presence, Nature increases violent crime rates with 60 percent probability and decreases them with 40 percent probability. If Nature moves first, it increases or decreases violent crime rates with equal probabilities.

Without specifying the police chief's utility function, suppose an OE is given by Figure 7. When the police chief moves first, he chooses a large police presence or a small police presence with equal probabilities. When Nature moves first and violent crimes increase, he chooses a large police presence with 60 percent probability. When Nature moves first and violent

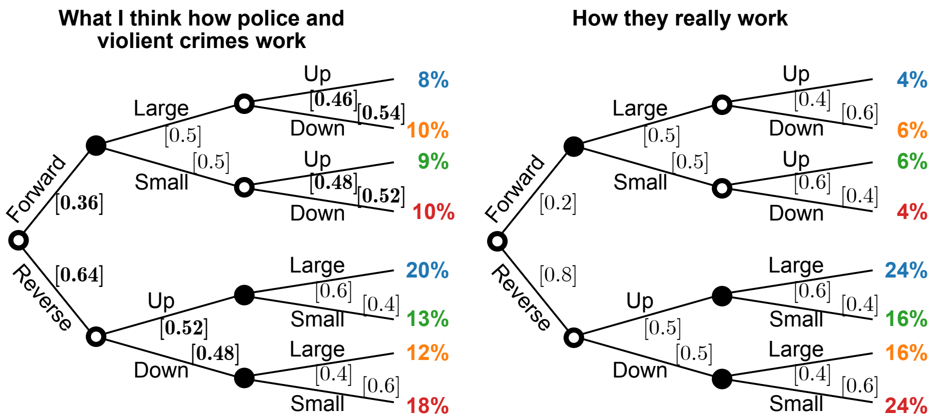


Figure 7: Simultaneity bias about the effects of police size on violent crimes

Note: Empty circles are Nature's decision nodes. Solid circles are Police Chief's decision nodes. *Forward* means that Police moves first, whereas *Reverse* means that Nature moves first. *Large* and *Small* indicates the size of police presence. *Up* and *Down* indicates the change in violent crime rates. Bold numbers in brackets are Police's subjective probabilities of Nature's moves. Other numbers in brackets are the objective probabilities of the player's and Nature's moves. Numbers on terminal nodes are subjective (left) and objective (right) probabilities of reaching those nodes.

crimes decrease, he chooses a small police presence with 60 percent probability. The police chief's belief about Nature's actions are indicated in the left panel, whereas the objective probabilities are shown in the right panel. The belief is observation-consistent because it generates the same probabilities of action-consequence pairs as the objective probabilities. Namely, according to both the subjective and objective probabilities, Large-Up pairs occur 28 percent of the time, Large-Down pairs occur 22 percent of the time, Small-Up pairs occur 22 percent of the time, and Small-Down pairs occur 28 percent of the time. Because Large-Up and Small-Down pairs arise more frequently than Large-Down and Small-Up pairs, police size and violent crime rates are overall positively correlated.

Given the police chief's decision problem and his strategy, the particular belief indicated in the left panel of Figure 7 turns out to be the maximum-entropy belief. It differs substantially from the actual probabilities of moves by Nature. Most importantly, the police chief thinks that the effect of placing a large police presence is merely a 2-percentage-points reduction (from 48 to 46 percent) in the probability of rising violent crimes, whereas the objective effect is a 20-percentage-points reduction (from 60 to 40 percent). An important reason behind this misperception is that he does not know the relative frequency of the Forward and Reverse causality: he thinks that the police affects the crime rates 36 percent of the time, when in fact it does only 20 percent of the time. Because he thinks that the Forward direction is more frequent than it really is, its effect on reducing crime cannot be so large, given the observed positive correlation between police size and violent crime rates.

As this numerical example illustrates, MOE of the simultaneity game can capture people's tendency to ignore or downplay the existence of simultaneous causal relationships. In a major newspaper article, Bump (2020) examines a scatter plot of year-over-year changes in violent crime rates and changes in police spending in the United States from 1960 to 2018. Finding near-zero correlation between the two, the article suggests that there is no causal effect of greater police spending on violent crime. However, we do not know from the scatter plot how much of the data are driven by police spending affecting crime rates or vice versa. The large econometric literature on the instrumental-variables design and similar research designs aim at isolating the forward direction to estimate the causal effects of various policy actions on consequences. On the effects of police size on crimes specifically, the literature generally finds that a greater police size reduces crimes, using instrumental variables such as electoral cycles (Levitt, 1997; McCrary, 2002), number of firefighters (Levitt, 2002), terrorist attacks (Di Tella and Schargrodsky, 2004), and government grants (Evans and Owens, 2007; Mello, 2019).

4.4 Discussion: Causality and causal misperception in games

I have so far appealed to an intuitive notion of causality in terms of probabilities without providing a precise definition. That "smoking causes cancer" meant that smoking increased the probability of cancer (Section 2), "investment banks caused the financial crisis" meant that investment banks increased the probability of households' joint default (Section 4.1), and so on. Although this probabilistic interpretation is intuitive, it is not an obvious choice. In the Potential Outcomes framework widely adopted in applied econometrics (Rubin, 1974; Angrist and Pischke, 2009; Imbens and Rubin, 2015), a causal effect of an action or treatment to an individual is defined as the difference between the observed realized outcome and the unobserved potential outcome under an alternative action, both of which are assumed to be fixed. This assumption works in econometrics because the focus is to infer the average causal effect among a population of individuals from *ex post* observations. However, this approach is not applicable for analyzing causality in games *ex ante*, where strategic players interact and causality itself may be endogenously determined from that interaction.

For this reason, I use the following probabilistic definition of subjective and objective causality. Consider a finite extensive-form game with perfect recall and an observational structure. Recall that Ω is the set of terminal histories, and π is the objective probabilities of moves by Nature. Let h be a non-terminal history of the game and let $a, b \in A(h)$ be two actions available at h. Recall that $p(\sigma_i, \beta_i)(\omega|h)$ denotes the subjective probability of reaching terminal history ω given a history h, when a player's strategy is σ_i and his belief is β_i. Let $p(\sigma_i, \beta_i)(E|h)$ denote the sum $\sum_{\omega \in E} p(\sigma_i, \beta_i)(\omega|h)$ for any event $E \subset \Omega$.

Definition 8 (Causality)**.** Let (σ, β, μ) be an observation-consistent equilibrium. An action a instead of b is a *subjective cause* of an event $E \subset \Omega$ given h to player i if

$$p(\sigma_i, \beta_i)(E|h, a) > p(\sigma_i, \beta_i)(E|h, b).$$

The action a instead of b is an *objective cause* of an event E given h if this inequality holds after replacing the player's belief β_i with objective probabilities (σ_{-i}, π).

The difference between the left-hand and right-hand sides of the inequality is player i's *subjective causal effect* of action a instead of b on event E given h. It is the *objective causal effect* of action a instead of b on event E given h if the belief β_i is replaced with the objective probabilities (σ_{-i}, π). A player has a *causal misperception* if his subjective causal effect of some action differs from its objective causal effect.

These definitions imply that causality depends on players' strategies and beliefs in the equilibrium. Objectively, an action causes an event if it increases the probability of that event according to the how everyone (including Nature) plays the game. Subjectively to a player, an action causes an event if it increases the perceived probability of that event according to how he plays and how he thinks others play.

In Appendix B, I explore a logical foundation for this probabilitic interpretation. I consider three basic axioms of causation: exclusivity, counterfactuality, and acyclicality. First, a cause has an exclusive effect: If an act x causes another act y, then x does not cause $\neg y$ (not y). Second, a cause has a counterfactual effect: If act x causes act y, then $\neg x$ is a cause of $\neg y$. Third, there are no causal cycles: For example, if act x causes act y and y causes act z, then z cannot cause x. I show that every set of causal relationships satisfying these axioms can be represented with an *extensive-form structure of acts*, defined as a probability distribution over a set of possible sequences of acts (Proposition 7). Thus, regardless of one's view on the true nature of causation, acts governed by the axioms unfold as though their underlying probabilities exist. This result provides a justification to use the probabilistic definition of causality in extensive-form games.

This game-theoretic approach to causality relates to a larger interdisciplinary literature in philosophy and the social sciences. On the one hand, philosophers have two logical (non-probabilistic) accounts: *regularity* and *counterfactual* theories. Regularity theorists (Hume, 1748; Mill, 1843; Mackie, 1965) say that an event causes another if the former is regularly followed by the latter. In contrast, counterfactual theorists (Lewis, 1973; Vihvelin, 1995) say that an event causes another if, had the former not happened, the latter would not have happened. On the other hand, social scientists broadly use probabilistic definitions of causation: an event causes another if the former raises the probability of the latter. Specifically, they define causation within controlled experiments (Wold, 1954; Rubin, 1974; Angrist and

Pischke, 2009), structural causal models (Pearl, 2009), and simultaneous equations models (Haavelmo, 1944; Heckman and Pinto, 2022), all of which include probabilistic components. Examining both sides of the views, philsophers of science (Illari, Russo, and Williamson, 2011) urge scientists to axiomatize causation.

5 Comparison with related solution concepts

Observation-consistent equilibrium (OE) and maximum-entropy observation-consistent equilibrium (MOE) are extensive-form solution concepts in which every player best responds to an observationally consistent—but possibly incorrect—belief about how other players and Nature play. MOE refines OE by selecting the belief that assumes the least information, adhering to the spirit of Occam's razor. These equilibria are closely related to two existing solution concepts in the literature: self-confirming equilibrium (Battigalli and Guaitoli, 1988, 1997; Fudenberg and Levine, 1993; Fudenberg and Kreps, 1995) and analogy-based expectation equilibrium (Jehiel, 2005; Jehiel and Koessler, 2008; Jehiel, 2022).

5.1 Self-confirming equilibrium

Consider a finite extensive-form game $(N, H, \iota, \pi, \mathcal{I}, u)$ with perfect recall and an observational structure C as described in Section 3.1. To define and relate self-confirming equilibrium to OE, I establish additional terms about players' beliefs.

Definition 9. Let a strategy profile σ be given.

1. A belief β_i is *everywhere correct* if $\beta_i = (\sigma_{-i}, \pi)$.
2. A belief β_i is *correct on the path* of σ if for all $I \in \mathcal{I}_{-i}$ such that $p(\sigma_i, (\sigma_{-i}, \pi))(I) > 0$ and for all $a \in A(I)$, the belief satisfies $\beta_i(a|I) = \sigma_{\iota(I)}$ or $\beta_i(a|I) = \pi(a|I)$.

In other words, a belief is everywhere correct if it exactly matches other player's and Nature's strategies. A belief is correct on path if it places the same probabilities on actions as other players' actual strategies and Nature's objective probabilities, at all information sets that are reached with positive probability.

I now define self-confirming and perfect Bayesian equilibrium using this language.

Definition 10. A strategy-belief-posterior triple (σ, β, μ) is a *(sequential) self-confirming equilibrium* if for every player i,

1. the strategy σ_i is sequentially rational given (β_i, μ_i),
2. the belief β_i is correct on the path of σ, and
3. the posterior function μ_i is Bayes-consistent given (σ_i, β_i).

A self-confirming equilibrium is a *(weak) perfect Bayesian equilibrium (PBE)* if every player's belief β_i is correct everywhere.

That is to say, every player in a self-confirming equilibrium best responds to a belief about how others play the game and how Nature moves, and that belief is correct about things that happen with positive proability. I note that this definition deviates from the original definition by Fudenberg and Levine (1993) in two aspects. First, the original definition does not require each player's strategy be sequentially rational but simply be an ex-ante best response. Second, they define beliefs as a probability distribution over the set of other players' (not including Nature's) strategies. Nonetheless, my reformulation above retains the essential features of self-confirming equilibrium that players best respond to beliefs that are correct about actions taken on the equilibrium path. Meanwhile, the definition of perfect Bayesian equilibrium is standard.

Under this definition, self-confirming equilibrium and perfect Bayesian equilibrium are special cases of OE and MOE, respectively, when players have perfect observation of a game's terminal histories.

Proposition 3. *Suppose the observational structure C is the identity matrix.*

1. *An assessment (σ, β, μ) is an OE if and only if it is a self-confirming equilibrium.*
2. *An assessment (σ, β, μ) is a MOE if and only if it is a perfect Bayesian equilibrium.*

Put differently, when players of a game observe the terminal histories of a game perfectly, the set of self-confirming equilibria coincides with the set of OEs, and the set of perfect Bayesian equilibria coincides with the set of MOEs. The simple intuition behind this result is that under perfect observation, observational consistency is equivalent to on-path correctness. That is, the condition for a belief β_i to be observation-consistent under perfect observation,

$$p(\sigma_i, \beta_i)(\omega) = p(\sigma_i, (\sigma_{-i}, \pi))(\omega) \quad \text{for all terminal histories } \omega \in \Omega,$$

is equivalent to the condition that the belief β_i is correct on the path of σ. To see this, notice that

$$\beta(a|h) = \frac{p(\sigma_i, \beta_i)(h, a)}{p(\sigma_i, \beta_i)(h)} = \frac{p(\sigma_i, (\sigma_{-i}, \pi))(h, a)}{p(\sigma_i, (\sigma_{-i}, \pi))(h)} = \sigma_{-i}(a|h),$$

where the equality in the middle is due to the observational consistency. From this reasoning, MOE is equivalent to perfect Bayesian equilibrium because the former requires the belief to be a limit of on-path correct beliefs. Consequently, the limit of on-path correct beliefs is an everywhere-correct belief, establishing the equivalence.

This result shows that self-confirming equilibrium and MOE are appropriate solution concepts for different types of causal misperception. Self-confirming equilibrium captures players' misperceptions about things that happen with zero probability, whereas MOE captures players' misperceptions even about things that do happen occasionally yet observed imperfectly by the players. For example, in the Manager-Worker game from Section 3.2, every information set is reached with positive probability, so self-confirming equilibrium coincides with perfect Bayesian equilibrium, having everywhere-correct beliefs. However, because the terminal histories of the game are imperfectly observed—specifically, the worker never finds out the source of an unfair bonus—he believes that the manager and Nature are equally likely causes behind any unfair bonus.

5.2 Analogy-based expectation equilibrium

Consider a finite extensive-form game with complete information: $\Gamma = (N, H, \iota, \pi, \mathcal{I}, u)$ where all information sets $I \in \mathcal{I}$ are singletons. A Player i's *analogy class* α_i is a partition of all non-terminal histories such that two histories h and h' in the same partition element are assigned to the same player j and the same set of actions $A_j(h) = A_j(h')$. An *analogy class profile* is a collection $\alpha = (\alpha_i)_{i \in N}$ of analogy classes. This environment is the original setting by Jehiel (2005). Later, Jehiel and Koessler (2008) apply the concepts to static games with incomplete information, and Ettinger and Jehiel (2010) extend them to multi-stage games with incomplete information.

For example, in the centipede game represented in Figure 8, there are four non-terminal histories: the root history ($\cdot$), the one after Alice passes (P), the one after each player passes once (PP), and the one after both players pass and Alice passes once again (PPP). One analogy class is the partition

$$\{\{\,\cdot\,, PP\}, \{P, PPP\}\}.$$

This partition is an analogy class because the two histories h_0 and PP are assigned to Player 1 and action set $\{Pass, Take\}$ and the two histories P and PPP are assigned to Player 2 with the action set $\{Pass, Take\}$. This partition is also the *coarsest analogy class* among all analogy classes of this game, as there is no other analogy class that is a coarser partition.

Let a pair (Γ, α) be a finite extensive-form game with complete information and an analogy class profile α. Recall that $p(\sigma_i, \beta_i)(h)$ is the subjective probability of reaching a history h given one's strategy σ_i and belief β_i. Write $p(\sigma)(h) = p(\sigma_i, \sigma_{-i})(h)$ to denote the objective probability of reaching history h given a strategy profile σ.

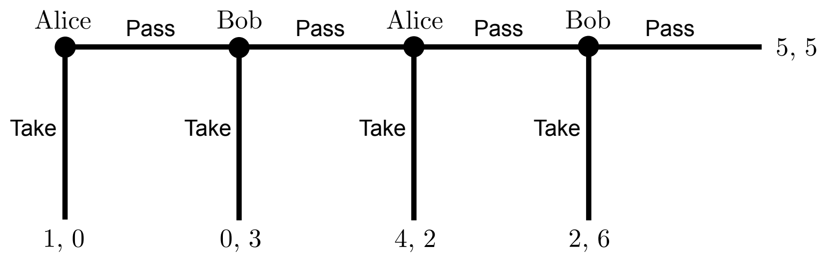


Figure 8: A four-node centipede game

Definition 11. Let a strategy profile σ be given. A belief β_i is an *analogy-based expectation* if for every non-terminal histories h assigned to a different player j,

$$\beta_i(a|h) = \sum_{h' \in \alpha_i(h)} \sigma_j(a|h')w(h')$$

where $w(\cdot)$ is a weighting function such that

$$w(h') = \lim_{k \to \infty} \frac{p(\sigma^k)(h')}{\sum_{\tilde{h} \in \alpha_i(h')} p(\sigma^k)(\tilde{h})}$$

for some sequence $\{\sigma^k\}_{k=1}^{\infty}$ of totally mixed strategy profiles that converges to σ.

Put differently, an analogy-based expectation is a player's belief that others behave in the same way in analogous situations. Specifically, each player believes that others play a mixed strategy that is a weighted average of what others actually play in those circumstances, weighted by the objective likelihoods of reaching those circumstances. The weight function is defined as a limit so that it is well-defined even when some analogy classes have zero probability of being reached.

At first glance, the interpretations of analogy-based expectations and maximum-entropy beliefs appear similar: they are the simplest explanations about others' strategies that are consistent with observation. However, the implicit assumptions about what players observe are different. The former assume that players observe others' marginal frequencies of action at analogous non-terminal histories the game. The latter assume that players observe a vector of linearly transformed frequencies of a game's terminal histories. In short, analogy-based expectations observe actions directly, whereas maximum-entropy beliefs infer actions from observed outcomes. Because of this difference, the two typically produce different point-predictions of players' beliefs about how others play.

An analogy-based expectation equilibrium (ABEE) is a solution concept in which every player best responds to one's analogy-based expectation.

Definition 12 (ABEE)**.** A pair (σ, β) of strategy profile and belief profile is an *analogy-based expectation equilibrium (ABEE)* if for every player i,

1. the strategy σ_i is sequentially rational given β_i, and

2. the belief β_i is an analogy-based expectation given the strategy profile σ.

Because analogy-based expectations and maximum-entropy beliefs produce different predictions about players' beliefs, ABEE and MOE produce different predictions about how players actually play. To illustrate the different predictions, consider the centipede game in Figure 8 and its coarsest analogy class, $\{\{\,\cdot\,, PP\}, \{P, PPP\}\}$.

Claim 6. *There are two strategy profiles supported as ABEE of the centipede game:*

1. *Take (with probability 1) at every non-terminal history, and*

2. *Pass (with probability 1) at every non-terminal history before the last one and Take at the last one.*

Proof. See Jehiel (2005), Proposition 3. ■

It is easy to see how the first strategy profile is supported as an ABEE. If both players' strategy is to always Take, each player's analogy-based expectation given this strategy profile is that the other always Takes, too. Given this belief, it is rational for everyone to always Take. Note that this strategy profile is the unique subgame-perfect Nash equilibrium. ABE equilibrium does not rule out this uncooperative, rational outcome.

However, ABEE also allows the second strategy profile, a more cooperative outcome. Since Alice Passes at both non-terminal history, Bob's belief is that Alice always Passes. Given this belief, Pass-Take is Bob's best response. Since Bob Passes and Takes with equal frequencies overall, Alice believes that he mixes between the two equally. Given this belief, always Passing is Alice's best response. Figure 9 illustrates this ABEE.

To contrast these ABEE predictions with MOE predictions, let the observational structure be given as $C = [0\ 1\ 2\ 3\ 4]$, where each element is the number of Passes played before the game ends. The interpretation is that players observe the average number of passes played but not the exact outcomes of the centipede game; that is, players have only a rough idea about the duration of cooperation.

Whereas ABEE fails to rule out the completely uncooperative outcome, MOE rules it out in this particular centipede game and observational structure.

Claim 7. *There exists no MOE of the centipede game in which Alice Takes with probability 1 at the beginning.*

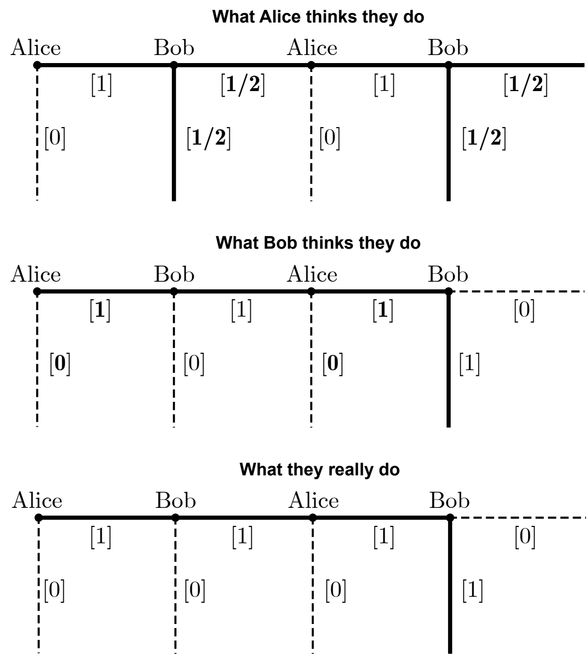


Figure 9: A non-trivial analogy-based expectation equilibrium of the centipede game

Although I include the formal proof in the Appendix, the rough reasoning is as follows. Suppose Alice Takes at the beginning with probability 1 in equilibrium, she has no observational evidence about what Bob does at the final non-terminal node. Because she has no clue, she believes that Take and Pass are equally likely. Thinking this, her best response at the node before that is to Take, because it gives her a payoff of 4 whereas Pass gives her an expected payoff of $(2+5)/2 = 3.5$. Given her best response at the third non-terminal node, she believes that both Take and Pass are equally likely from Bob at the second node. Given this belief, Passing at the first node gives her an expected payoff of $(0+4)/2 = 2$, whereas Taking gives her a sure payoff of 1. Therefore, Taking with probability 1 is not rational for Alice, a contradiction.

Since there always exists a MOE (Theorem 1), the above claim implies that there exists an equilibrium where Alice behaves more cooperatively. Although it is difficult to solve for one analytically, one can solve it numerically by a brute-force search using the equilibrium's definition directly. Specifically, I implement the following procedure:

1. Pick a strategy profile σ.
2. If σ is not totally mixed, make it totally mixed by assigning every action at least $\epsilon > 0$ probability.
3. Compute each player's belief β_i by maximizing the entropy subject to the observational consistency constraint.

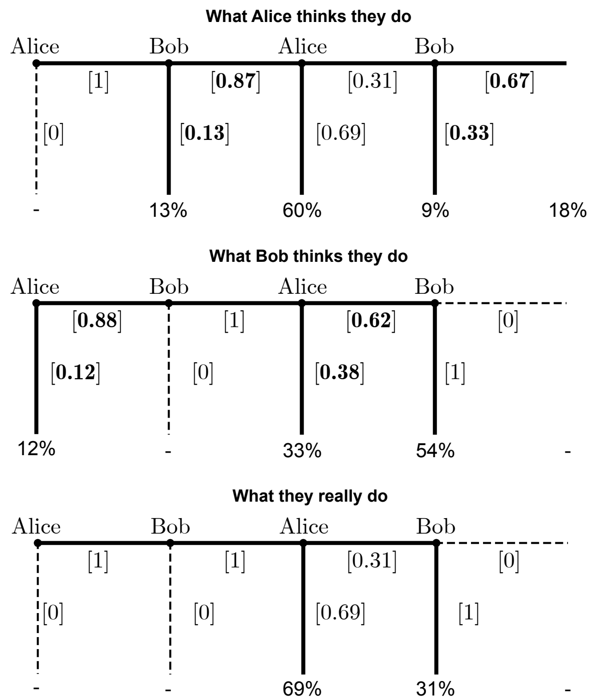


Figure 10: Unique maximum-entropy observation-consistent equilibrium (MOE) of the centipede game

4. Check whether any player has an incentive to deviate given their belief β_i.
5. If no player has an incentive to deviate, (σ, β) is an equilibrium.
6. Pick a different strategy profile and repeat from Step 2.

Figure 10 shows the unique MOE found from the above procedure. Alice's strategy is to Pass in the first node and to mix between Passing and Taking with 31 percent and 69 percent frequencies, respectively. Bob's strategy is to Pass in the second node and to Take in the fourth node. Given this strategy profile, both players observe that the average number of Pass plays per game is $0.69 \cdot 2 + 0.31 \cdot 3 = 2.3$. Given this observation and her own strategy, Alice believes that Bob plays Pass with 87 percent and 67 percent frequencies at the second and fourth nodes, respectively. This belief maximizes the entropy while being consistent with the observation that there are 2.3 Passes on average. Given this belief, her strategy is sequentially rational: she prefers Passing over Taking at the first node and is indifferent between the two at the third node. Meanwhile, given the observation and his own strategy, Bob believes that Alice plays Pass with 88 percent and 62 percent frequencies at the first and third nodes, respectively. This belief maximizes the entropy subject to the observational consistency constraint. Given this belief, his strategy is sequentially rational.

As a result, the MOE provides plausible predictions about this centipede game. It predicts probabilistic outcomes: the game will end with Alice Taking at the third node

69 percent of the time and Bob Taking at the fourth node 31 percent of the time. Both players are optimistic, believing that their opponents will Pass with moderately high probabilities. Their strategies and beliefs are sensitive to the underlying payoffs, as players' mixed strategies depend on their indifference conditions and their beliefs depend on the observed strategies. These qualitative predictions about outcomes, strategies, and beliefs are roughly in line with experimental evidence. For example, Healy (2017) finds that in a six-move centipede game with low-risk, high-reward payoffs, more than 70 percent of the games played in the lab end in the last two non-terminal nodes by either of the two players playing Take. The elicited beliefs from the subjects show that a majority of them have beliefs greater than about 40 percent that their opponents will Pass.

The centipede game shows how different ABEE and MOE can be. When do the two solution concepts look more alike? Despite ther substantial differences, there is a narrow class of games in which the two equilibria exactly coincide.

Definition 13 (Simple sequential game)**.** A *simple sequential game* is a finite extensive-form game with perfect information with the following order of moves. In the first stage, player 1 chooses an action $a_1 \in A_1$. In the second stage, player 2 chooses an action $a_2 \in A_2$. The remaining stages proceed similarly. In the nth stage, player n chooses an action $a_n \in A_n$. The game ends after the nth stage.

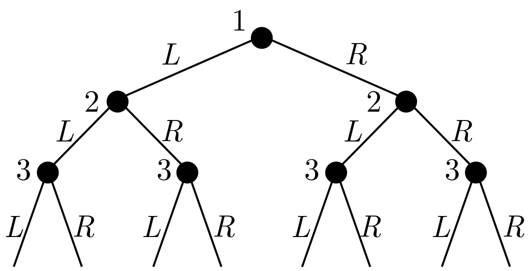
Note that, by this definition, the set of terminal histories of a simple sequential game is

$$\Omega = \prod_{i \in N} A_i$$

Importantly, this means that the set of actions available for player i is A_i regardless of the history h.

Proposition 4. *Let Γ be a simple sequential game. Let α be the coarsest analogy class profile of this game. Let C be an observational structure of this game such that players observe only the marginal probability distributions of every player's actions. Then a pair (σ, β) is an ABEE of (Γ, α) if and only if it is a MOE of (Γ, C).*

That is to say, if we restrict the class of games to simple sequential games, taking the coarsest analogy class and the marginal observational structure, ABEE and MOE are equivalent. The intuition behind this result is that under the analogy-based expectations with the coarsest analogy class, players believe that each opponent plays the same strategy as the marginal frequency of their actions in his analogy class. Under the maximum-entropy belief with the marginal observational structure, players observe the marginal frequencies of each opponent's actions but not their joint frequencies. The belief that assumes the least



$$\alpha = \{\{\,\cdot\,\}, \{L, R\}, \{LL, LR, RL, RR\}\}$$

$$C = \begin{bmatrix} 1 & 1 & 1 & 1 & \cdot & \cdot & \cdot & \cdot \\ 1 & 1 & \cdot & \cdot & 1 & 1 & \cdot & \cdot \\ 1 & \cdot & 1 & \cdot & 1 & \cdot & 1 & \cdot \end{bmatrix}$$

Figure 11: Illustration of a simple sequential game, the coarsest analogy class, and a marginal observational structure

information given this observation is that everyone's actions are independent, playing the same strategy as the observed marginal frequencies. Since the two beliefs—the analogy-based expectations and maximum-entropy beliefs—are the same, the two sets of equilibria are the same.

Figure 11 shows an illustration with three players and binary actions. Each player sequentially chooses Left (L) or Right (R). The analogy class lumps all of nodes by each player into a single group: For Player 1, the analogy group consists of the root node. For Player 2, the analogy group consists of the histories L and R. For Player 3, it consists of thie histories LL, LR, RL, and RR. In both an ABEE and a MOE (σ, β), each player's belief β_i is given by the marginal frequencies of his opponents' actions. Namely, for all $a_1, a_2, a_3 \in A = \{L, R\}$, the belief about Player 1's moves is $\beta_i(a_1|\,\cdot\,) = \sigma_1(a_1)$, The belief about Player 2's moves is $\beta_i(a_2|a_1) = \sum_{a \in A} \sigma_2(a_2|a)\sigma_1(a)$. The belief about Player 3's moves is $\beta_i(a_3|a_1, a_2) = \sum_{a,a' \in A} \sigma_3(a_3|a, a')\sigma_2(a'|a)\sigma_1(a)$.

5.3 Discussion: Interpretation of maximum-entropy observation-consistent equilibrium (MOE) revisited

I have so far interpreted observational structures of extensive-form games literally: in a large population of groups of players, players observe the terminal histories of a game in a coarse manner. In the centipede game of this section, players observe only the average number of Pass plays in the games played by other groups of players facing the same game. In a MOE, every player forms the belief consistent with the coarse observation and best responds to that belief.

(a) Prisoner's Dilemma

	Cooperate	Defect
Cooperate	$2, 2$	$-1, 3$
Defect	$3, -1$	$0, 0$

(b) Battle of the Sexes

	Ballet	Football
Ballet	$2, 1$	$0, 0$
Football	$0, 0$	$1, 2$

(c) Stag Hunt

	Stag	Rabbit
Stag	$4, 4$	$0, 3$
Rabbit	$3, 0$	$2, 2$

Figure 12: Classic simultaneous-move games

However, there is an alternative interpretation: that an observational constraint represents a degree of players' uncertainty about opponents' strategies. Even after observing the precise outcomes of games among a population of players, players may retain uncertainty about how others play, due factors such as sample size, noisy choices by players, the mental cost of figuring out others' strategies, and the differences in the precise environment. In Manager-Worker relationships (Section 3.2), even if one sees the exact outcomes of many similar worker-manager pairs, the next manager or worker one encounters may be different from the last. In lab experiments of centipede games, Healy (2017) finds that even after having played and observed the exact outcomes of multiple rounds of the centipede game, subjects retain substantial uncertainty about other players' strategies. For these reasons, MOE may be an appropriate solution concept to capture players' uncertainty about others' strategies even when they observe precise game outcomes.

This figurative interpretation is not only useful on its own, but is also useful for assessing the robustness of standard concepts to assumptions about strategic uncertainty. To illustrate this point most saliently, consider the classic simultaneous-move games in Figure 12, and suppose the observational structures of these games are such that players observe nothing about the outcomes: that is, $C = [1\ 1\ 1\ 1]$. In Prisoner's Dilemma, the MOE is to play Defect-Defect, with each player having no clue about the other's strategy and believing that both Cooperate and Defect are equally likely. Thus, unsurprisingly, the Nash equilibrium prediction of this game is highly robust to strategic uncertainty. In the Battle of the Sexes, the MOE is Ballet-Football, because each player believes that the other will be at Ballet or Football with equal chances. This result suggests that the Nash equilibrium prediction that the players will coordinate relies heavily on the assumption of strategic certainty. In the Stag Hunt game, the MOE is Rabbit-Rabbit, which similarly suggests that the cooperative outcome relies on strategic certainty.

6 Causal misperception in infinite-horizon games

Although my focus has been on finite extensive-form games, many applications of dynamic strategic interaction employ infinite horizons. Models in industrial organization, environmental economics, and macroeconomics often involve firms, governments, regulators, central

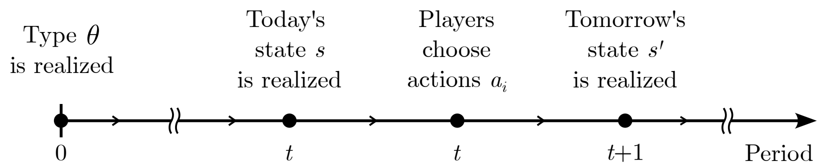


Figure 13: Summary of timing in stochastic games with permanent types

banks, and other nonhuman entities that iteract over infinitely many periods. In labor economics, even when agents are human (such as workers and students), they are sometimes modeled as living through an infinite number of periods to make the analysis tractable.

To address these settings, I extend the definitions of observation-consistent equilibrium (OE) and maximum-entropy observation-consistent equilibrium (MOE) for infinite-horizon games, particularly stochastic games. By allowing for causal misperceptions in agents that interact repeatedly, MOE offers a meaningful alternative to the standard solution concept of Markov perfect equilibrium (Maskin and Tirole, 1988a, 2001).

6.1 Equilibrium in stochastic games

Consider a *finite stochastic game with permanent types* as follows. There are a finite set N of players, a finite set Θ of permanent game types, and a finite set $S = \{s_1, s_2, \dots, s_m\}$ of states. At each state $s \in S$, each player $i \in N$ has a finite set $A_i(s)$ of available actions. Let $\pi_0(\theta, s)$ denote the initial probability of the type-state pair (θ, s) in period 0, with full support on θ. Given type θ, state s in period t, and an action vector $a = (a_i)_{i\in N}$, the objective transition probability of the next period's state being $s' \in S$ is $\pi(s'|\theta, s, a)$. Figure 13 illustrates the timing of the realization of types, states, and actions. Each player i's payoff in each period given type θ, state s, and action vector a is $u_i(\theta, s, a)$. Each player discounts their future payoffs with a discount factor $\delta \in (0, 1)$.

Each player i's *mixed Markov strategy* (or *strategy*) is a mapping σ_i that assigns each (θ, s) to a probability distribution on the set $A_i(s)$ of available actions. A *strategy profile* is a collection $\sigma = \{\sigma_i\}_{i\in N}$ of all players' strategies.

The transition probability function $\pi(s'|\theta, s, a)$ is *regular*: for every type $\theta \in \Theta$ and strategy profile σ, the stochastic matrix

$$M_\theta(\sigma_i, (\sigma_{-i}, \pi)) = \begin{bmatrix} \sum_{a\in A} \pi(s_1|\theta, s_1, a)\sigma(a|\theta, s_1) & \cdots & \sum_{a\in A} \pi(s_m|\theta, s_1, a)\sigma(a|\theta, s_1) \\ \vdots & \ddots & \vdots \\ \sum_{a\in A} \pi(s_1|\theta, s_m, a)\sigma(a|\theta, s_m) & \cdots & \sum_{a\in A} \pi(s_m|\theta, s_m, a)\sigma(a|\theta, s_m) \end{bmatrix}$$

has some power $M_\theta(\sigma_i, (\sigma_{-i}, \pi))^t$ whose elements are all nonzero. This implies that the Markov chain described by this stochastic matrix has a stationary probability vector.

Definition 14 (Belief)**.** Each player i's *belief* is a subjective probability function β_i such that

1. $\beta_i(a_{-i}|\theta, s)$ is the subjective probability of other players' actions a_{-i} given the given type θ and current state s, and
2. $\beta_i(s'|\theta, s, a)$ is subjective transition probability of the next state being s' given type θ, current state s, and current action vector $a \in A$.

Let $p(\sigma_i, \beta_i)(\theta, s)$ be the limiting distribution of (θ, s) given one's strategy σ_i and belief β_i. That is,

$$\begin{bmatrix} p(\sigma_i, \beta_i)(\theta, s_1) & \cdots & p(\sigma_i, \beta_i)(\theta, s_m) \end{bmatrix} = \lim_{t \to \infty} \begin{bmatrix} \pi_0(\theta, s_1) & \cdots & \pi_0(\theta, s_m) \end{bmatrix} M_\theta(\sigma_i, \beta_i)^t.$$

Let $p(\sigma_i, \beta_i)(\theta, s, a, s')$ be the limiting distribution of (θ, s, a, s') given one's strategy σ_i and belief β_i. That is,

$$p(\sigma_i, \beta_i)(\theta, s, a, s') = \beta_i(s'|\theta, s, a)\beta_i(a_{-i}|\theta, s)\sigma_i(a_i|\theta, s)p(\sigma_i, \beta_i)(\theta, s).$$

By substituting (σ_{-i}, π) for β_i, we have the objective probability

$$p(\sigma_i, (\sigma_{-i}, \pi))(\theta, s, a, s') = \pi(s'|\theta, s, a)\sigma_{-i}(a_{-i}|\theta, s)\sigma_i(a_i|\theta, s)p(\sigma_i, \beta_i)(\theta, s),$$

alternatively written as $p(\sigma, \pi)(\theta, s, a, s')$.

Let $\mathbf{p}(\sigma_i, \beta_i)$ denote the probability vector whose elements are $p(\sigma_i, \beta_i)(\theta, s, a, s')$ for all $(\theta, s, a, s') \in \Omega = \Theta \times S \times A \times S$. An *observational structure* C is a full-row-rank matrix C with $|\Omega|$ rows.

Definition 15. Let a strategy profile σ be given. A belief β_i is an *observation consistent* if

$$C\mathbf{p}(\sigma_i, \beta_i) = C\mathbf{p}(\sigma_i, (\sigma_{-i}, \pi)).$$

Definition 16. Let a totally-mixed strategy profile σ be given. An observation-consistent belief β_i *maximizes the entropy* if

$$\beta_i \in \operatorname*{argmax}_{\beta'_i} G(\mathbf{p}(\sigma_i, \beta'_i))$$

subject to β'_i being observation-consistent, where $G(\cdot)$ is the Shannon entropy function.

Let each player i's subjective expected lifetime payoff be defined as the value functions

$$V_i(\sigma_i, \beta_i)(\theta, s) = \mathbb{E}_{\sigma,\beta_i}\left[\sum_{t=0}^{\infty} \delta^t u_i(\theta, s_t, a_t) \,\middle|\, \theta, s_0 = s\right],$$

where the expectation is taken over the state-action pairs $\{(s_t, a_t)\}_{t=0}^{\infty}$ conditional on one's strategy σ_i, belief β_i, type θ, and the initial state $s_0 = s$. A strategy σ_i is *(subjectively) sequentially rational* if for every θ and s,

$$\sigma_i \in \operatorname*{argmax}_{\sigma_i' \in \mathcal{S}_i} V_i\left(\sigma_i', \beta_i\right)(\theta, s).$$

Definition 17 (Equilibrium)**.** An assessment (σ, β) is an *observation-consistent equilibrium (OE)* if for each $i \in N$,

1. the strategy σ_i is sequentially rational, and
2. the belief β_i is observation-consistent.

An OE (σ, β) is a maximum-entropy observation-consistent equilibrium (MOE) if there exists a sequence $\{\sigma^k, \beta^k\}_{k=1}^{\infty}$ that converges to (σ, β) and each β_i^k maximizes the entropy given σ^k.

An assessment (σ, β) is a *Markov perfect equilibrium (MPE)* if σ is sequentially rational and each player's belief β_i is everywhere correct, that is, $\beta_i = (\sigma_{-i}, \pi)$. Defined this way, MPE is a special case of MOE when agents have perfect observation, echoing the previous result that PBE is a special case of MOE under perfect observation (Proposition 3).

Proposition 5. *Suppose the observational structure C is the identity matrix. Then an assessment (σ, β) is a MOE if and only if it is a Markov perfect equilibrium.*

An interesting cases is when the observational structure C is such that players observe the marginal probability probability distributions of the triples (θ, s, a) and (s, a, s') but not the full joint probability distribution of (θ, s, a, s'). That is, a belief β_i is observation-consistent if and only if

$$\sum_{s' \in S} p(\sigma_i, \beta_i)(\theta, s, a, s') = \sum_{s' \in S} p(\sigma_i, \beta_i)(\theta, s, a, s'), \text{ and}$$
$$\sum_{\theta \in \Theta} p(\sigma_i, \beta_i)(\theta, s, a, s') = \sum_{\theta \in \Theta} p(\sigma_i, \beta_i)(\theta, s, a, s').$$

This observational structure implies that players perceive others' strategies correctly but have an omitted-variable bias on Nature's transition probabilities, by neglecting the effect of types θ on tomorrow's state s'.

Proposition 6. *Suppose the observational structure C is such that players observe only the marginal probability probability distributions of (θ, s, a) and (s, a, s'). An observation-consistent equilibrium (σ, β) is a MOE if and only if the belief β_i satisfies, for all θ, s, a, s',*

$$\beta_i(a_{-i}|\theta, s) = \sigma_{-i}(a_{-i}|\theta, s), \text{ and}$$
$$\beta_i(s'|\theta, s, a) = \sum_{\theta \in \Theta} \pi(s'|\theta, s, a) w(\theta', s, a),$$

where $w(\cdot)$ is a weight function such that

$$w(\theta', s, a) = \lim_{k \to \infty} \frac{\sigma^k(a|\theta', s) \cdot p(\sigma^k, \pi)(\theta', s)}{\sum_\theta \sigma^k(a|\theta, s) \cdot p(\sigma^k, \pi)(\theta, s)},$$

for some sequence $\{\sigma^k\}_{k=1}^\infty$ of totally mixed strategy profiles that converges to σ.

In other words, the defining features of MOE is that players have correct beliefs about other players' strategies, but have incorrect beliefs about Nature, the state transition probabilities. Specifically, they fail to account for the effect of the type θ on the next period's state. As a result, they believe that the next period's state depends only on todays' state s and players actions a. They believe that this transition probability is is some weighted average of the true transition probabilities. This intuition is the same as that in the omitted-variable game of Section 4.2.

6.2 Example: Parent-Child game on social media use

To illustrate an application, I consider a finite stochastic game involving a Child (Player 1, he) and his Parent (Player 2, she), inspired by an example mentioned in Kendall and Charles (2022). There are two permanent game types θ: one in which the Child is emotionally sensitive ($\theta = 1$) and not sensitive ($\theta = 0$). There are two states s: one in which the child is in good mood ($s = 1$) and in bad mood ($s = 0$). Every day, the child can either use social media ($a_1 = 1$) or not use it ($a_1 = 0$), whereas the parent can either enforce a lenient policy ($a_2 = 1$) or a strict policy ($a_2 = 0$) about the child using social media.

Both Parent and Child want the child to be happily in a good mood, but they differ a little about the usage of social media. The child's payoff from each period is:

$$u_1(s, a_1, a_2) = s + [0.1 + 0.1a_2 + 0.025(1 - s)]a_1.$$

That is, the child benefits temporarily from using social media, with an extra kick when under a lenient policy or when in a bad mood. The Parent's payoff from each period is:

$$u_2(\theta, s, a_1, a_2) = s - 0.05(1 - a_2).$$

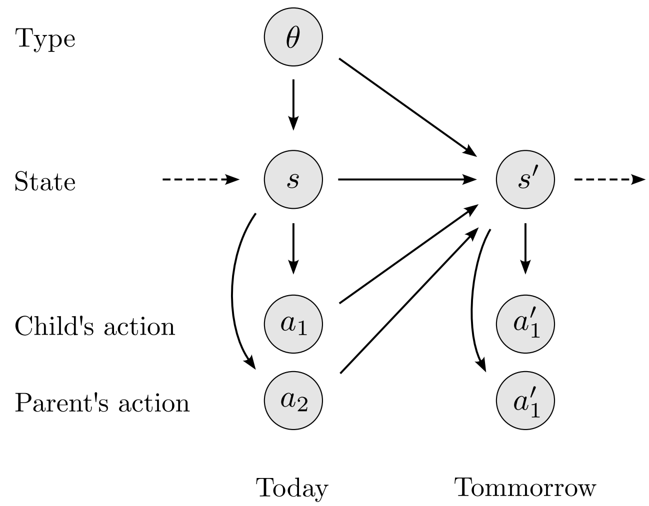


Figure 14: Causal relationships in the Parent-Child game of social media use

That is, it is costly for the parent to enforce a strict policy. The discount factor for both the child and the parent is $\delta = 0.9$.

The objective transition probabilities into a good mood in the next period ($s' = 1$) are

$$\pi(1|\theta, s, a_1, a_2) = 0.3 + 0.4s + \theta \cdot [-0.1(1-s) + 0.1s](1 + a_2)a_1.$$

This transition function captures plausible mood swings in children. First, moods are persistent: other things equal, having a good mood today makes having a good mood tomorrow more likely by 0.4. Second, social media affects a sensitive child's future mood but does not affect that of an insenstivie child. by Third, social media has opposite effects depending on the child's current mood: if a child is in a bad mood, social media makes him more likely to stay in bad mood; if he is in a good mood, social media makes him more likely to stay in good mood. Fourth, the lenient social media use policy amplifies the effect of social media. Figure 14 illustrates the direction of causality between the variables of this model.

Numerical method. I compute the pure-strategy Markov perfect equilibrium and pure-strategy MOE of the Parent-Child game with a brute-force search as follows.

1. Fix a pure-strategy profile σ.
2. Given σ, compute the belief profile β:
   (a) For MPE, set $\beta_i = (\sigma_{-i}, \pi)$ for both players.
   (b) For MOE, pick a sequence $\{\sigma^k\}$ of ϵ-constrained strategy profiles that converges to σ as $\epsilon \to 0$. Set β using Proposition 6.

Table 1: Equilibrium strategies in the Parent-Child game of social media use

		Child's strategy (σ_1)		Parent's strategy (σ_2)	
Equilibrium	Type (θ)	Bad mood	Good mood	Bad mood	Good mood
MPE	Not sensitive	Use	Use	Lenient	Lenient
	Sensitive	Don't	Use	Lenient	Lenient
MOE	Not sensitive	Use	Use	Strict	Lenient
	Sensitive	Use	Use	Strict	Lenient

Note: MPE refers to Markov perfect equilibrium. MOE refers to maximum-entropy observation-consistent equilibrium.

3. Compute the value functions $V_i(\sigma_i, \beta_i)$ by value function iteration.
4. For every (θ, s), check if anyone has an incentive to deviate: that is,

$$u_i(\theta, s, (a_i, a_{-i})) + \delta V_i(\sigma_i, \beta_i)(\theta, s) > V_i(\sigma_i, \beta_i)(\theta, s), \quad \text{for some } a_i \in A_i(s).$$

   If no player has an incentive to deviate, (σ, β) is an equilibrium.
5. Pick a different pure-strategy profile and repeat from Step 2.

In Step 2(b), this particular choice of a sequence requiring all actions to have at least $\epsilon > 0$ probability uniformly is a substantive restriction. If one allows different ϵ's for different actions, the resulting belief profile may look highly dissimilar.

Result. Using the above procedure, I find that there is a unique pure-strategy Markov perfect equilibrium and a unique pure-strategy MOE, organized in Table 1. The result shows that agents' causal misperception under observational constraint alters the predictions considerably, by inducing them to choose suboptimal actions.

In the Markov perfect equilibrium, the child's strategy is to not use social media if one is sensitive and is in a bad mood; otherwise, they use social media. The parent's strategy is to place a lenient policy at all times for all types of children. The reason behind this result is that both players have rational expectations—correct beliefs about others strategies and the transition probabilities. Specifically, the child knows that he hurts himself by using social media when he is a sensitive type in a bad mood. Knowing this, he rationally refrain from using social media. Knowing that the child has such self-discipline, the parent can afford to be lenient at all times.

In the MOE, in contrast, the child has no such self-restraint; his strategy is to use social media at all times. The parent's strategy is to enforce a strict policy when the child is having a bad mood, and have a lenient policy when he is having a good mood. The reason behind this result is that both the child and the parent have a causal misperception. They correctly perceive each other's strategy, but they incorrectly perceive that the transition probability does not depend on the child's type. Neglecting that only sensitive types are affected by social media, they think all types are affected by social media. Specifically, they believe in the the subjective transition probabilities

$$\beta(1|\theta, s, a_1, a_2) = 0.3 + 0.4s + [-0.04(1-s) + 0.06s](1+a_2)a_1,$$

These probabilities mean that the players believe using social media in a bad mood reduces tomorrow's likelihood of good mood by 4 percentage points, whereas using social media in a good mood increases it by 6 percentage points, under the lenient policy. The actual effects for the sensitive type is a 10 percentage-point reduction in a bad mood and a 10 percentage-point increase in a good mood. The actual effect for the non-sensitive type is zero. Because the players neglect the effect of these types, their perceived effect is an average of the actual effects. As a result, the child does not think social media is too harmful and use it at all times—recall that there is a temporary utiliy benefit. Consequently, the parent, who does not enjoy such temporary benefit but only sees her child in harm's way, enforces a strict policy for the child in a bad mood.

6.3 Discussion: Applications to dynamic structural econometrics

Extending maximum-entropy observation-consistent equilibrium (MOE) to infinite-horizon games makes it a viable alternative to the rational expectations paradigm in estimating parameters of dynamic decision problems or dynamic games. While it may not fully replace standard solution concepts like perfect Bayesian equilibrium or Markov Perfect Equilibrium, it offers a useful benchmark at the opposite end of the spectrum. In standard concepts, agents are assumed to know much more than the econometrician, knowing the objective causal effects of their actions whereas the econometrician does not. With an appropriate choice of the observational structure, MOE assumes that agents know only as much as the econometrician, by not observing the full joint probability distribution of variables. Thus, traditional concepts and MOE can complement each other by representing two extremes in how much rational expectations we attribute to agents. I briefly discuss potential applications in labor economics and industrial organization.

Schooling, employment, and family. In the dynamic discrete choice models by Keane and Wolpin (1994, 1997), individuals with different permanent types make lifetime decisions such as attending school, working in white- or blue-collar jobs, joining the military, or staying home. Rather than assuming they have rational expectations about the impact of these choices on future wages, we can restrict their knowledge assuming maximum-entropy beliefs. For instance, agents might only observe marginal distributions of type-action and action-outcome pairs, without knowing the full joint probability distribution, similar to the omitted-variable bias game in Section 4.2. Consequently, the standard solution concept and MOE will produce different predictions, resulting in different parameter estimates. While Keane and Wolpin (1997) find that the private gains from a college tuition subsidy are small, estimating the model with MOE could reveal larger gains, as it allows students to overestimate or underestimate the benefits of higher education on earnings. A similar approach could apply to dynamic models of family labor supply (Eckstein and Wolpin, 1989; Eckstein and Lifshitz, 2011; Eckstein, Keane, and Lifshitz, 2019), where agents may misperceive the effects of marriage or fertility on future earnings, or models of unemployment insurance (Auray, Fuller, and Lkhagvasuren, 2019; Auray and Fuller, 2020; Blasco and Fontaine, 2021), where individuals may misjudge how unemployment insurance impacts the length of their unemployment spells.

Dynamic Oligopoly. In Benkard (2004)'s dynamic oligopoly model, firms in the commercial aircraft industry compete in markets for small, medium, and large aircraft. Firms' productivity improves through learning-by-doing—gaining efficiency from past production. Instead of assuming firms have rational expectations about how production affects experience, maximum-entropy beliefs can be applied to limit their knowledge of experience transitions. Firms might observe marginal probability distributions of type-action and action-outcome pairs but not the full joint distribution, leading to over- or underestimation of the effects of learning-by-doing. Similarly, this approach could be applied to R&D competition (Goettler and Gordon, 2011; Igami, 2017) or environmental regulation (Ryan, 2012; Fowlie, Reguant, and Ryan, 2016), where firms misperceive the impacts of their actions on innovation or pollution, respectively.

7 Conclusion

I argue that maximum-entropy observation-consistent equilibrium (MOE) is a reasonable alternative to standard solution concepts in dynamic games. Despite requiring players' beliefs to be the simplest explanation consistent with observation, MOE exists in every finite extensive-form game with perfect recall. Rather than assuming a specific form of causal misperception, it takes the game's observational structure as given and lets causality and misperception arise endogenously. When agents have imperfect observation of the game's outcomes, MOE captures common causal misperceptions such as correlation neglect, omitted-variable bias, and simultaneity bias. When agents have perfect observation, MOE coincides with perfect Bayesian equilibrium (with finite horizons) and Markov perfect equilibrium (with infinite horizons).

One limitation of MOE is its computational complexity. Except in cases where agents observe marginal probability distributions, there is no closed-form solution for maximum-entropy beliefs, making them reliant on numerical optimization techniques. Although the examples I consider are simple enough that this computation is straightforward, the computational demands may present challenges for more complex games.

All in all, MOE serves as a widely applicable framework for understanding strategic behavior in dynamic games, offering sharp predictions in the presence of causal misperception and expanding the toolkit for studying games with boundedly rational players. Its ability to nest standard solution concepts further demonstrates its value as a unifying framework.

References

Acharya, Viral V. and Matthew Richardson (2009) "Causes of the financial crisis," *Critical Review*, 21 (2-3), 195–210.

Aguirregabiria, Victor and Arvind Magesan (2020) "Identification and estimation of dynamic games when players' beliefs are not in equilibrium," *The Review of Economic Studies*, 87 (2), 582–625.

Aguirregabiria, Victor and Pedro Mira (2010) "Dynamic discrete choice structural models: A survey," *Journal of Econometrics*, 156 (1), 38–67.

Angrist, Joshua D and Jörn-Steffen Pischke (2009) *Mostly harmless econometrics: An empiricist's companion*: Princeton University Press.

Aoyagi, Masaki, Venkataraman Bhaskar, and Guillaume R Fréchette (2019) "The impact of monitoring in infinitely repeated games: Perfect, public, and private," *American Economic Journal: Microeconomics*, 11 (1), 1–43.

Auray, Stéphane and David L Fuller (2020) "Eligibility, experience rating, and unemployment insurance take-up," *Quantitative Economics*, 11 (3), 1059–1107.

Auray, Stephane, David L Fuller, and Damba Lkhagvasuren (2019) "Unemployment insurance take-up rates in an equilibrium search model," *European Economic Review*, 112, 1–31.

Azrieli, Yaron (2009) "On pure conjectural equilibrium with non-manipulable information," *International Journal of Game Theory*, 38, 209–219.

Azrieli, Yaron, Christopher P Chambers, and Paul J Healy (2018) "Incentives in experiments: A theoretical analysis," *Journal of Political Economy*, 126 (4), 1472–1503.

Battigalli, Pierpaolo and Danilo Guaitoli (1988) *Conjectural equilibria and rationalizability in a macroeconomic game with incomplete information*: Università Commerciale L. Bocconi.

——— (1997) "Conjectural equilibria and rationalizability in a game with incomplete information," in *Decisions, Games and Markets*, 97–124: Springer.

Benkard, C Lanier (2004) "A dynamic analysis of the market for wide-bodied commercial aircraft," *The Review of Economic Studies*, 71 (3), 581–611.

Blasco, Sylvie and François Fontaine (2021) "Unemployment Duration and the Take-up of Unemployment Insurance."

Bump, Philip (2020) "Over the past 60 years, more spending on police hasn't necessarily meant less crime," *The Washington Post*.

Chakrabarti, Subir K and Iryna Topolyan (2016) "An extensive form-based proof of the existence of sequential equilibrium," *Economic Theory Bulletin*, 4, 355–365.

Csiszar, Imre (1991) "Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems," *The Annals of Statistics*, 2032–2066.

Di Tella, Rafael and Ernesto Schargrodsky (2004) "Do police reduce crime? Estimates using the allocation of police forces after a terrorist attack," *American Economic Review*, 94 (1), 115–133.

Eckstein, Zvi, Michael Keane, and Osnat Lifshitz (2019) "Career and family decisions: Cohorts born 1935–1975," *Econometrica*, 87 (1), 217–253.

Eckstein, Zvi and Osnat Lifshitz (2011) "Dynamic female labor supply," *Econometrica*, 79 (6), 1675–1726.

Eckstein, Zvi and Kenneth I Wolpin (1989) "The specification and estimation of dynamic stochastic discrete choice models: A survey," *The Journal of Human Resources*, 24 (4), 562–598.

Eliaz, Kfir, Simone Galperti, and Ran Spiegler (2022) "False narratives and political mobilization," *arXiv preprint arXiv:2206.12621*.

Eliaz, Kfir and Ran Spiegler (2020) "A model of competing narratives," *American Economic Review*, 110 (12), 3786–3816.

Enke, Benjamin and Florian Zimmermann (2019) "Correlation neglect in belief formation," *the Review of Economic Studies*, 86 (1), 313–332.

Ettinger, David and Philippe Jehiel (2010) "A theory of deception," *American Economic Journal: Microeconomics*, 2 (1), 1–20.

Evans, William N and Emily G Owens (2007) "COPS and Crime," *Journal of Public Economics*, 91 (1-2), 181–201.

Eyster, Erik and Georg Weizsacker (2016) "Correlation neglect in portfolio choice: Lab evidence," *Available at SSRN 2914526*.

Fowlie, Meredith, Mar Reguant, and Stephen P Ryan (2016) "Market-based emissions regulation and industry dynamics," *Journal of Political Economy*, 124 (1), 249–302.

Fudenberg, Drew and David M Kreps (1995) "Learning in extensive-form games I. Self-confirming equilibria," *Games and Economic Behavior*, 8 (1), 20–55.

Fudenberg, Drew and David K Levine (1993) "Self-confirming equilibrium," *Econometrica: Journal of the Econometric Society*, 523–545.

Fudenberg, Drew and Jean Tirole (1991) "Perfect Bayesian equilibrium and sequential equilibrium," *Journal of Economic Theory*, 53 (2), 236–260.

Goettler, Ronald L and Brett R Gordon (2011) "Does AMD spur Intel to innovate more?" *Journal of Political Economy*, 119 (6), 1141–1200.

Haavelmo, Trygve (1944) "The probability approach in econometrics," *Econometrica: Journal of the Econometric Society*, iii–115.

Healy, Paul J (2017) "Epistemic experiments: Utilities, beliefs, and irrational play," *Unpublished manuscript, Ohio State University, Columbus, OH*.

Heckman, James J and Rodrigo Pinto (2022) "Causality and econometrics,"Technical report.

Hume, David (1748) *An enquiry concerning human understanding: A dissertation on the passions. An enquiry concerning the principles of morals; the natural history of religion*, London, England: Printed for A. Miller, T. Cadell, A. Donaldson and W. Creech.

Igami, Mitsuru (2017) "Estimating the innovator's dilemma: Structural analysis of creative destruction in the hard disk drive industry, 1981–1998," *Journal of Political Economy*, 125 (3), 798–847.

Illari, Phyllis McKay, Federica Russo, and Jon Williamson (2011) *Causality in the sciences*: Oxford University Press.

Imbens, Guido W and Donald B Rubin (2015) *Causal inference in statistics, social, and biomedical sciences*: Cambridge University Press.

Jaynes, Edwin T (1957a) "Information theory and statistical mechanics," *Physical Review*, 106 (4), 620.

——— (1957b) "Information theory and statistical mechanics. II," *Physical Review*, 108 (2), 171.

Jehiel, Philippe (2005) "Analogy-based expectation equilibrium," *Journal of Economic Theory*, 123 (2), 81–104.

——— (2018) "Investment strategy and selection bias: An equilibrium perspective on overoptimism," *American Economic Review*, 108 (6), 1582–1597.

——— (2022) "Analogy-based expectation equilibrium and related concepts: Theory, applications, and beyond."

Jehiel, Philippe and Frédéric Koessler (2008) "Revisiting games of incomplete information with analogy-based expectations," *Games and Economic Behavior*, 62 (2), 533–557.

Jehiel, Philippe and Giacomo Weber (2024) "Endogenous Clustering and Analogy-Based Expectation Equilibrium."

Keane, Michael P and Kenneth I Wolpin (1994) "The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte Carlo evidence," *Review of Economics and Statistics*, 648–672.

——— (1997) "The career decisions of young men," *Journal of Political Economy*, 105 (3), 473–522.

Kendall, Chad W and Constantin Charles (2022) "Causal narratives,"Technical report.

Koehler, Jonathan J and Molly Mercer (2009) "Selection neglect in mutual fund advertisements," *Management Science*, 55 (7), 1107–1121.

Kreps, David M and Robert Wilson (1982) "Sequential equilibria," *Econometrica: Journal of the Econometric Society*, 863–894.

Laudenbach, Christine, Michael Ungeheuer, and Martin Weber (2023) "How to alleviate correlation neglect in investment decisions," *Management Science*, 69 (6), 3400–3414.

Levitt, Steven D. (1997) "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime," *American Economic Review*, 87 (3), 270–290.

——— (2002) "Using Electoral Cycles in Police Hiring to Estimate the Effects of Police on Crime: Reply," *American Economic Review*, 92 (4), 1244–1250.

Levy, Gilat and Ronny Razin (2015) "Correlation neglect, voting behavior, and information aggregation," *American Economic Review*, 105 (4), 1634–1645.

Lewis, David (1973) "Causation," *The Journal of Philosophy*, 70 (17), 556–567.

Mackie, John L (1965) "Causes and conditions," *American Philosophical Quarterly*, 2 (4), 245–264.

Maskin, Eric and Jean Tirole (1988a) "A theory of dynamic oligopoly, I: Overview and quantity competition with large fixed costs," *Econometrica: Journal of the Econometric Society*, 549–569.

——— (1988b) "A theory of dynamic oligopoly, II: Price competition, kinked demand curves, and Edgeworth cycles," *Econometrica: Journal of the Econometric Society*, 571–599.

——— (2001) "Markov perfect equilibrium: I. Observable actions," *Journal of Economic Theory*, 100 (2), 191–219.

McCrary, Justin (2002) "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime: Comment," *American Economic Review*, 92 (4), 1236–1243.

Mello, Steven (2019) "More COPS, less crime," *Journal of Public Economics*, 172, 174–200.

Mill, John Stuart (1843) *A system of logic, ratiocinative and inductive*: University of Toronto Press.

Ortoleva, Pietro and Erik Snowberg (2015) "Overconfidence in political behavior," *American Economic Review*, 105 (2), 504–535.

Osborne, Martin J and Ariel Rubinstein (1994) *A course in game theory*: MIT Press.

Pearl, Judea (2009) *Causality*: Cambridge University Press.

Proctor, Robert N (2012) "The history of the discovery of the cigarette–lung cancer link: evidentiary traditions, corporate denial, global toll," *Tobacco control*, 21 (2), 87–91.

Rubin, Donald B (1974) "Estimating causal effects of treatments in randomized and nonrandomized studies.," *Journal of Educational Psychology*, 66 (5), 688.

Rubinstein, Ariel and Asher Wolinsky (1994) "Rationalizable conjectural equilibrium: between Nash and rationalizability," *Games and Economic Behavior*, 6 (2), 299–311.

Ryan, Stephen P (2012) "The costs of environmental regulation in a concentrated industry," *Econometrica*, 80 (3), 1019–1061.

Shore, John and Rodney Johnson (1980) "Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy," *IEEE Transactions on Information Theory*, 26 (1), 26–37.

Spiegler, Ran (2016) "Bayesian networks and boundedly rational expectations," *The Quarterly Journal of Economics*, 131 (3), 1243–1290.

——— (2017) ""Data monkeys": a procedural model of extrapolation from partial statistics," *The Review of Economic Studies*, 84 (4), 1818–1841.

——— (2020) "Behavioral implications of causal misperceptions," *Annual Review of Economics*, 12, 81–106.

——— (2021) "Modeling players with random "data access"," *Journal of Economic Theory*, 198, 105374.

——— (2022) "On the behavioral consequences of reverse causality," *European Economic Review*, 149, 104258.

——— (2023) "Behavioral causal inference," *arXiv preprint arXiv:2305.18916*.

Vihvelin, Kadri (1995) "Causes, effects and counterfactual dependence," *Australasian Journal of Philosophy*, 73 (4), 560–573.

Wiswall, Matthew and Basit Zafar (2021) "Human capital investments and expectations about career and family," *Journal of Political Economy*, 129 (5), 1361–1424.

Wold, Herman (1954) "Causality and econometrics," *Econometrica: Journal of the Econometric Society*, 162–177.

A Proofs of statements omitted in the main text

A.1 Proof of Claim 1

Observe from the maximization problem that the objective function is strictly concave in β and the constraint set is compact. Thus, there exists a unique solution β. The first-order necessary (and sufficient) conditions for β are

$$
\begin{gathered}
(1-\sigma)\log((1-\sigma)\beta_0)-(1-\sigma)\log((1-\sigma)(1-\beta_0))=(1-\sigma)\lambda, \\
-\sigma\log(\sigma\beta_1)-\sigma\log(\sigma(1-\beta_1))=\sigma\lambda, \quad \text{and} \\
(1-\sigma)\beta_0+\sigma\beta_1=(1-\sigma)\pi_0+\sigma\pi_1,
\end{gathered}
$$

where λ is the Lagrange multiplier for the constraint. The unique solution of this system of equations is $\beta_0=\beta_1=(1-\sigma)\pi_0+\sigma\pi_1$. ■

A.2 Proof of Claim 2

Suppose $\sigma = 0$ in equilibrium. Observational consistency implies $\beta_0 = \pi_0$. The strategy $\sigma = 0$ being a best response implies that the perceived causal effect of smoking on cancer is greater than or equal to the reward, i.e. $\beta_1 - \beta_0 \geq r$. Thus $\beta_1 \geq \beta_0 + r$.

Suppose $\sigma = 1$ in equilibrium. Observational consistency implies $\beta_1 = \pi_1$. The strategy $\sigma = 1$ being a best response implies that the perceived causal effect of smoking on cancer is less than or equal to the reward, i.e. $\beta_1 - \beta_0 \leq r$. Thus $\beta_0 \geq \beta_1 - r$.

Suppose $\sigma \in (0,1)$ in equilibrium. Observational consistency implies $(1-\sigma)\beta_0 + \sigma\beta_1 = (1-\sigma)\pi_0 + \sigma\pi_1$. The strategy $\sigma \in (0,1)$ being a best response implies that the perceived causal effect of smoking on cancer is equal to the reward, i.e. $\beta_1 - \beta_0 = r$. The only belief that satisfies both conditions is $\beta_0 = (1-\sigma)\pi_0 + \sigma(\pi_1 - r)$ and $\beta_1 = (1-\sigma)(\pi_0 + r) + \sigma\pi_1$. ■

A.3 Proof of Claim 3

Suppose (σ, β) is a MOE. First, suppose $\sigma = 0$. By definition, there exists a sequence $(\sigma^k, \beta^k)_{k=1}^{\infty}$ converging to $(0, \beta)$ such that every σ^k is totally mixed and every β^k maximizes the entropy given σ^k. By Claim 1, $\beta^k = (\pi_0, \pi_0)$ for all k, thus $a = b = \beta$. However, $\sigma = 0$ is not a best response to the belief $\beta = (\beta, \beta)$, because the reward r is greater than the perceived null effect of smoking. It follows that this strategy-belief pair cannot be a MOE.

Second, suppose $\sigma \in (0,1)$. By Proposition 1, $\beta_0 = \beta_1 = (1-\sigma)\pi_0 + \sigma\pi_1$. Then σ is not a best response, as the reward r is greater than the perceived null effect of smoking. so this strategy-belief pair cannot be a MOE.

Third, suppose $\sigma = 1$. By definition, there exists a sequence $(\sigma^k, \beta^k)_{k=1}^{\infty}$ converging to $(1, \beta)$ such that every σ^k is totally mixed and every β^k maximizes the entropy given σ^k. By Claim 1, $\beta^k = (\pi_1, \pi_1)$ for all k, thus $\beta = (\pi_1, \pi_1)$. Then the strategy $\sigma = 1$ is a best response to β as the reward r is greater than the perceived null effect of smoking.

Therefore, the only MOE is $(1, (\pi_1, \pi_1))$. ■

A.4 Proof of Theorem 1

Let $\Gamma = (N, H, \iota, \pi, \mathcal{I}, u, C)$ be a finite extensive-form game with perfect recall and observational constraint. Let $\mathcal{S}^\epsilon \subset \mathcal{S}$ be the set of ϵ-constrained profiles, that is, the set of strategy profiles such that each action is taken with at least $\epsilon > 0$ probability. Let $\varphi : \mathcal{S}^\epsilon \to \operatorname{int} \mathcal{B}$ map each strategy profile $\sigma \in \mathcal{S}^\epsilon$ to the belief profile β such that each $\beta_i \in \mathcal{B}_i$ maximizes the entropy given σ_{-i}: For all $i \in N$,

$$\beta_i \in \operatorname*{argmax}_{\beta_i' \in \mathcal{B}_i} G(\mathbf{p}(\sigma_i, \beta_i')) \qquad \text{subject to} \qquad C\mathbf{p}(\sigma_i, \beta_i') = C\mathbf{p}(\sigma_i, (\sigma_{-i}, \pi)). \tag{1}$$

Lemma 1. *The function φ is well-defined and continuous.*

Proof. Let $\sigma_i \in \mathcal{S}_i$ and $\beta_i \in \mathcal{B}$ be given, and let $\mathcal{Q}$ denote the corresponding constraint set from (1). Consider the vector $\mathbf{p}(\sigma_i, \beta_i)$, each of whose elements is $p(\sigma_i, \beta_i)(\omega)$, the probability of reaching the terminal history $\omega \in \Omega$. Observe that for every terminal history $\omega = (a_1, a_2, \ldots, a_K) \in \Omega$ of depth K, we have

$$p(\sigma_i, \beta_i)(\omega) = \prod_{0 \geq k \geq K-1,\ \iota(h_k) = i} \sigma_i(a_{k+1} | h_k) \prod_{0 \geq k \geq K-1,\ \iota(h_k) \neq i} \beta_i(a_{k+1} | h_k).$$

Because player i's behavioral strategy $\sigma_i(\cdot|\cdot)$ and belief $\beta_i(\cdot|\cdot)$ enter the probabilities $p(\sigma_i, \beta_i)(\omega)$ multiplicatively, there exist a matrix $\widetilde{C}$ and a homeomorphism $\widetilde{\mathbf{p}}(\cdot) : \mathcal{B}_i \to [0,1]^{|\Omega|}$ such that

$$\mathcal{Q} = \left\{ \beta_i \in \mathcal{B}_i : \widetilde{C}\widetilde{\mathbf{p}}(\beta_i) = \widetilde{C}\widetilde{\mathbf{p}}(\sigma_{-i}, \pi) \right\}.$$

This implies that β_i maximizes the entropy satisfying (1) if and only if

$$\widetilde{\mathbf{p}}(\beta_i) \in \operatorname*{argmax}_{\mathbf{q} \in [0,1]^{|\Omega|}} \sum_{\omega \in \Omega} -q(\omega) \log(q(\omega)) \qquad \text{subject to} \qquad C\mathbf{q} = C\widetilde{\mathbf{p}}(\sigma_{-i}, \pi).$$

The constraint set of the above maximization problem is nonempty, compact, and continuous in σ_{-i}. Thus, by the Theorem of the Maximum, the correspondence maping σ_{-i} to the solution set is nonempty-valued and continuous. Furthermore, because the objective function is strictly concave and infinitely steep at the boundary, the solution set is a singleton. This implies that the correspondence that maps σ to the solution set (1) is a singleton and continuous. As a result, the function φ continuously maps each strategy profile σ to a unique belief profile $\beta = (\beta_i)_{i \in N}$. This completes the proof of Lemma 1.

Let $\psi : \mathcal{S}^\epsilon \times \operatorname{int} \mathcal{B} \to \mathcal{M}$ map each pair (σ, β) to the Bayes-consistent belief profile: $\mu = (\mu_i)_{i \in N}$ such that for all $i \in N$, $I \in \mathcal{I}_i$, and $h \in I$,

$$\mu_i(h|I) = \frac{p(\sigma_i, \beta_i)(h)}{\sum_{h \in I} p(\sigma_i, \beta_i)(h)}. \tag{2}$$

Lemma 2. *The function ψ is well-defined and continuous.*

Proof. Observe that ψ is defined on a domain such that $\sigma_i(a|h)$ and $\beta_i(a'|h')$ are strictly positive for all $i \in N$, $a \in h$, $a' \in h'$, $h \in \mathcal{I}_i$, and $h' \in \mathcal{I}_{-i}$. Because the denominator in (2) is always positive, the posterior function μ_i for every player $i \in N$ is well-defined. In addition, because $\sigma_i(a|h)$ and $\beta_i(a'|h')$ enter multiplicatively to every $p(\sigma_i, \beta_i|h)$, the belief system μ_i is continuous in (σ_i, β_i). Therefore, the function ψ that maps every pair (σ, β) to the belief profile $\mu = (\mu_i)_{i \in N}$ is well-defined and continuous. This completes the proof of Lemma 2.

Let a correspondence $\Phi : \mathcal{P} \times \mathcal{M} \twoheadrightarrow \mathcal{S}^\epsilon$ map each pair (β, μ) of belief and posterior function profiles to the set of strategy profiles σ that are sequentially rational given (β, μ). That is, $\sigma \in \Phi(\beta, \mu)$ if and only if for every strategic player $i \in N$ and every information set $I \in \mathcal{I}_i$ of that player,

$$\sigma(\cdot|I) \in \operatorname*{argmax}_{\tilde{\sigma}_i(\cdot|I) \in \mathcal{S}_i^\epsilon(I)} U_i(\tilde{\sigma}_i, \beta_i, \mu_i|I), \tag{3}$$

where $\mathcal{S}_i^\epsilon(I) \subset \Delta(A(I))$ denotes the set of ϵ-constrained mixed strategies at the information set I.

Lemma 3. *The correspondence Φ is nonempty-valued, convex-valued, and upper hemicontinuous.*

Proof. Observe that player i's expected utility at an information set I is

$$U_i(\sigma_i, \beta_i, \mu_i|I) = \sum_{h \in I} \sum_{\omega \in \Omega(h)} u_i(\omega) \cdot p(\sigma_i, \beta_i)(\omega|h) \cdot \mu_i(h|I),$$

where for every non-terminal history $h = (a_1, a_2, \ldots, a_K) \in H$ and its terminal successor $\omega = (h, a_{K+1}, a_{K+2}, \ldots, a_L) \in \Omega$, we have

$$p(\sigma_i, \beta_i)(\omega|h) = \prod_{K \le k \le L-1,\ \iota(h_k)=i} \sigma_i(a_{k+1}|h_k) \prod_{K \le k \le L-1,\ \iota(h_k) \neq i} \beta_i(a_{k+1}|h_k).$$

After substituting for all $p(\sigma_i, \beta_i)(\omega|h)$ in the expected utility, we see that the maximization problem in (3) has a nonempty compact constraint set that is (trivially) continuous in (β_i, μ_i). In addition, the objective function is continuous in σ_i. Therefore, by the Theorem of the Maximum, the correspondence Φ is nonempty-valued and upper hemicontinuous.

Furthermore, suppose two strategy profiles $\sigma, \sigma' \in \mathcal{S}$ yield the same expected utility for every strategic player at every information set, given a belief profile $\beta \in \mathcal{B}$ and a posterior function profile $\mu \in \mathcal{M}$. That is,

$$U_i(\sigma_i, \beta_i, \mu_i|I) = U_i(\sigma_i', \beta_i, \mu_i|I),$$

for every $i \in N_{-0}$ and $I \in \mathcal{I}_i$. Fix $\lambda \in [0,1]$, and let σ'' be the strategy profile such that

$$\sigma_i''(a|I) = (1-\lambda)\sigma_i(a|I) + \lambda\sigma_i(a|I),$$

for every player i, every information set $I \in \mathcal{I}_i$, and every action $a \in A(I)$. From the definition of the expected utility, the strategy profile σ'' yields the same expected utility as σ and σ'. Thus, $\sigma'' \in \Phi(\beta, \mu)$ whenever $\sigma \in \Phi(\beta, \mu)$ and $\sigma' \in \Phi(\beta, \mu)$.

Therefore, the correspondence Φ is nonempty-valued, convex-valued, and upper hemicontinuous. This completes the proof of Lemma 3.

Consider the correspondence $F : \mathcal{S}^\epsilon \twoheadrightarrow \mathcal{S}^\epsilon$ such that

$$F(\sigma) = \Phi\Big(\varphi(\sigma), \psi\big(\sigma, \varphi(\sigma)\big)\Big).$$

That is, $F(\sigma)$ is the set of strategy profiles in $\mathcal{S}^\varepsilon$ such that every player is sequentially rational given the beliefs β_i and posterior functions μ_i generated by σ. From Lemmas 1–2, the functions φ and ψ are well-defined and continuous. From Lemma 3, the correspondence Φ is nonempty-valued, convex-valued, and upper hemicontinuous. Therefore, the correspondence F is nonempty-valued, convex-valued, and upper hemicontinuous. Hence by Kakutani's fixed point theorem, the correspondence F has a fixed point σ^ε.

Consider a sequence of positive real numbers $\epsilon_1, \epsilon_2, \ldots$ that converges to 0. Pick a corresponding sequence of fixed points of $F : \mathcal{S}^\epsilon \twoheadrightarrow \mathcal{S}^\epsilon$. By Bolzano-Weirstrass theorem, there exists a subsequence $\{\sigma^k\}_{k=1}^\infty$ that converges to some σ^* as $k \to \infty$. Let

$$\beta^* = \lim_{k\to\infty} \varphi(\sigma^k) \qquad \text{and} \qquad \mu^* = \lim_{k\to\infty} \psi(\sigma^k, \varphi(\sigma^k)),$$

whose limits exist because the functions φ and ψ are continuous (Lemmas 1–2).

I claim that $(\sigma^*, \beta^*, \mu^*)$ is a MOE. First, the strategy profile σ^* is sequentially rational given (β^*, μ^*). Because each σ^k is a fixed point of $F : \mathcal{S}^\epsilon \twoheadrightarrow \mathcal{S}^\epsilon$, we have

$$\sigma^k \in \Phi\Big(\varphi(\sigma^k), \psi\big(\sigma^k, \varphi(\sigma^k)\big)\Big).$$

By Lemma 3, the correspondence Φ is upper hemicontinuous. This implies that $\sigma^* \in \Phi(\beta^*, \mu^*)$. That is, the strategy profile σ^* is sequentially rational given (β^*, μ^*).

Second, β^* satisfies observational consistency and maximum entropy condition. This is because by construction, each $\beta^k = \varphi(\sigma^k)$ satisfies

$$\beta_i^k \in \operatorname*{argmax}_{\beta_i \in \mathcal{B}_i} G(p(\sigma_i^k, \beta_i)) \qquad \text{subject to} \qquad Cp(\sigma_i^k, \beta_i) = Cp(\sigma_i^k, \sigma_{-i}^k),$$

for every strategic player $i \in N$.

Third, the belief profile μ^* satisfies Bayes rule on the equilibriu path of play. Observe that each $\mu^k = \psi(\sigma^k, \varphi(\sigma^k))$ satisfies

$$\mu_i^k(h|I) = \frac{p(\sigma_i^k, \beta_i^k)(h)}{\sum_{h \in I} p(\sigma_i^k, \beta_i^k)(h)}.$$

for every player $i \in N$, every information set $I \in \mathcal{I}_i$, and every terminal history $h \in I$. This implies that μ^* satisfies

$$\mu^*(h|I) = \frac{p(\sigma^*, \beta^*)(h)}{p(\sigma^*, \beta^*)(I)},$$

for every information set I reached with positive probability.

Since $(\sigma^*, \beta^*, \mu^*)$ is a MOE, it completes the proof that there exists one in every finite extensive-form game with perfect recall and observational constraint. ■

A.5 Proof of Claim 4

Suppose $((\sigma_1, \sigma_2), \mu)$ is a PBE.

Case 1: Suppose $\sigma_1 = 0$. By Bayes rule, $\mu = 0$. Given this posterior, Worker's expected utility of accepting an unfair offer is 1 whereas rejecting is 0, thus $\sigma_2 = 1$. However, $sigma_1 = 0$ is not a best response to $\sigma_2 = 1$, as deviating to choosing an unfair offer gives Manager an expected payoff of 9 instead of 5. Thus, this case cannot be a PBE.

Case 2: Suppose $\sigma_1 = 1$. By Bayes rule, $\mu = 1$. Given this posterior, Worker's expected utility of accepting an unfair offer is 1 whereas rejecting is 3, thus $\sigma_2 = 0$. However, $\sigma_1 = 1$ is not a best response to $\sigma_2 = 0$, as deviating to choosing a fair offer gives Manager an expected payoff of 2.5 instead of 0. Thus, this case cannot be a PBE.

Case 3: Suppose $\sigma_1 \in (0,1)$. Since σ_1 is a best response to σ_2, Manager is indifferent between choosing fair or unfair offer. That is,

$$9 \cdot \sigma_2 = 9 \cdot 0.5\sigma_2 + 5 \cdot 0.5,$$

implying that $\sigma_2 = 5/9$. Since σ_2 is a best response to σ_1, Worker is indifferent between accepting and rejecting. That is,

$$1 = 3\mu,$$

implying that $\mu = 1/3$. Observe that by Bayes rule, $\mu = \frac{\sigma_1}{\sigma_1 + 0.5(1-\sigma_1)}$. Thus $\sigma_1 = 1/5$.

Cases 1–3 show that $((\sigma_1, \sigma_2), \mu) = ((1/5, 5/9), 1/3)$ is the only PBE. ■

A.6 Proof of Claim 5

Suppose $((\sigma_1^*, \sigma_2^*), (\beta_1^*, \beta_2^*), \mu^*)$ is an MOE. Let $\pi = 1/3$ denote Nature's probability of choosing unfair offer conditional on Manager having chosen the unfair offer.

Step 1: Showing that $\beta_1^* = (\sigma_2^*, \pi)$. Let $(\sigma_1, \sigma_2) \in (0,1)^2$ be given, and suppose Manager's belief $\beta_1 = (\beta_{10}, \beta_{12})$ maximizes the entropy. By definition, it satisfies $C\mathbf{p}(\sigma_1, \beta_1) = C\mathbf{p}(\sigma_1, (\sigma_2, \pi))$, where

$$\mathbf{p}(\sigma_1, \beta_1) = \begin{bmatrix} \sigma_1\beta_{12} \\ \sigma_1(1-\beta_{12}) \\ (1-\sigma_1)\beta_{10}\beta_{12} \\ (1-\sigma_1)\beta_{10}(1-\beta_{12}) \\ (1-\sigma_1)(1-\beta_{10}) \end{bmatrix}.$$

Equivalently, we have

$$\begin{bmatrix} 1 & \cdot & 1 & \cdot & \cdot \\ \cdot & 1 & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 \end{bmatrix} \begin{bmatrix} \sigma_1\beta_{12} \\ \sigma_1(1-\beta_{12}) \\ (1-\sigma_1)\beta_{10}\beta_{12} \\ (1-\sigma_1)\beta_{10}(1-\beta_{21}) \\ (1-\sigma_1)(1-\beta_{10}) \end{bmatrix} = \begin{bmatrix} 1 & \cdot & 1 & \cdot & \cdot \\ \cdot & 1 & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 \end{bmatrix} \begin{bmatrix} \sigma_1\sigma_2 \\ \sigma_1(1-\sigma_2) \\ (1-\sigma_1)\pi\sigma_2 \\ (1-\sigma_1)\pi(1-\sigma_2) \\ (1-\sigma_1)(1-\pi) \end{bmatrix}.$$

The third equation in the above system of equations imply that $\beta_{10} = \pi$. By substituting for $\beta_{10} = \pi$ in the first equation, we get $\beta_{12} = \sigma_2$. Therefore, every maximum-entropy belief of Manager is (σ_2, π). Because (σ_2^*, β_1^*) is the limit of some sequence $\{(\sigma_2^k, \beta_1^k)\}_{k=1}^{\infty}$ where each β_1^k maximizes the entropy given σ_2^k, we have $\beta_1^* = (\sigma_2^*, \pi)$.

Step 2: Showing that $\beta_{21}^* = (1-\beta_{21}^*)\beta_{20}^* = \frac{1}{2}\left[1-(1-\sigma_1^*)(1-\pi)\right]$. Let $(\sigma_1, \sigma_2) \in (0,1)^2$ be given, and suppose Worker's belief $\beta_2 = (\beta_{20}, \beta_{21})$ maximizes the entropy. By definition, it satisfies $C\mathbf{p}(\sigma_2, (\beta_{20}, \beta_{21})) = C\mathbf{p}(\sigma_2, (\sigma_1, \pi))$. That is,

$$\begin{bmatrix} 1 & \cdot & 1 & \cdot & \cdot \\ \cdot & 1 & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 \end{bmatrix} \begin{bmatrix} \beta_{21}\sigma_2 \\ \beta_{21}(1-\sigma_2) \\ (1-\beta_{21})\beta_{20}\sigma_2 \\ (1-\beta_{21})\beta_{20}(1-\sigma_2) \\ (1-\beta_{21})(1-\beta_{20}) \end{bmatrix} = \begin{bmatrix} 1 & \cdot & 1 & \cdot & \cdot \\ \cdot & 1 & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 \end{bmatrix} \begin{bmatrix} \sigma_1\sigma_2 \\ \sigma_1(1-\sigma_2) \\ (1-\sigma_1)\pi\sigma_2 \\ (1-\sigma_1)\pi(1-\sigma_2) \\ (1-\sigma_1)(1-\pi) \end{bmatrix}.$$

After eliminating σ_2, the above system of equations is equivalent to

$$(1-\beta_{21})(1-\beta_{20}) = (1-\sigma_1)(1-\pi).$$

Since (β_{20}, β_{21}) maximizes the entropy for Worker, it satisfies

$$\begin{aligned} (\beta_{20}, \beta_{21}) \in \underset{(\beta'_{20}, \beta'_{21}) \in [0,1]^2}{\operatorname{argmax}} \; & G(\mathbf{p}(\sigma_2, (\beta'_{20}, \beta'_{21}))) \\ \text{subject to} \quad & (1-\beta'_{21})(1-\beta'_{20}) = (1-\sigma_1)(1-\pi). \end{aligned}$$

With a change of variables $a = \beta_{21}$ and $b = (1-\beta_{21})\beta_{20}$, we have

$$\begin{aligned} (a,b) \in \underset{(a,b) \in [0,1]^2}{\operatorname{argmax}} \; & -a\log a - b\log b \\ \text{subject to} \quad & a + b = 1-(1-\sigma_1)(1-\pi), \end{aligned}$$

which implies that $a = b = \frac{1}{2}\left[1-(1-\sigma_1)(1-\pi)\right]$. It follows that

$$\beta_{21} = (1-\beta_{21})\beta_{20} = \frac{1}{2}\left[1-(1-\sigma_1)(1-\pi)\right].$$

Since (σ_2^*, β_2^*) is the limit of a sequence of maximum-entropy beliefs, the values β_{21}^* and $(1-\beta_{21}^*)\beta_{20}^*$ also equal $\frac{1}{2}\left[1-(1-\sigma_1)(1-\pi)\right]$.

Step 3: Showing that $(\sigma_1^*, \sigma_2^*) = (0,0)$, $(\beta_{10}^*, \beta_{12}^*) = (1/3, 0)$, $(\beta_{20}^*, \beta_{21}^*) = (1/5, 1/6)$, and $\mu^* = 0.5$. By Bayes rule,

$$\mu^* = \frac{\beta_{21}^*}{\beta_{21}^* + (1-\beta_{21}^*)\beta_{20}^*} = \frac{1}{2}.$$

Given this posterior belief, Worker's best response is to always reject, thus $\sigma_2^* = 0$. From Step 1, Manager's belief is $(\beta_{10}^*, \beta_{12}^*) = (1/3, 1/2)$. Given this belief, Manager's best response is to always choose the fair offer, thus $\sigma_1^* = 0$. From Step 2, this implies that $(\beta_{20}^*, \beta_{21}^*) = (1/5, 1/6)$. ■

A.7 Proof of Proposition 1

Suppose an assessment (σ, β, μ) is an OE. First, suppose the strategy profile σ is totally mixed. From the observational structure, we have, for every player $i \in N$ and action profile $x \in \mathcal{X}$,

$$\sigma_i(x_i)\beta_i(x_{-i}) = \sigma_i(x_i)\sigma_{-i}(x_{-i}) \neq 0,$$

thus $\beta_i(x_{-i}) = \sigma_{-i}(x_{-i})$. By definition, a belief β_i maximizes the entropy if and only if

$$\beta_i \in \underset{\beta_i' \in \mathcal{B}_i}{\operatorname{argmax}} \sum_{x,y_1,y_2} -\beta(y_1, y_2|x)\sigma(x) \log\Big(\beta(y_1, y_2|x)\sigma(x)\Big)$$

subject to

$$\begin{aligned}
\sum_{y_2'} \beta_i(y_1, y_2'|x)\sigma(x) &= \sum_{y_2'} \pi(y_1, y_2'|x)\sigma(x) \quad \text{for all } y_1 \in \mathcal{Y}_1 \text{ and} \\
\sum_{y_1'} \beta_i(y_1', y_2|x)\sigma(x) &= \sum_{y_1'} \pi(y_1', y_2|x)\sigma(x) \quad \text{for all } y_2 \in \mathcal{Y}_2.
\end{aligned}$$

The first-order necessary and sufficient conditions for β_i are

$$-\log(\beta_i(y_1, y_2|x)\sigma(x)) - 1 = \lambda(x, y_1) + \gamma(x, y_2) \quad \text{for all } x, y_1, y_2,$$

where $\lambda(x, y_1)$ and $\gamma(x, y_2)$ are the Lagrange multipliers for the first and second constraints. Note from the above that

$$\log(\beta_i(y_1, y_2|x)\sigma(x)) + \log(\beta_i(y_1', y_2'|x)\sigma(x)) = \log(\beta_i(y_1, y_2'|x)\sigma(x)) + \log(\beta_i(y_1', y_2|x)\sigma(x)),$$

for all y_1, y_2, y_1', y_2', therefore

$$\beta_i(y_1, y_2|x)\beta_i(y_1', y_2'|x) = \beta_i(y_1, y_2'|x)\beta_i(y_1', y_2|x).$$

By summing both sides of the equaiton with respect to all y_1' and y_2', we get $\beta_i(y_1, y_2|x) = \beta_i(y_1|x)\beta_i(y_2|x)$. By substituting this into the original observational consistency constraints, we get

$$\beta_i(y_1, y_2|x) = \pi(y_1|x)\pi(y_2|x).$$

Therefore, an OE (σ, β, μ) with a totally mixed strategy profile σ is a MOE if and only if every player's belief β_i is the one such that $\beta_i(x_{-i}) = \sigma_{-i}(x_{-i})$ and $\beta_i(y_1, y_2|x) = \pi(y_1|x)\pi(y_2|x)$.

Next, suppose (σ, β, μ) is a MOE, and suppose the strategy profile σ is not totally mixed. By definition, there exists a sequence $(\sigma^k, \beta^k)_{k=1}^{\infty}$ converging to (σ, β) such that each σ^k is totally mixed and each β^k is the MOE given σ^k. Then from the same first-order conditions as above, we have, for every k and every player i,

$$\beta_i^k(x_{-i}) = \sigma_{-i}^k(x_{-i}) \quad \text{for all } x_{-i}, \text{ and} \tag{4}$$

$$\beta_i^k(y_1, y_2|x) = \pi(y_1|x)\pi(y_2|x), \quad \text{for all } x, y_1, y_2. \tag{5}$$

Since $\beta^k \to \beta_i$ and $\sigma^k \to \sigma$, we have $\beta_i(x_{-i}) = \sigma_{-i}(x_{-i})$ and $\beta_i(y_1, y_2|x) = \pi(y_1|x)\pi(y_2|x)$. Suppose, conversely, that β_i satisfies this condition. Let $\{\sigma^k\}_{k=1}^{\infty}$ be a sequence of totally mixed strategy profiles that converges to σ. For every k, let β^k maximize the entropy given σ^k. From the first-order conditions, we know that each β^k satisfies (4)–(5), and thus converges to β. Therefore, (σ, β, μ) is a MOE.

A.8 Proof of Proposition 2

Suppose β_i is a maximum-entropy belief given a totally mixed strategy profile σ. By observational consistency, $\beta_i(t) = \pi(t)$ and $\beta_i(a_{-i}|t) = \sigma_{-i}(a_{-i}|t)$ for all players i. Furthermore, because each β_i maximizes the entropy, it maximizes

$$\sum_{t,x,y} -\beta_i(y|t,x)\sigma(x|t)\pi(t)\log(\beta_i(y|t,x)\sigma(x|t)\pi(t))$$

subject to

$$\sum_{y} \beta_i(y|t,x)\sigma(x|t)\pi(t) = \sigma(x|t)\pi(t) \quad \text{for all } (t,x) \in \mathcal{T} \times \mathcal{X}, \text{ and}$$

$$\sum_{t} \beta_i(y|t,x)\sigma(x|t)\pi(t) = \sum_{t} \pi(y|t,x)\sigma(x|t)\pi(t) \quad \text{for all } (x,y) \in \mathcal{X} \times \mathcal{Y}.$$

The first-order necessary conditions require that for all (t, x, y),

$$-\log(\beta_i(y|t,x)\sigma(x|t)\pi(t)) - 1 = \lambda(t,x) + \gamma(x,y)$$

where $\lambda(t, x)$ and $\gamma(x, y)$ are the Lagrange multipliers for the two constraints. By substituting (t', x, y'), (t', x, y), and (t, x, y') for (t, x, y) and eliminating the multipliers from the resulting equations, we get

$$\beta_i(y|t,x)\sigma(x|t)\pi(t) \cdot \beta_i(y'|t',x)\sigma(x|t')\pi(t') = \beta_i(y|t',x)\sigma(x|t')\pi(t') \cdot \beta_i(y'|t,x)\sigma(x|t)\pi(t).$$

for all t, t', x, y, y'. This equation implies that $\beta_i(y|t,x) = \beta_i(y|t',x)$. Then from the second equation of the observational consistency condition, we have

$$\beta_i(y|t,x) = \sum_{t' \in \mathcal{T}} \frac{\pi(y|t',x)\sigma(x|t')\pi(t')}{\sum_{t'' \in \mathcal{T}} \sigma(x|t'')\pi(t'')}.$$

Suppose an assessment (σ, β, μ) is a MOE. By definition, there exists a sequence $\{(\sigma^k, \beta^k)\} \to (\sigma, \beta)$ where each β_i^k is maximizes the entropy given a totally mixed strategy profile σ^k. We thus have

$$\beta_i(y|t,x) = \lim_{k \to \infty} \sum_{t' \in \mathcal{T}} \frac{\pi(y|t',x)\sigma^k(x|t')\pi(t')}{\sum_{t'' \in \mathcal{T}} \sigma^k(x|t'')\pi(t'')} = \lim_{k \to \infty} \sum_{t \in \mathcal{T}} w(t, x; \{\sigma^k\})\pi(y|t,x),$$

which is the desired result. ■

A.9 Proof of Proposition 3

I first establish the following lemma: in a self-confirming equilibrium, every player's belief has the correct support on the set of terminal histories.

Lemma 4. *In a self-confirming equilibrium (σ, β, μ), every player's belief β_i satisfies*

$$\big\{\omega \in \Omega \bigm| p(\sigma_i, \beta_i)(\omega) > 0\big\} = \big\{\omega \in \Omega \bigm| p(\sigma_i, (\sigma_{-i}, \pi))(\omega) > 0\big\}.$$

Proof. Let (σ, β, μ) be a self-confirming equilibrium, and let a terminal history $\omega \in \Omega$ be given. Suppose $p(\sigma_i, (\sigma_{-i}, \pi))(\omega) > 0$. Then every history preceding ω is reached with positive probability, so the belief satisfies $\beta_i(a|I) = \sigma_{-i}(a|I)$ or $\beta_i(a|I) = \pi(a|I)$ on every information set I containing a history that precedes ω for every action $a \in A(I)$. This implies that $p(\sigma_i, \beta_i)(\omega) > 0$. Conversely, suppose $p(\sigma_i, (\sigma_{-i}, \pi))(\omega) = 0$. There exists a predecessor h of ω such that $p(\sigma_i, (\sigma_{-i}, \pi))(h) > 0$ and $p(\sigma_i, (\sigma_{-i}, \pi))(h') = 0$, where h' is an immediate successor of h and a predecessor of ω. This implies that $p(\sigma_i, \beta_i)(h') = 0$. Therefore, $p(\sigma_i, \beta_i)(\omega) > 0$ if and only if $p(\sigma_i, (\sigma_{-i}, \pi))$, completing the proof of Lemma 4.

I now prove Proposition 3. To prove the "if" part of the first statement, suppose an assessment (σ, β, μ) is a self-confirming equilibrium. Suppose that (σ, β, μ) is not an OE. Then there exists a player i and a terminal history $\omega \in \Omega$ such that

$$p(\sigma_i, \beta_i)(\omega) \neq p(\sigma_i, (\sigma_{-i}, \pi))(\omega),$$

where, by Lemma 4, $p(\sigma_i, (\sigma_{-i}, \pi))(\omega) > 0$. This implies that $\beta_i(a|I) \neq \sigma_{\iota(I)}(a|I)$ and $\beta_i(a|I) \neq \pi(a|I)$ for some information set I and action $a \in A(I)$ that contains a predecessor of ω. Since the terminal history ω is reached with positive probability, the information set I is reached with positive probability. This contradicts that (σ, β, μ) is a self-confirming equilibrium. Therefore, (σ, β, μ) is an OE if it is a self-confirming equilibrium and every player's belief β_i has the correct terminal support.

Conversely, to prove the "only if" part of the first statement, suppose an assessment (σ, β, μ) is an OE. From the definition of observational consistency, every player's belief β_i satisfies

$$p(\sigma_i, \beta_i)(\omega) = p(\sigma_i, (\sigma_{-i}, \pi))(\omega)$$

for every terminal history $\omega \in \Omega$. This implies that every player's belief has the correct terminal support. Suppose (σ, β, μ) is not a self-confirming equilibrium. Then there exists a belief β_i such that $\beta_i(a|I) \neq \sigma_{\iota(I)}(a|I)$ or $\beta_i(a|I) \neq \pi(a|I)$ for some information set I reached with positive probability and action $a \in A(I)$. That is, for all histories $h \in I$,

$$p(\sigma_i, \beta_i)(h, a|h) \neq p(\sigma_i, (\sigma_{-i}, \pi))(h, a|h).$$

Let $\Omega(h, a)$ denote the set of terminal histories that are successors of the history (h, a). The above inequality implies that

$$\sum_{\omega \in \Omega(h,a)} p(\sigma_i, \beta_i)(\omega) \neq \sum_{\omega \in \Omega(h,a)} p(\sigma_i, (\sigma_{-i}, \pi))(\omega).$$

This contradicts the fact that $p(\sigma_i, \beta_i)(\omega) = p(\sigma_i, (\sigma_{-i}, \pi))(\omega)$. Therefore, (σ, β, μ) is an OE only if it is a self-confirming equilibrium.

To prove "if" part of the second statement, suppose (σ, β, μ) is a perfect Bayesian equilibrium. By definition, β is correct everywhere; that is, $\beta = (\sigma_{-i}, \pi)$. Let $\{\sigma^k\}_{k=1}^{\infty}$ be a sequence of totally mixed strategy profiles that converge to σ. Let $\{\beta^k\}_{k=1}^{\infty}$ be the sequence of belief profiles such that each β_i^k maximizes the entropy given σ^k. Because every terminal history is reached with positive probability given the totally mixed strategy profile σ^k, the observational consistency constraint implies that $\beta_i^k = (\sigma_{-i}^k, \pi)$ for every player i. Since $(\sigma_{-i}^k, \pi) \to (\sigma_{-i}, \pi)$, every belief β_i^k converges to $(\sigma_{-i}, \pi) = \beta$. Therefore, (σ, β, μ) is a MOE.

Conversely, to prove the "only if" part of the second statement, suppose (σ, β, μ) is a MOE. By definition, there exists a sequence $\{\sigma^k, \beta^k\}_{k=1}^{\infty}$ converging to (σ, β) where each σ^k is a totally mixed strategy profiles and each β_i^k maximizes the entropy given σ^k. Because every terminal history is reached with positive probability given σ^k, the observational consistency constraint implies that $\beta_i^k = (\sigma_{-i}^k, \pi)$ for every player i. Since $\beta_i^k \to \beta$ and $(\sigma_{-i}^k, \pi) \to (\sigma_{-i}, \pi)$, it follows that $\beta_i = (\sigma_{-i}, \pi)$ for every player i; that is, β_i is correct everywhere. Therefore, (σ, β, μ) is a perfect Bayesian equilibrium). ■

A.10 Proof of Proposition 4

Given player i's assessment (σ_i, β_i), let $p_j(\sigma_i, \beta_i)(a)$ denote the marginal probability of player j taking an action $a \in A_j$. That is,

$$p_j(\sigma_i, \beta_i)(a) = \sum_{a_j = a,\, (a_1, \ldots, a_n) \in \Omega} p(\sigma_i, \beta_i)(a_1, a_2, \ldots, a_j, \ldots, a_n).$$

Suppose (σ, β) is an ABEE of (Γ, α). Suppose first that σ is a totally mixed strategy profile. By definition, each player's belief β_i satisfies, for every history h assigned to a different player j and action $a \in A(h)$,

$$\beta_i(a|h) = \sum_{h' \in \alpha_i(h)} \sigma_j(a|h') \frac{p(\sigma_i, \beta_i)(h')}{\sum_{\tilde{h} \in \alpha_i(h')} p(\sigma_i, \beta_i)(\tilde{h})} = p_j(\sigma_i, \sigma_{-i})(a). \tag{6}$$

Suppose σ is not totally mixed. Take a sequence $\{\sigma^k\}$ of totally mixed strategy profiles that converges to σ. For every σ^k, its ABBE is given by β_i^k that satisfies $\beta_i^k(a|h) = p_j(\sigma^k)(a)$, for every history h assigned to a different player j and action $a \in A(h)$. Thus, β_i^k converges to β_i as given by (6).

Suppose (σ', β') is a MOE of (Γ, C). And Suppose the strategy profile σ is totally mixed and is arbitrarily close to σ'. For every player i, his belief β_i satisfies

$$\beta_i \in \operatorname*{argmax}_{\beta_i' \in \mathcal{B}_i} \sum_{\omega \in \Omega} -p(\sigma_i, \beta_i')(\omega) \log(p(\sigma_i, \beta_i')(\omega))$$

subject to

$$p_j(\sigma_i, \beta_i')(a) = p_j(\sigma_i, \sigma_{-i})(a) \quad \text{for all } j \neq i \text{ and } a \in A_j.$$

The necessary and sufficient condition for β_i is that for all $j \neq i$, $a \in A_j$, $\beta_i(a|h)$ satisfies the first-order condition

$$\sum_{\omega \in \Omega(h,a),\, \iota(h)=j} -p(\sigma_i, \beta_i)(\omega) \cdot \log\left(p(\sigma_i, \beta_i)(\omega)\right) = \sum_{\omega \in \Omega(h,a),\, \iota(h)=j} p(\sigma_i, \beta_i)(\omega) \cdot (\lambda_j(a) + 1),$$

where $\lambda_j(a)$ are Lagrangian multipliers. This condition is satisfied if $\beta_i(a|h) = p_{\iota(h)}(\sigma_i, \sigma_{-i})(a)$ for all h and $a \in A(h)$. Because this condition is also sufficient, this implies that every player's belief in an MOE satisfies $\beta_i(a|h) = p_{\iota(h)}(\sigma_i, \sigma_{-i})(a)$.

Since the requirements on players' strategies and beliefs are the same for ABEE and MOE, a pair (σ, β) is an ABBE if and only if it is MOE. ■

A.11 Proof of Claim 7

Suppose there exists a MOE (σ^*, β^*) such that Alice Passes with positive probability at the beginning. Let σ denote a totaly mixed strategy profile arbitrarily close to σ^*. Let β denote the maximum-entropy belief given σ^*. Since (σ^*, β^*) is a MOE, $\beta \longrightarrow \beta^*$ as $\sigma \longrightarrow \sigma^*$.

For easier notation, let the strategies be represented by

$$a = \sigma_{\text{Alice}}(P),\ b = \sigma_{\text{Bob}}(P|P),\ c = \sigma_{\text{Alice}}(P|PP),\ \text{ and } d = \sigma_{\text{Bob}}(P|PPP).$$

Similarly, let the beliefs be represented by

$$\hat{a} = \beta_{\text{Bob}}(P),\ \hat{b} = \beta_{\text{Alice}}(P|P),\ \hat{c} = \beta_{\text{Bob}}(P|PP),\ \text{ and } \hat{d} = \beta_{\text{Alice}}(P|PPP).$$

By definition, Alice's belief $(\hat{b}, \hat{d})$ satisfies

$$\begin{aligned}(\hat{b}, \hat{d}) \in \underset{(\tilde{b},\tilde{d})\in[0,1]^2}{\operatorname{argmax}}\ & -(1-a)\log(1-a) - a(1-\tilde{b})\log(a(1-\tilde{b})) \\ & - a\tilde{b}(1-c)\log(a\tilde{b}(1-c)) - a\tilde{b}c(1-\tilde{d})\log(a\tilde{b}c(1-\tilde{d})) - a\tilde{b}c\tilde{d}\log(a\tilde{b}c\tilde{d})\end{aligned}$$

subject to

$$a(1-\tilde{b}) + 2a\tilde{b}(1-c) + 3a\tilde{b}c(1-\tilde{d}) + 4a\tilde{b}c\tilde{d} = a(1-b) + 2ab(1-c) + 3abc(1-d) + 4abcd.$$

The first-order condition with respect to $\tilde{d}$ yields

$$\log\frac{1-\hat{d}}{\hat{d}} = \lambda,$$

where λ is the Lagrange multiplier for the constraint. Observe that as $a \longrightarrow 0$, both sides of the constraint approaches 0, so $\lambda \longrightarrow 0$. Thus, $\hat{d} \longrightarrow \frac{1}{2}$.

Given this belief $\hat{d} = \frac{1}{2}$, Alice's strategy is sequentially rational only if $c = 0$. Given this strategy, Alice's belief $\hat{c}$ approaches $\frac{1}{2}$ as a approaches 0 because of the first-order condition as the above. Given this belief, Alice's strategy is sequentially rational only if $a = 0$, a contradiction. ■

A.12 Proof of Proposition 5

Suppose (σ, β) is a MOE, and suppose σ is totally mixed. Observational consistency requires that for all $(\theta, s, a, s') \in \Omega$,

$$\beta_i(s'|\theta, s, a)\beta_i(a_{-i}|\theta, s)\sigma_i(a_i|\theta, s)p(\sigma_i, \beta_i)(\theta, s) = p(\sigma_i, (\sigma_{-i}, \pi))(\theta, s, a, s').$$

This implies that

$$\beta_i(s'|\theta,a,s) = \frac{p(\sigma_i,(\sigma_{-i},\pi))(\theta,s,a,s')}{\sum_{s,a,s'} p(\sigma_i,(\sigma_{-i},\pi))(\theta,s,a,s')} = \pi(s'|\theta,a,s), \text{ and}$$

$$\beta_i(a_{-i}|\theta,s) = \frac{\sum_{a_i,s'} p(\sigma_i,(\sigma_{-i},\pi))(\theta,s,(a_i,a_{-i}),s')}{\sum_{a,s'} p(\sigma_i,(\sigma_{-i},\pi))(\theta,s,a,s')} = \sigma_{-i}(a_{-i}|\theta,s),$$

thus β_i is everywhere correct. If σ is not totally mixed, let $\{\sigma^k,\beta^k\}_{k=1}^{\infty}$ be a sequence such that each σ^k is totally mixed strategy profiles converging to σ and each β_i^k maximizes the entropy given σ^k. Since each β_i^k is correct everywhere, each sequence $\{\beta_i^k\}$ converges to (σ_{-i},π). Therefore, in every MOE (σ,β), each belief β_i is correct everywhere. This means that the condition for a MOE and that for Markof perfect equilibrium are equivalent. ■

A.13 Proof of Proposition 6

Suppose an assessment (σ,β) is a MOE. Suppose σ is totally mixed. Observational consistency implies that $\beta_i(a_{-i}|\theta,s) = \sigma_{-i}(a_{-i}|\theta,s)$ and $p(\sigma_i,\beta_i)(\theta,s) = p(\sigma,\pi)(\theta,s)$ for all θ,s,a_{-i}. From the definition of MOE, the belief β_i maximizes

$$\sum_{\theta,s,a,s'} -\beta_i(s'|\theta,s,a)\sigma(a|\theta,s)p(\sigma,\pi)(\theta,s)\cdot\log\left(\beta_i(s'|\theta,s,a)\sigma(a|\theta,s)p(\sigma,\pi)(\theta,s)\right)$$

subject to

$$\sum_{s'} \beta_i(s'|\theta,s,a)\sigma(a|\theta,s)p(\sigma,\pi)(\theta,s) = \sum_{s'} \pi(s'|\theta,s,a)\sigma(a|\theta,s)p(\sigma,\pi)(\theta,s), \text{ and}$$

$$\sum_{\theta} \beta_i(s'|\theta,s,a)\sigma(a|\theta,s)p(\sigma,\pi)(\theta,s) = \sum_{\theta} \pi(s'|\theta,s,a)\sigma(a|\theta,s)p(\sigma,\pi)(\theta,s).$$

The first-order necessary and sufficient conditions are

$$-\log(\beta_i(s'|\theta,s,a)\sigma(a|\theta,s)p(\sigma,\pi)(\theta,s)) - 1 = \lambda(\theta,s,a) + \gamma(s,a,s'),$$

for all i,θ,s,a, and s', where $\lambda(\theta,s,a)$ and $\gamma(s,a,s')$ are the Lagrange multipliers for the two constraints, respectively. These conditions imply that

$$\begin{aligned}&\log(\beta_i(s'|\theta,s,a)\sigma(a|\theta,s)p(\sigma,\pi)(\theta,s)) + \log(\beta_i(\bar{s}'|\theta,\bar{s},a)\sigma(a|\bar{\theta},s)p(\sigma,\pi)(\bar{\theta},\bar{s}))\\ =&\log(\beta_i(s'|\bar{\theta},s,a)\sigma(a|\bar{\theta},s)p(\sigma,\pi)(\bar{\theta},s)) + \log(\beta_i(\bar{s}'|\theta,s,a)\sigma(a|\theta,s)p(\sigma,\pi)(\theta,s)),\end{aligned}$$

for all $i, \theta, \bar{\theta}, s, a, s'$, and $\bar{s}'$. It follows that

$$\beta_i(s'|\theta, s, a) = \beta_i(s'|\bar{\theta}, s, a) \quad \text{for all } i, \theta, \bar{\theta}, s, a, \text{ and } s'.$$

This equation and the observational consistency imply that

$$\beta_i(s'|\theta, s, a) = \sum_{\theta' \in \Theta} \pi(s'|\theta', s, a) w(\theta', s, a) \quad \text{for all } \theta, s, a, s',$$

where $w(\cdot)$ is a weight function such that

$$w(\theta', s, a) = \frac{\sigma(a|\theta', s) \cdot p(\sigma, \pi)(\theta', s)}{\sum_\theta \sigma(a|\theta, s) \cdot p(\sigma, \pi)(\theta, s)}.$$

Suppose σ is not totally mixed. There exists a sequence $\{\sigma^k, \beta^k\}$ converging to (σ, β) where each σ^k is a totally mixed strategy profile and each β_i^k maximizes the entropy given σ^k. We know that every player's belief β_i^k satisfies

$$\begin{aligned}
&\beta_i^k(a_{-i}|\theta, s) = \sigma_{-i}^k(a_{-i}|\theta, s) \text{ and} \\
&\beta_i^k(s'|\theta, s, a) = \sum_{\theta' \in \Theta} \pi(s'|\theta', s, a) \cdot \frac{\sigma^k(a|\theta', s) \cdot p(\sigma^k, \pi)(\theta', s)}{\sum_\theta \sigma^k(a|\theta, s) \cdot p(\sigma^k, \pi)(\theta, s)}.
\end{aligned}$$

Because $\{\sigma^k, \beta^k\}$ converges to (σ, β), we have

$$\begin{aligned}
&\beta_i(a_{-i}|\theta, s) = \sigma_{-i}(a_{-i}|\theta, s) \text{ and} \\
&\beta_i(s'|\theta, s, a) = \sum_{\theta' \in \Theta} \pi(s'|\theta', s, a) \cdot \lim_{k \to \infty} \frac{\sigma^k(a|\theta', s) \cdot p(\sigma^k, \pi)(\theta', s)}{\sum_\theta \sigma^k(a|\theta, s) \cdot p(\sigma^k, \pi)(\theta, s)}.
\end{aligned}$$

Conversely, if an OE (σ, β) satisfies the above, it is a MOE. Therefore, an OE is a MOE if each player's belief satisfies the above conditions. ■

B Causality: an axiomatic foundation

In the main text, I appeal to common intuitions about causality and used a definition of causation loosely as changes in probabilities. In the Smoker's problem of Section 2, smoking causes cancer in the sense that it increases the probability of cancer. In the Manager-Worker game of Section 3, Manager has a causal effect on Worker's bonus in the sense that her actions change the probability distribution of Worker's bonus. Causal misperception has loosely meant a wrong belief about these probabilities. In this section, I provide an axiomatic foundation for a precise definition.

B.1 Probabilistic characterization

Consider a finite set $\mathcal{A}$ of *elementary acts* or simply *acts*, which are natural phenomena or actions by agents. Note that an act refers an abstract element rather than a set of outcomes. The *act set* $\mathcal{A}$ has $2n$ elements: for every act $x \in \mathcal{A}$ (*e.g.*, it rains), its *negation* $\neg x$ (*e.g.*, it does not rain) also belongs to $\mathcal{A}$. The negation of a negation is the original act, *i.e.*, $\neg(\neg x) = x$. For example, the set $\mathcal{A}$ may be $\{x, \neg x, y, \neg y, z, \neg z\}$, where $x =$ "it rains", $y =$ "tennis court is wet", and $z =$ "one plays tennis".

A *causal relation* is a binary relation on the set of acts $\mathcal{A}$. Namely, a causal relation R specifies *causation* between acts: whether an act x *causes* another act y, alternatively written as $(x, y) \in R$. If x causes y, x is a *cause* of y and y is an *effect* of x. For example, a causal relation may specify that a rain causes a wet tennis court (*i.e.*, $(x, y) \in R$) and that a wet tennis court causes one to not play tennis (*i.e.*, $(y, \neg z) \in R$).

I examine the following *axioms of causation.* Let $x, y, x_1, x_2, \dots$ be any acts in $\mathcal{A}$.

Axiom 1 (Exclusivity)**.** If $(x, y) \in R$, then $(x, \neg y) \notin R$.

Axiom 2 (Counterfactuality)**.** If $(x, y) \in R$, then $(\neg x, \neg y) \in R$.

Axiom 3 (Acyclicality)**.** If $(x_1, x_2), (x_2, x_3), \dots, (x_{k_1}, x_k) \in R$, then $(x_k, x_1) \notin R$.

These axioms capture essential aspects of causation as commonly understood in everyday life. Since a rainfall causes a wet tennis court, it does not cause a dry court (Axiom 1). Not having a rainfall is a cause—perhaps as a necessary condition—of a dry court (Axiom 2). Causation does not run backward in a cycle; if rain causes a wet tennis court and a wet tennis court causes one to not play tennis, not playing tennis cannot be a cause of the rain (Axiom 3).

To arrive at a characterization of the Axioms 1–3, I define an *extensive-form structure of acts* or simply an *act structure.* An *act history* or *history* is a finite sequence $h = (x_1, x_2, \dots, x_k)$ of acts that contains each act or its negation at most once. An empty history $h_0 = (\ )$ is called a *root.* A history h is *terminal* if it contains each act or its negation exactly once. An act history h is a *predecessor* of an act history h' if $h' = (h, x_{k+1}, \dots, x_{k+\ell})$; then h' is a *successor* of h. If $\ell = 1$, then h is an *immediate predecessor* of h' and h' is an *immediate successor* of h.

Definition 18 (Act structure)**.** An *act structure* is a pair (H, p) consisting of

1. An *act tree* H: a set of histories such that

    (a) for each terminal history $h \in H$, all of its predecessors are in H,

    (b) for each non-terminal history $h \in H$, there are at most two immediate successors (h, x) or $(h, \neg x)$ in H, and

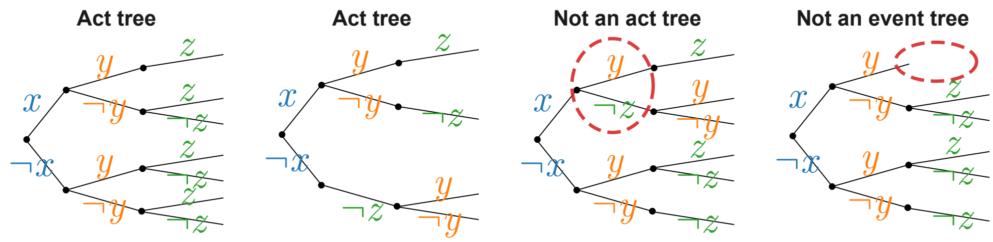


Figure 15: Examples and nonexamples of act trees

(c) for each non-terminal history $h \in H$, there is a terminal successor $\omega \in H$.

2. An *act probability function* p: a function such that, for every non-terminal history $h \in H$ and its immediate successor $h' \in H$, $p(h'|h)$ is the conditional probability of h' given h.

To clarify the definition of an act tree, I show examples and nonexamples of act trees in Figure 15. The act set is $\mathcal{A} = \{x, \neg x, y, \neg y, z, \neg z\}$. The first and second trees are act trees, because every non-terminal history h has at most two immediate successors (h, w) or $(h, \neg w)$ for some act w and has a terminal successor. The third tree is not an act tree because the history x has two immediate successors whose succeeding acts are y and $\neg z$. The fourth tree is not an act tree because the history (x, y) does not have a terminal successor.

I only consider act structures with full support: $p(h'|h) > 0$ for every successor h' of each history $h \in H$. For any act structure (H, p), I extend p to define conditional probabilities: for every history $h_j \in H$ and its successor $h_J \in H$,

$$p(h_J|h_j) = p(h_J|h_{J-1})p(h_{J-1}|h_{J-2}) \dots p(h_{j+2}|h_{j+1})p(h_{j+1}|h_j),$$

where each h_{k+1} is an immediate successor of h_k. I write the unconditional probability of any history $h \in H$ as $p(h) = p(h|h_0)$ where $h_0 = (\ )$ is the root.

I further extend p to define conditional probabilities of an act given a history. Let Ω denote the set of terminal histories in H. Let $\Omega(h)$ denote the set of terminal successors of $h \in H$. Let $\Omega(x)$ denote the set of terminal histories that contain the act $x \in \mathcal{A}$. For every act y and history h, define the conditional probability

$$p(y|h) = \frac{\sum_{h' \in \Omega(h) \cap \Omega(y)} p(h')}{p(h)}.$$

Note with this definition that if an act y is already in history h, the conditional probability of y given h is $p(y|h) = 1$ and that of $\neg y$ given h is $p(\neg y|h) = 0$.

We are ready to define what it means for an act structure to represent a causal relation.

Definition 19 (Act-structure representation of causation)**.** An act structure (H, p) *represents* a causal relation R if, for all acts $x, y \in \mathcal{A}$, $(x, y) \in R$ is equivalent to the following: (a) there exists a pair of immediate successors $((h, x), (h, \neg x))$ for some $h \in H$, and (b) for every such pair

$$0 < p(y|h, x) - p(y|h, \neg x) < 1.$$

Roughly speaking, an act structure represents a causal relation if the statement "x cause y" is equivalent to the statement "x strictly raises the probability of y relative to $\neg x$ whenever both x and $\neg x$ are feasible", or x is a *universal probability increaser* of y. If such act structure exists, I say that the causal relation R has an act-structure representation.

The following result provides a logical foundation for a probabilistic interpretation of causality.

Proposition 7. *A causal relation R satisfies the axioms of causation if and only if R has an act-structure representation.*

Both the "if" and "only if" parts of this statement are interesting. The former means that if there is a probability distribution over an act tree where we define causation as a unversal increase in probabilities, this definition satisfies Axioms 1–3. The latter means that if a causal relation on a set of acts satisfies the axioms, one can find an act tree and an associated probability distribution to represent that causal relation.

It is easier to see why "if" part is true. Suppose a causal relation R is represented by an act structure (H, p). This means that if $(x, y) \in R$ for some two acts x and y, the act x rather than $\neg x$ increases the probability of y and decreases the probability of $\neg y$, hence $(x, \neg y) \notin R$ as required by Axiom 1 (exclusivity). On the flip side, the act $\neg x$ rather than x decreases the probability of y and increases the probabilty of $\neg y$, hence $(\neg x, \neg y) \in R$ as required by Axiom 2 (counterfactuality). Lastly, if $(x_1, x_2), (x_2, x_3), \ldots, (x_{k-1}, x_k) \in R$, the act x_k or $\neg x_k$ occurs later everywhere in the act tree than the act x_1, so x_k cannot raise the probability of x_1; thus, $(x_k, x_1) \notin R$ as required by Axiom 3 (acyclicality).

The "only if" part is less straightforward: Given a causal relation satisfying the axioms, how do we find an act structure (H, p) that represents it? Although I include the general proof in subsection B.2, it is informative to walk through the key steps in a simple setting. Consider an act set $\mathcal{A} = \{x, \neg x, y, \neg y, z, \neg z\}$ and a causal relation R on $\mathcal{A}$ satisfying Axioms 1–3:

$$R = \{(x, y), (\neg x, \neg y), (y, \neg z), (\neg y, z), (x, \neg z), (\neg x, z)\}.$$

For example, a rain (x) causes the tennis court to be wet (y), a wet tennis court causes one to not play tennis ($\neg z$), and a rain (x) causes one to not play tennis ($\neg z$). Counterfactually,

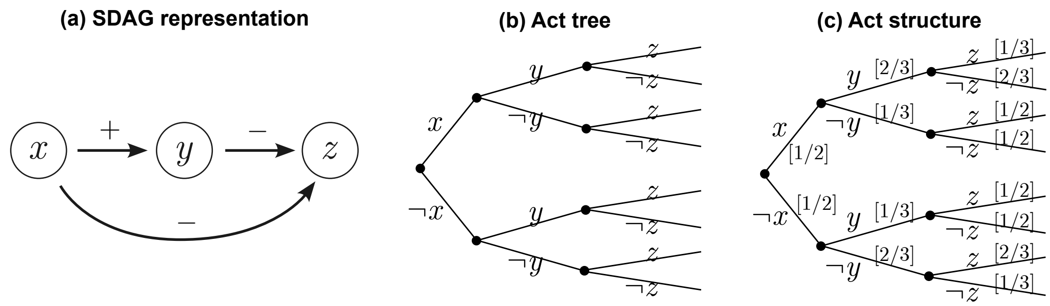


Figure 16: Constructing an act-structure representation from a causal relation

an absence of rain ($\neg x$) causes the tennis court to be dry ($\neg y$), a dry tennis court ($\neg y$) causes one to play tennis (z), and an absence of rain ($\neg x$) causes one to play tennis (z).

The first step in finding a act-structure representation of R is to form an act tree. An alternative way to express R is to say that x has a positive effect on y , y has a negative effect on z, and x has a negative effect on z. In short, the relation R is represented by a *signed directed acylcic graph* (SDAG) with $x \xrightarrow{+} y$, $y \xrightarrow{-} z$, and $x \xrightarrow{-} z$ as in Figure 16(a).[1] Once we have a SDAG representation, we can order the acts x, y, and z by their *depths* or the maximal number of edges from the root of the graph. The depths of x, y, and z are 0, 1, and 2, respectively. We can then form a *complete and totally ordered* act tree H where acts always occur in the order (1) x or $\neg x$, (2) y or $\neg y$, and (3) z or $\neg z$, as in Figure 16(b).

Now that we have the act tree H, the second step is to find an associated probability function p. Let $\mathbf{e} = (e_1, e_2, e_3, e_4) \in [0,1]^4$ and let p be an unknown probability function on the act tree H. Consider the system of functional equations

$$\begin{aligned} p(y|x) - p(y|\neg x) &= e_1, \\ p(z|x,y) - p(z|x,\neg y) &= -e_2, \\ p(z|\neg x,y) - p(z|\neg x,\neg y) &= -e_3, \text{ and} \\ p(z|x) - p(z|\neg x) &= -e_4. \end{aligned}$$

In other words, in the act structure (H,p), the effect of x on y is e_1, the effect of y on z given x is $-e_2$, the effect of y on z given $\neg x$ is $-e_3$, and the effect of x on z is $-e_4$.

Observe that all conditional probabilities can be rewritten as a continuous function of a probability vector $\mathbf{p}$ over the set terminal histories $\Omega \subset H$. Thus, we can write the above system of equation as $\mathbf{f}(\mathbf{p}, \mathbf{e}) = 0$ for some continuous vector-valued function $\mathbf{f}$. Consider

[1] Formally, I say that a SDAG *represent* a causal relation R if $(x,y) \in R$ is equivalent to $x \xrightarrow{+} y$, $x \xrightarrow{-} \neg y$, $\neg x \xrightarrow{-} y$, or $\neg x \xrightarrow{+} \neg y$ (Definition 21). In the Appendix, I show that every causal relation satisfying the axioms of causation has such a *SDAG representation* (Lemma 5).

the correspondence $G : [0,1]^4 \twoheadrightarrow \Delta(\Omega)$ such that

$$G(\mathbf{e}) = \{\mathbf{p} \in \operatorname{int} \Delta(\Omega) : \mathbf{f}(\mathbf{p}, \mathbf{e}) = \mathbf{0}\},$$

which is continuous because $\mathbf{f}$ is continuous. Let $\bar{\mathbf{p}} = (1/8, 1/8, \ldots, 1/8)$ be the uniform probability distribution on Ω, meaning that every conditional probability in the above system of equations equals one half. This uniform probability distribution results in zero effect between any two acts, thus $\mathbf{p} \in G(\mathbf{0})$. Because the correspondence G is lower hemicontinuous at $\mathbf{0}$, there exists a sequence $\{(\mathbf{p}^k, \mathbf{e}^k)\}_{k=1}^{\infty}$ in $(\operatorname{int} \Delta(\Omega)) \times (0,1)$ that converges to $(\bar{\mathbf{p}}, \mathbf{0})$ and $\mathbf{p}^k \in G(\mathbf{e}^k)$ for all k. Then by construction, all of the act structures (H, p^k) represent the causal relation R. Note, however, by this argument, that there exist infinitely many act structures (H, p^k) that represent the same causal relation.

Figure 16(c) shows a concrete example of an act-structure representation of the causal relation R. In this example, the effect of x on y is positive ($p(y|x) - p(y|\neg x) = 1/3$), the effect of y on z given x is negative ($p(z|x, y) - p(z|x, \neg y) = -1/6$), the effect of y on z given $\neg x$ is negative ($p(z|\neg x, y) - p(z|\neg, \neg y) = -1/6$), and the effect of x on z is negative ($p(z|x) - p(z|\neg x) = -2/9$). Therefore, this act structure represents the causal relation R.

B.2 Proof of Proposition 7

Definition 20. A *directed acyclic graph* (DAG) is a directed graph that has no directed cycles. A *signed directed acyclic graph* (signed DAG or SDAG) is a DAG that has a positive or negative sign (but not both) for each of its edges.

We can write an SDAG as the triple $(\mathcal{V}, E^-, E^+)$, where $\mathcal{V}$ is the set of nodes (vertices), E^- is the set of minus-signed edges, and E^+ is the set of plus-signed edges. That is, E^- is the set of ordered pairs of vertices $(x, y) \in \mathcal{V}^2$ that have minus-signed directed edges from x to y. Similarly, E^+ is the set of ordered pairs of vertices $(x, y) \in \mathcal{V}^2$ that have plus-signed directed edges from x to y.

A *basis* of an act set $\mathcal{A}$ is any subset $\mathcal{B} \subset \mathcal{A}$ such that $x \in \mathcal{B}$ if and only if $\neg x \notin \mathcal{B}$.

Definition 21. A SDAG *represents* a causal relation R on an act set $\mathcal{A}$ if

1. the set of its nodes is a basis of $\mathcal{A}$, and
2. for all $x, y \in \mathcal{A}$, $(x, y) \in R$ is equivalent to the following:

$$x \xrightarrow{+} y, \quad x \xrightarrow{-} \neg y, \quad \neg x \xrightarrow{-} y, \quad \text{or} \quad \neg x \xrightarrow{+} \neg y.$$

Lemma 5. *If a causal relation satisfies the axioms of causation, it has a SDAG representation.*

Proof. Suppose a causal relation R on an act set $\mathcal{A}$ satisfies the axioms of causation. Consider a triple $(\mathcal{B}, E^-, E^+)$ where $\mathcal{B}$ is a basis of $\mathcal{A}$, $E^- = \{(x,y) \in \mathcal{B}^2 : (x, \neg y) \in R\}$, and $E^+ = \{(x,y) \in \mathcal{B}^2 : (x,y) \in R\}$. By Axiom 1, E^- and E^+ are mutually exclusive. By Axiom 2, $(x,y) \in E$ implies $(\neg x, \neg y) \in E$, for both $E \in \{E^-, E^+\}$. By Axiom 3, there does not exist a cycle $x_1, x_2, \ldots, x_k$ such that

$$(x_1, x_2), (x_2, x_3), \ldots, (x_{k-1}, x_k), (x_k, x_1) \in E^- \cup E^+.$$

This implies that the triple $(\mathcal{B}, E^-, E^+)$ is an SDAG that represents the causal relation R, completing the proof of Lemma 5.

Lemma 6. *If a causal relation has an SDAG representation, it has an act-structure representation.*

Proof. Suppose a SDAG $(\mathcal{B}, E^-, E^+)$ represents a causal relation R on $\mathcal{A}$ with $2n$ elements. For any node $x \in \mathcal{B}$, let $\delta(x)$ denote the depth of x, that is, the maximum length of a directed path between x and the SDAG's root node. Without loss of generality, suppose $\mathcal{B} = \{x_1, x_2, \ldots, x_n\}$ and

$$\delta(x_1) \leq \delta(x_2) \leq \cdots \leq \delta(x_n).$$

Ronstruct an act structure on $\mathcal{A}$ as follows. Let H denote the *complete and totally ordered* act tree on $\mathcal{A}$ with the above order: every non-terminal history $h \in H \backslash \Omega$ has two immediate successors, and every terminal history $\omega \in \Omega$ is a sequence of acts $(\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n)$ where $\tilde{x}_1 \in \{x_1, \neg x_1\}$, $\tilde{x}_2 \in \{x_2, \neg x_2\}$, $\ldots$, and $\tilde{x}_n \in \{x_n, \neg x_n\}$.

Let p denote an act probability function on the act tree H. For every i, j, h such that $i < j$ and $(h, x_j), (h, \neg x_j) \in H$, Suppose $e_{ijh} \in [0,1]$, and consider the equations

$$p(x_j|h, x_i) - p(x_j|h, \neg x_i) = \begin{cases} -e_{ijh} & \text{if } (x_i, x_j) \in E^-, \\ e_{ijh} & \text{if } (x_i, x_j) \in E^+, \\ 0 & \text{if } (x_i, x_j) \in \mathcal{B} \backslash (E^- \cup E^+) \end{cases}$$

for all h such that $(h, x_i) \in H$ and $(h, \neg x_i) \in H$. Equivalently, this system of equations can be written as $\mathbf{f}(\mathbf{p}, \mathbf{e}) = \mathbf{0}$ for some continuous vector-valued function $\mathbf{f} : \operatorname{int} \Delta(\Omega) \times [0,1]^L \to [-1,1]^{|\Omega|}$ and the vector $\mathbf{e}$ consisting of each e_{ijh}.

Let $G : [0,1]^L \twoheadrightarrow \operatorname{int} \Delta(\Omega)$ be defined as the correspondence such that $G(\mathbf{e}) = \{\mathbf{p} \in \operatorname{int} \Delta(\Omega) : \mathbf{f}(\mathbf{p}, \mathbf{e}) = \mathbf{0}\}$. Since the vector-valued function f is continuous, the correspondence G is continuous. In particular, G is lower hemicontinuous at 0. Furthermore, the uniform probability distribution $\bar{\mathbf{p}} = (1/|\Omega|, 1/|\Omega|, \ldots, 1/|\Omega|)$ is an element of $G(\mathbf{0})$.

Consider a sequence $\{\mathbf{e}^m\}_{m=2}^{\infty}$ where $\mathbf{e}^m = (1/m, 1/m, \ldots, 1/m)$. By the lower hemi-continuity of G, there exists a subsequence $\{\mathbf{e}^{m^k}\}_{k=1}^{\infty}$ and a sequence $\{\mathbf{p}^k\}_{k=1}^{\infty}$ in $\Delta(\Omega)$ that converges to the uniform probability distribution $\bar{\mathbf{p}}$ such that $\mathbf{p}^k \in G(\mathbf{e}^{m^k})$ for all k. Then for all k and for all h such that $(h, x_i) \in H$ and $(h, \neg x_i) \in H$, we have

$$\begin{aligned}
-1 < p^k(x_j|h, x_i) - p^k(x_j|h, \neg x_i) < 0 \qquad & \text{if } (x_i, x_j) \in E^-, \\
0 < p^k(x_j|h, x_i) - p^k(x_j|h, \neg x_i) < 1 \qquad & \text{if } (x_i, x_j) \in E^+, \text{ and} \\
p^k(x_j|h, x_i) - p^k(x_j|h, \neg x_i) = 0 \qquad & \text{if } (x_i, x_j) \in \mathcal{B}\backslash(E^- \cup E^+).
\end{aligned}$$

Equivalently, $(x, y) \in R$ if and only if $0 < p^k(y|h, x) - p^k(y|h, \neg x) < 1$ for all h such that $(h, x) \in H$ and $(h, \neg x) \in H$. Therefore, the act structures (H, p^k) for all k represent the causal relation R, completing the proof of Lemma 6.

Lemmas 5 and 6 imply that a causal relation satisfies the axioms of causation only if it has an act-structure representation.

Conversely, suppose an act structure (H, p) represents a causal relation R on an act set $\mathcal{A}$. In contrast to the theorem's statement, suppose R does not satisfy all axioms of causation.

Case 1: R violates Axiom 1 (exclusivity). That is, $(x, y) \in R$ and $(x, \neg y) \in R$ for some $x, y \in \mathcal{A}$. Since (H, p) represents R, $(x, y) \in R$ implies that for some history h with immediate succesors (h, x) and $(h, \neg x)$,

$$p(y|h, x) > p(y|h, \neg x).$$

For the same reason, $(x, \neg y) \in R$ implies that

$$p(\neg y|h, x) > p(\neg y|h, \neg x).$$

It follows that $p(y|h, x) = 1 - p(\neg y|h, x) < 1 - p(\neg y|h, \neg x) = p(y|h, \neg x)$, a contradiction.

Case 2: R violates Axiom 2 (counterfactuality). That is, $(x, y) \in R$ and $(\neg x, \neg y) \notin R$ for some $x, y \in \mathcal{A}$. Since (H, p) represents R, $(x, y) \in R$ implies that for some history h with immediate succesors (h, x) and $(h, \neg x)$,

$$0 < p(y|h, x) - p(y|h, \neg x) < 1. \tag{7}$$

Observe that $p(y|h, x) = 1 - p(\neg y|h, \neg(\neg x))$ and $p(y|h, \neg x) = 1 - p(\neg y|h, \neg x)$. Thus, the above is equivalent to

$$0 < p(\neg y|h, \neg x) - p(\neg y|h, \neg(\neg x)) < 1.$$

This implies that $(x \neg y) \in R$, a contradiction.

Case 3: R violates Axiom 3 (acyclicality). That is, $(x_1, x_2) \in R$, $(x_2, x_3) \in R$, $\ldots$, $(x_{k-1}, x_k) \in R$, $(x_k, x_1) \in R$ for some $x_1, \ldots, x_k \in \mathcal{A}$. Since (H, p) represents H, $(x_1, x_2) \in R$ implies that for every history h with immediate succesors (h, x_1) and $(h, \neg x_1)$,

$$p(x_2|h, x_1) > p(x_2|h, \neg x_1).$$

This implies that x_2 and $\neg x_2$ are not in the histories h with immediate succesors (h, x_1) and $(h, \neg x_1)$. Similarly, $(x_2, x_3) \in R$ implies that x_3 and $\neg x_3$ are not in the histories h with immediate successors (h, x_2) and $(h, \neg x_2)$, and so on. It follows that x_k and $\neg x_k$ are not in the histories h with immeidate successors (h, x_{k_1}) and $(h, \neg x_{k_1})$, and x_1 or $\neg x_1$ are in the histories h with immediate successors (h, x_k) and $(h, \neg x_k)$. Then for all h with immediate successors (h, x_k) and $(h, \neg x_k)$,

$$p(x_1|h, x_k) = p(x_1|h, \neg x_k).$$

However, (x_k, x_1) implies $p(x_1|h, x_k) > p(x_1|h, \neg x_k)$, a contradiction.

Cases 1–3 show that any violation of the axioms of causation leads to a contradiction. Therefore, a causal relation satisfies the axioms of causation only if it has an act-structure representation. ∎