Deriving Predictions from Economic Theory: Solution Concepts vs. Epistemic Conditions

Jason Campbell[†] Nicolas Mäder[†]

April 12, 2024

Click here for the latest version.

Abstract

To derive empirical predictions from economic theory, the literature conventionally proceeds by studying, or 'solving', said theory through the lens of a solution concept (e.g. Nash equilibrium). In this paper, we argue that theoretical predictions are more suitably derived from an explicit set of epistemic conditions, even (or especially) if the latter map into an existing solution concept. To illustrate its transparency, versatility, and conceptual cogency, we apply the epistemic approach to various well-known economic frameworks.

Keywords: Nash equilibrium, solution concepts, epistemic game theory, epistemic economic theory

JEL codes: D80, D70, C70

[†]Knauss School of Business, University of San Diego; *jasoncampbell@sandiego.edu*, *nmaeder@sandiego.edu*.

1 Introduction

A key objective of economic theory is to generate empirically relevant predictions; that is, plausible predictions concerning actual, real-life behavior. To this end, the literature conventionally proceeds by studying, or 'solving', a given theory through the lens of a solution concept such as Nash equilibrium.[1] More precisely, solution concepts are used to restrict the set of principally possible outcomes to a corresponding, concept-implied subset. In effect, contemporary convention may be summarily described as follows:

A1. Propose an economic theory

A2. Specify a solution concept (e.g. Nash equilibrium)

A3. Restrict pool of strategy profiles to concept-implied set of 'solutions'

A4. Predict outcome by selecting among set of concept-implied solutions

Evidently, a key premise that underlies Procedure A is that a strategy profile represents a suitable candidate for the ultimately observed outcome (if and) only if it is contained in the set of solutions derived in step A3. For example, a particularly popular approach to derive such candidates is to discard any profile that is not Nash.

The canonical choice of Nash equilibrium in step A2 is typically motivated by the fact that any non-Nash profile is 'self-destabilizing' in that it must feature at least one player who would benefit from deviating.[2] Thus, to the extent that humans are rational, one might naturally expect real-life behavior to form such equilibrium. However, Bernheim (1984) and Pearce (1984) famously show that rationality — even if it is commonly known — is in and of itself insufficient to imply that a game's outcome must be Nash; suggesting that Nash equilibrium may be too restrictive a solution concept. Conversely, there exists an extensive literature on equilibrium 'refinements' which argues that Nash equilibrium is in fact not restrictive enough (see Harsanyi and Selten, 1988). This seeming contradiction can be reconciled by distinguishing between two types of errors.

- ⋆ *Type 1 (too restrictive).* Exclusion of a 'plausible' outcome from set of solutions
- ⋆ *Type 2 (not restrictive enough).* Inclusion of an 'implausible' outcome in set of solutions

Since type-I and type-2 errors are not mutually exclusive, it is easy to see that Nash equilibrium

[1] See Hu and Sobel (2022), Altinoglu and Stiglitz (2023), Budish and Bhave (2023), or Liu and Sun (2023).

[2] "Clearly, [...] the solution of a noncooperative game has to be a Nash equilibrium since every other strategy combination is self-destabilizing" (van Damme, 1991)

(or any other solution concept) can be both too restrictive and not restrictive enough. Indeed, this is not only true across games, but even *within* a particular game.

Following the above logic, a tempting way to operationalize Procedure A would be to select a solution concept so as to minimize some sort of weighted sum of type-1 and type-2 of errors. Such an approach is problematic for at least two reasons. First, what constitutes a 'plausible' prediction is precisely what a theorist wishes to derive from their theory and, as such, should not be invoked primitively.[3] Second, and this is the starting point of the present paper, it is well-known (in epistemic game theory) that choosing a solution concept effectively amounts to making assumptions about the modeled players' rationality, their knowledge, and/or the assumptions that they might make when determining their strategy (see Figure 2). For example, Nash equilibrium reflects an epistemic state, whereby each agent knows their own payoffs, acts rationally, and correctly anticipates every other player's actual/final choice. Thus, Nash equilibrium should be invoked if (and only if) there is reason to believe that all real-world counterparts of a particular model's theoretical agents do, in fact, fulfill the listed requirements.[4]

Following the above logic, our primary contribution lies in the translation of existing insights from epistemic game theory into an operational paradigm (akin to Procedure A) for predicting the plausible outcomes of a particular economic theory. To this end, we first collect various results from the epistemic literature in the form of a mapping from epistemic states into well-known solution concepts (Section 2). In turn, we illustrate the relative merits of the epistemic approach by applying it to a series of well-known economic models (Section 3). Finally, in Section 4, we take stock and distill the main insights from our analysis into the following two alternatives to Procedure A,

B1/C1. Propose an economic theory

B2/C2. Specify a suitable epistemic state capturing the specific strategic environment at hand

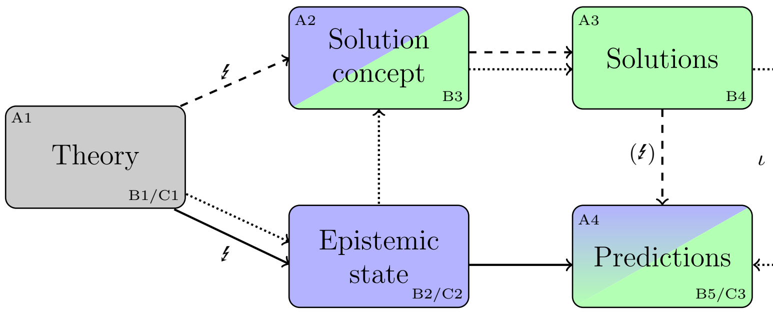
B3/C3. Derive state-implied set of potential outcomes, either via an existing solution concept (B) or from the epistemic state directly (C)

[3] That is, a theorist should not be able to (cherry-)pick a set of 'plausible' outcomes and then, without any constraints or further justification, select a solution concept so as to legitimize and/or obfuscate said choice.

[4] It is tempting, for two reasons, to reject the 'only if' portion of this argument. First, while jointly sufficient, the listed conditions, including the ability to correctly anticipate others' choices, are *not* individually necessary. However, this is only because such necessary conditions do not exist. Indeed, note that players can always "blunder into a Nash equilibrium by accident" (see Aumann and Brandenburger, 1995), a fact that is surely insufficient for a theorist to restrict their attention to Nash outcomes for purposes of prediction. Second, Aumann and Brandenburger (1995) provide an epistemic characterization of Nash equilibrium in which choices are mutually unknown (see Section 2). However, even if said conditions were fulfilled, invoking Nash equilibrium for purposes of prediction would still be inappropriate, namely because they only induce equilibrium in beliefs, not actual play.

When comparing Procedures B and C to the canonical Procedure A, first note that, to the extent that any solution concept can be derived from an underlying epistemic state[5], Procedure B nests Procedure A. For example, the 'Nash specification' of Procedure A can easily be mimicked by way of specifying, in step B2, the aforementioned epistemic state that underpins Nash equilibrium.

Figure 1. Three procedures to derive predictions from economic theory



Notes: Figure 1 contrasts the traditional solution-concept-based approach (i.e Procedure A) with the epistemic approach advocated for herein (Procedures B and C). Here, ι is used to denote the identity map, the lightning symbol (↯) and blue represent assumptions, and green denotes results. The main insight of Figure 1 is then that the derivation of plausible predictions neither requires nor benefits necessarily from the use of solution concepts. This is because, for purposes of deriving plausible predictions, specifying an epistemic state is not only sufficient, but it is also more versatile and more transparent.

While interesting, the fact that Procedures B nests Procedure A is relatively inconsequential. Instead, the primary virtues of the epistemic approach lie in its transparency and its versatility.

In terms of transparency, the primary benefit of the epistemic approach is that it requires an explicit disclosure of the epistemic assumptions that are required to yield the proclaimed predictions. For example, when invoking Nash equilibrium, a theorist should disclose that, and ideally why, players are assumed to correctly anticipate each other's choices. In turn, such disclosures allow the reader to interrogate more transparently the plausibility of the presented argument. In particular, if the asserted epistemic state seems contextually inappropriate, the reader can discern more transparently the practical limitations of the corresponding predictions. Conversely, if the asserted epistemic state is uncontroversial, the proclaimed predictions' credibility is strengthened.[6]

Beyond transparency, the main benefit of the epistemic approach lies in its versatility. This addresses an important practical limitation of existing solution concepts, all of which reflect epis-

[5]Even if this were not true, it is entirely unclear why or under which circumstances a solution concept that fails this very basic requirement would ever be of practical relevance or interest.

[6]For example, in a game of Prisoner's dilemma (see G_2 in Appendix A) in which neither player knows anything about the other player, appealing to Nash equilibrium is not only epistemically inappropriate, but it is also entirely unnecessary. Indeed, so long as each player acts rationally, we can eliminate a single round of strictly dominated strategies, which is sufficient to imply defection by both players.

temic states that are perfectly symmetric. For example, to be able to invoke Nash equilibrium, *each* player must correctly anticipate *every* other player's choice. Similarly, to invoke rationalizability, rationality must be *commonly* assumed. But what if a certain player, or a certain fraction of players, is known to play irrationally? For example, a theorist who wishes to explain the actually observed behavior in real-world Keynesian beauty contests inevitably must account for the fact that some players play irrationally, whereas others seemingly (and falsely) assume that rationality is commonly assumed (see Thaler, 2015).[7] It is for this reason, to account for such asymmetries, that we introduce Procedure C as a complement to Procedure B.

To illustrate its merits in more detail, Section 3 applies the epistemic approach to various well-known economic frameworks: The Diamond-Dybvig model of bank runs (1983), Calvo's indeterminate sovereign debt auction (1988), Akerlof's market for lemons (1970), and various market economies (Arrow and Debreu, 1954; Bertrand, 1883; Cournot, 1838). Specifically, for each application, we first specify a suitable epistemic state and then derive our empirical predictions therefrom. Whenever applicable, following the logic of Procedure B, we reference the relevant solution concept. Especially in the context of our Akerlof-inspired analysis of the market for a used car, however, departing from existing solution concepts and allowing for epistemic asymmetries (i.e. Procedure C) yields additional insights.

Related literature. Broadly, the present paper lies at the intersection of two vast theoretical literatures: (i) the applied literature that leverages solution concepts for purposes of prediction (i.e. game theory, micro- and macroeconomic theory); and (ii) the epistemic literature that studies solution concepts' decision-theoretic roots.[8] Since Section 2 is devoted to collecting insights from the latter, an additional review of this literature is omitted here. Instead, a comment regarding our paper's distinctive objective is in order: Unlike the epistemic literature, we do not seek to establish the epistemic underpinnings of an existing solution concept.[9] In particular, this is because such characterizations seek to answer whether a particular solution concept can (i.e. in general/across all finite games) be motivated by way of some underlying epistemic state. However, the fact that a

[7] In a Keynesian beauty contest, players are asked to pick a number in the closed interval $[0, 100]$ with the objective of matching two thirds of the mean response.

[8] The list of papers which introduce a novel solution concept is too long to present here, so we limit ourselves to the following, particularly well-known subset: Nash (1950), Harsanyi (1973), Aumann (1974), Selten (1975), Myerson (1978), Kreps and Wilson (1982), van Damme (1984), Gilboa and Matsui (1991), Carlsson and van Damme (1993).

[9] See Bernheim (1984), Pearce (1984), Aumann (1987), Blume, Brandenburger, and Dekel (1991b), Aumann and Brandenburger (1995), Brandenburger, Friedenberg, and Keisler (2008).

solution concept *can* be motivated by way of some underlying epistemic state does not mean that said state serves as an empirically suitable description of any arbitrary strategic environment at hand.[10] Thus, rather than fixing a solution concept and examining whether it can be epistemically characterized across all games, we fix various games and ask which solution concept most suitably reflects the corresponding strategic environment.

To illustrate the insight that solution concepts should be chosen to accurately reflect the strategic environment at hand, we now briefly review a theoretical argument presented in Hu and Sobel (2022). For this, suppose two 'experts', $i \in \{1, 2\}$, with utility $u_i(x) \equiv 1 - \max_{i \in I}\{x_i\}$ simultaneously and independently choose $x_i \in X \equiv \{0, 0.1, ..., 0.9, 1\}$ so as to determine which projects in X a manager (with utility $u_m = \max_{i \in I}\{x_i\}$) is permitted to undertake.[11] Interestingly, this game features as many Nash equilibria as it features actions (i.e. eleven, all symmetric), but $x_i = 0$ is weakly dominant for both experts and, as such, might be perceived as the only equilibrium that is 'plausible'. Following this logic, the referenced authors write: "the possibility of multiple equilibria leads us to consider a more restrictive solution concept" (i.e. weak dominance).

With reference to the theory described above, we concur with the authors that weak dominance represents a more suitable solution concept than Nash equilibrium. Nevertheless, our paper contributes — via Figure 2 — to their analysis in two ways: First, the reason why Nash equilibrium is contextually inappropriate is *not* that it generates multiple solutions, which it does, but rather because its epistemic requirements do not plausibly reflect the strategic setting at hand.[12] Second, the reason why weak dominance is contextually appropriate is *not* that it generates a unique and intuitive prediction, which it does, but rather because its epistemic requirements plausibly reflect the epistemic state at hand. Indeed, in our specification, an immediate elimination of all strategies other than $x_i = 0$ is warranted so long as we are willing to assume that both experts know their own payoffs, that they are rational, and that they are cautious (see Appendix B).[13]

[10]For example, the fact that Nash equilibrium *can* be rationalized across the universe of all finite games (namely via rationality, knowledge of payoffs, and mutual knowledge of others' choices) does not imply that it does, in fact, represent a suitable solution concept across all such games. Indeed, in most games, there is no reason to believe that players correctly anticipate opponents' actual/final choices.

[11]The referenced authors' framework is substantially more general in that players' utility is only assumed to satisfy $u_i = v_i(M(x))$, where $x = (x_1, x_2, ..., x_I)$, $x_i \in X$ for each i in the finite set of players I, $X \subset \mathbb{R}$ is also finite, $M(x) \equiv \max_{i \in I}\{x\}$, and $v_i : X \mapsto \mathbb{R}$ is bijective for each i. Nevertheless, since our formulation satisfies all of the aforementioned conditions, their results apply.

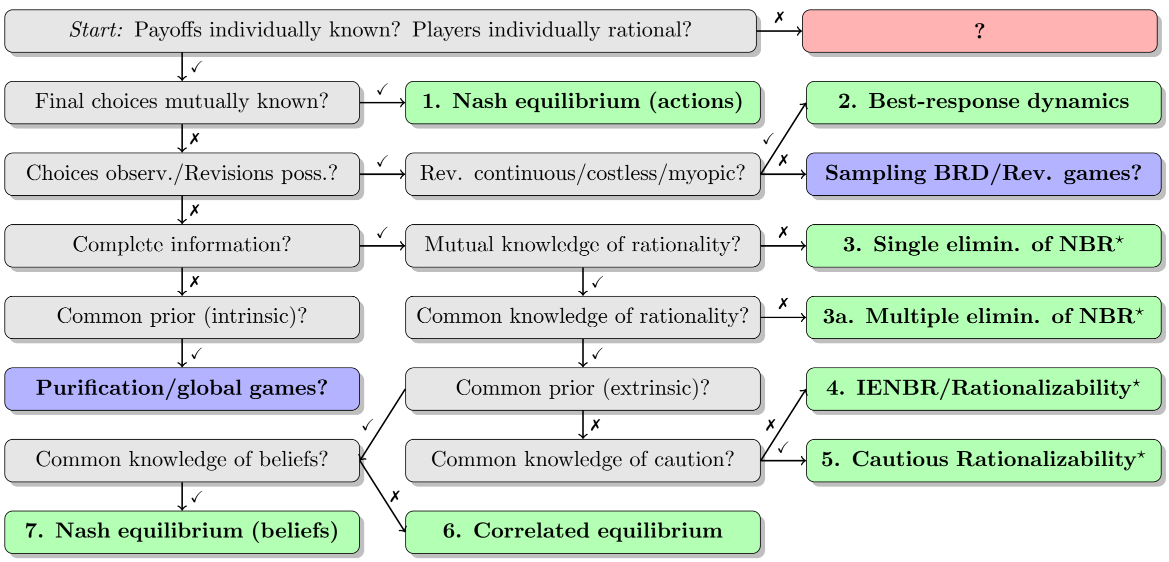
[12]Indeed, without further justification, there is no reason to believe that the experts know each other's actual/final choice, nor that their beliefs regarding the other's play are mutually known (see Figure 2).

[13]Rationality and caution yield a single round of elimination of weakly dominated strategies.

2 The epistemic approach

In this section, we illustrate our paper's main motivating observation — that epistemic conditions can be used to narrow down a game's set of conceivable outcomes — by collecting a number of key results from the epistemic literature in the form of a flowchart (Figure 2).[14] Aside from serving as a useful tool for future research, Figure 2 also illustrates the implicit epistemic assumptions that remain hidden when solution concepts are invoked as model primitives (i.e. Procedure A).

Figure 2. A map of epistemic conditions into well-known solution concepts



Notes: Figure 2 maps various sets of epistemic conditions into their corresponding solution concept and, as such, represents an extension of Table 1 in Brandenburger (1992b). While all of the depicted conditions are sufficient, some, to keep the flowchart as compact as possible, are not 'tight' in that weaker conditions would be sufficient to yield the same concept (e.g. Nash equilibrium only requires mutual knowledge of the game and rationality). Fields in blue contain possibly suitable solution concepts that would require further assumptions, whereas the solution concepts denoted by a star (*) vary based on whether players' can correlate their random choices.[15]

To motivate the depicted results within a unified framework, we consider the class of all finite N-person games $\mathbf{\Gamma}$ such that each game $\Gamma \in \mathbf{\Gamma}$ features a finite set of players I^Γ all of whom choose a possibly mixed strategy σ_i^Γ in a finite space of pure strategies Σ_i^Γ taking as given their utility function $u_i^\Gamma : \Sigma^\Gamma \mapsto \mathbb{R}$.[16] Let us then start by defining as follows,

[14] Importantly, Figure 2 depicts epistemic conditions that are *sufficient* to invoke a particular solution concept. As pointed out by Aumann and Brandenburger (1995) in the context of Nash equilibrium, it makes little sense to contemplate 'necessary' epistemic conditions as players can, even if they don't know their payoffs and/or are irrational, always tumble (i.e. by accident) into a profile that just so happens to align with a particular solution concept.

[15] If correlated play is permitted and the strategy space is finite, SENBR, MENBR, and IENBR are equivalent to single (SESDS), multiple (MESDS), and iterated elimination of strictly dominated strategies (IESDS) respectively. This is because a strictly dominated strategy is always a never-best response, but a never-best response is not always, unless correlated play is permitted, strictly dominated (see Pearce, 1984).

[16] The results depicted in Figure 2 are based on the assumption that each player's strategy space is finite and, as

Solution concept. A *solution concept* is a function $S : \Gamma \mapsto \Sigma^\Gamma$ that restricts the set of a game's possible outcomes Σ^Γ to a concept-implied subset $S(\Gamma) \subseteq \Sigma^\Gamma$ for each $\Gamma \in \mathbf{\Gamma}$.

Solutions. The *solutions* $\mathcal{S}$ of a game $\Gamma \in \mathbf{\Gamma}$ under a solution concept $S : \Gamma \mapsto \Sigma^\Gamma$ are given by $\mathcal{S} = S(\Gamma)$.

The primary benefit of our formal treatment above lies in the insight that, in general, a game's set of solutions not only depends on the game itself, but also on the employed solution concept. But then, if the choice of solution concept is so crucial, how should one go about choosing it? For this, we follow the decision-theoretic literature which emphasizes that solution concepts reflect an underlying epistemic state.

Epistemic state. A game's *epistemic state* $\mathcal{E} \in \boldsymbol{\mathcal{E}}$ consists of players' individual and collective knowledge (or assumptions[17]) "about the game and about each other's rationality, actions, knowledge, and beliefs" (Aumann and Brandenburger, 1995).[18]

Throughout the following, we broadly distinguish between three types of epistemic states: (i) when players' choices are mutually known; (ii) when choices are mutually observable; and (iii) when they are neither mutually known nor mutually observable.[19] In the last case, we further distinguish between various levels of knowledge of payoffs, rationality, caution, and beliefs (see Figure 2).

2.1 When choices are mutually known

In our construction of Figure 2, we start with the (seemingly) innocuous assumptions that each player $i \in I$ knows their own action space Σ_i, their utility function u_i, and that all players are rational in that they choose their actions so as to maximize utility.[20] In this context, let us first consider the special case in which all players' actual/final choices are mutually known,

Rationality (choices mutually known). If a player i knows all other players' actual/final

such, may not extend to infinite strategy spaces (see Lipman, 1994). However, the main insight contained in Figure 2 — that varying solution concepts reflect varying epistemic states — *does* extend to infinite strategy spaces.

[17] In our construction of Figure 2, we limit our analysis to the canonical notion of 'knowledge' as proposed by Aumann (1976). However, to motivate our use of weak dominance/admissibility in Section 3, Appendix B discusses the alternate, LPS-based notion of 'assumption' due to Brandenburger, Friedenberg, and Keisler (2008).

[18] Since most epistemic states do not map into a well-known solution concept, our analysis only scratches the surface of $\boldsymbol{\mathcal{E}}$. As such, the present paper might be viewed as fertile ground for future research. However, as alluded to in Section 1 and illustrated in Section 3, the power of Procedure C is precisely that it operates independently of solution concepts.

[19] A fact is said to be mutual knowledge if all players know said fact.

[20] To simplify notation, we fix a particular game $\Gamma \in \mathbf{\Gamma}$ and, thus, drop the superscripts.

choices σ_{-i}, they are said to be *rational* if and only if their own final choice constitutes a best response with respect to σ_{-i}, i.e. $u_i(\sigma_i|\sigma_{-i}) \geq u_i(\sigma_i'|\sigma_{-i})$ for each $\sigma_i' \in \Sigma_i$ and each $\sigma_{-i} \in \Sigma_{-i}$.

Result 1 (Aumann and Brandenburger, 1995). In finite N-person games where actual/final choices are mutually known, and all players are rational, the resulting play σ must be Nash, i.e. $u_i(\sigma_i|\sigma_{-i}) \geq u_i(\sigma_i'|\sigma_{-i})$ for each $\sigma_i' \in \Sigma_i$, each i, and each $\sigma \in \Sigma_{NE}$.

At this point, two remarks are in order, one in defense of Nash equilibrium and one in opposition to Nash equilibrium. First, it is tempting to think that the above result only applies to finite games since Nash equilibrium may not exist in games with infinite strategy spaces.[21] However, this is not true because if all players' actual/final actions are mutually known and all players are rational as defined above, then the resulting play, by definition, must be Nash. Thus, the observation that certain games fail to feature a Nash equilibrium simply reflects the fact that, in some games, it is impossible for all players to jointly satisfy the above definition of rationality. As such — and this is our second remark — although appealing from a mathematical perspective, the above result is economically artificial in that it is difficult to imagine a strategic environment where all players truly *know* (i.e. with certainty) everyone else's actual/final choice when they make their own final choice.[22] To relax this assumption, we now consider the case where choices are mutually unknown.

2.2 When choices are mutually observable

If choices are mutually unknown, the task of selecting an action is complicated by the fact that utility depends on a player's beliefs $\phi^i : \Sigma_{-i} \mapsto [0,1]$. Let us then define,

Rationality (choices mutually unknown). If player i does not know all other players' actual/final choices σ_{-i}, they are said to be *rational* if and only if their own final choice constitutes a best response with respect to their beliefs, i.e. $\mathbb{E}[u_i(\sigma_i|\sigma_{-i})] \geq \mathbb{E}[u_i(\sigma_i'|\sigma_{-i})]$ for each $\sigma_i' \in \Sigma_i$ and some beliefs ϕ^i over σ_{-i}.[23]

For example, if actions are mutually observable and strategic revisions are possible, a natural way to proceed is to assume that each player's (posterior) beliefs coincide with the presently pre-

[21] For example, suppose two players choose a natural number and whoever chooses a strictly higher number wins, i.e. $u_i = \mathbb{1}[\sigma_i > \sigma_{-i}]$ for both i. Clearly, this game does not feature any Nash equilibria.

[22] "Nash equilibrium does make sense if one starts by assuming that, for some specified reason, each player knows which strategies the other players are using. But this assumption appears rather restrictive" (Aumann, 1987).

[23] While this definition is standard in the epistemic literature, Brandenburger et al. (2008) consider an alternate definition that includes the avoidance of weakly dominated strategies as part of a player's rationality.

vailing strategic state (which supersedes any preexisting priors). In this case, if strategic revisions are continuous, costless, and myopic, best-response dynamics constitute a contextually appropriate solution concept.[24]

Result 2 (Gilboa and Matsui, 1991; Matsui, 1992). In finite N-person games, if the prevailing strategic state $\sigma(t) \in \Sigma$ is mutually observable and strategic revisions are continuously possible, costless, and myopic (with or without strategic trembles), then the resulting play takes the form of a best-response dynamic path (BRDP), i.e. $\sigma : T \mapsto \Sigma$.[25,26]

2.3 When choices are neither mutually known nor observable

If choices are neither known nor observable, players must resort to forming beliefs based on their knowledge of others' characteristics such as rationality and/or caution. In particular, players can use such knowledge to restrict the set of possible beliefs to a *permissible* subset.

Permissible beliefs. A profile of beliefs ϕ is said to be *permissible* if and only if it does not contradict the game's epistemic state $\mathcal{E}$.

For example, if a player knows that the game being played is mutual knowledge and that all players are rational, they can infer that no player will play a strictly dominated strategy. Following this logic, we continue our construction of Figure 2 by assuming that the game itself is common knowledge.[27]

Complete information. A game Γ is said to be of complete information if Γ itself (but not the actual/final choices $\sigma \in \Sigma$) is common knowledge.[28]

[24]Conversely, if strategic revisions are subject to frictions, alternate concepts such as *sampling best-response dynamics* (see Oyama et al., 2015) or *revision games* (see Kamada and Kandori, 2020) may be more appropriate.

[25]Gilboa and Matsui (1991) allow for strategic trembles, whereas Matsui (1992) does not.

[26]Much like Nash equilibrium, the solution concept of best-response dynamics itself naturally extends to infinite strategy spaces, although, once again, existence is not guaranteed as the best-response correspondence may not be well-defined across the entire strategy space.

[27]A fact is said to be *common knowledge* if all know, all know that all know, and so on ad infinitum (see Aumann, 1976).

[28]Many games of incomplete information can be transformed into games of complete information through the suitable use of random variables (and expected utility). For example, if players are unaware of others' payoffs (see Harsanyi, 1973) or even their own payoffs (see Carlsson and van Damme, 1993), a corresponding common prior in conjunction with expected utility is sufficient to permit an interpretation of such environments (i.e. *purification* and *global games* respectively) as ones of complete information: "Such games [incomplete information games] can also be modeled, more conveniently, as games with complete information involving appropriate random variables (chance moves), where the players' ignorance about any aspect of the game situation is represented as ignorance about the actual values of these random variables." (Harsanyi, 1973)

In games of complete information, knowledge about other players, especially their rationality, can be used to restrict a game's set of conceivable outcomes.

Knowledge of rationality

As famously shown by Bernheim (1984) and Pearce (1984), rationality is in and of itself is insufficient to imply Nash equilibrium. Indeed, even in a game of complete information where all players are individually rational, but no player knows this, an outside observer can at most rule out a single round of never-best responses.

Result 3. In finite N-person games of complete information where actual/final choices are mutually unknown, but all players are rational, no player plays a never-best response (NBR), i.e. $\sigma \in \Sigma_{NBR_1}$.

To obtain a 'tighter' restriction (beyond a single elimination of NBRs), further epistemic assumptions are required. For example, mutual knowledge of rationality allows for a second round of elimination, second-order mutual knowledge for a third, and so on.[29] In turn, perhaps unsurprisingly, common knowledge of rationality yields iterated elimination of never-best responses (IENBR), a solution concept more commonly known as *rationalizability*.

Result 4 (Bernheim, 1984; Pearce, 1984). In finite N-person games of complete information where actual/final choices are mutually unknown, but rationality is commonly known, the resulting play survives iterated elimination of never-best responses, i.e. $\sigma \in \Sigma_R \equiv \Sigma_{NBR_\infty}$.[30]

Aside from knowledge of rationality, there exist three further types of knowledge that have been used to discipline a game's set of solutions: caution, common priors, and beliefs.

Caution and knowledge thereof

Pearce (1984) proposed that a player might act 'cautiously' in that they consider each of their opponents' rationalizable strategies with positive probability. In turn, strategies that fail to be a best response to any such 'cautious beliefs' are then ruled out as 'not cautious'. Moreover, if players' caution is common knowledge, the two described operations — eliminating all non-rationalizable strategy profiles and all not-cautious responses thereto — are repeated until the remaining set of

[29]See Brandenburger, Friedenberg, and Kneeland (2020) for a comparative study of such iterated dominance and level-k analysis.

[30]Lipman (1994) shows that this result can be extended to the case of infinite strategy spaces, albeit only if we allow for "transfinite" (e.g. uncountably infinite) eliminations of never-best responses as part of common knowledge of rationality.

strategy profiles no longer varies.

Result 5 (Pearce, 1984). In finite N-person games of complete information where final choices are mutually unknown, but rationality and caution are commonly known, the resulting play σ must be cautiously rationalizable, i.e. $\sigma \in \Sigma_{CR} \subseteq \Sigma_R$.[31]

Common priors

To provide an epistemic justification for correlated play in non-cooperative settings, Aumann (1987) proposed that agents might coordinate their actions by way of an (extrinsic) state.

Result 6 (Aumann, 1987). In finite N-person games where final choices are mutually unknown, but rationality is commonly known and players share a common prior over observed play $\sigma : \Omega \mapsto \Sigma$ (via a common prior over the set of states of the world Ω), the latter forms a *correlated equilibrium*.[32]

Importantly, conditional on the state of the world ω, players are assumed to be able to infer others' actual/final choices by way of the mapping $\sigma_{-i} : \Omega \mapsto \Sigma_{-i}$. However, since they do not know the true state of the world, they proceed by deriving a posterior over σ_{-i} by way of their (individual) posterior over ω. In effect, in the described setting, a player's set of permissible beliefs is given by a singleton (and equal to said player's posterior distribution over σ_{-i} as implied their posterior over ω) for any given realization of ω.

In a correlated equilibrium, actually observed behavior must be rationalizable because rationality is common knowledge. Following this logic, one might then naturally wonder what additional epistemic conditions would be required for actually observed play to be Nash.

Knowledge of beliefs

To develop an intuition for the epistemic requirements of Nash equilibrium when choices are mutually unknown, we require two lemmas.

Lemma (Aumann and Brandenburger, 1995). In two-person games where final choices are mutually unknown, mutual knowledge of beliefs, payoffs, and rationality imply that said beliefs form a (possibly mixed strategy) Nash equilibrium.

[31] Following the logic from Lipman (1994), we suspect that this result extends to the case of infinite strategy spaces (so long as there exist strictly positive probability mass/density functions thereon), namely if we allow for transfinite eliminations of never-best responses and not cautious responses thereto as part of common knowledge.

[32] Intuitively, the defining element of correlated equilibrium is that players' actions may correlate. Thus, unlike in mixed strategy Nash equilibrium, the common prior over σ need not be equal to the product of its marginals $\{\sigma_i\}_{i \in I}$.

The above result is intuitive in that player 1 can leverage their knowledge of player 2's beliefs, payoffs, and rationality to infer which actions player 2 might choose. More precisely, player 1's beliefs ϕ^1 only assign positive probabilities to actions in player 2's best response set (relative to ϕ^2). Analogously, player 2's beliefs ϕ^2 only assign positive probabilities to actions in player 1's best response set (relative to ϕ^1). Thus, (ϕ^1, ϕ^2) must form a Nash equilibrium.[33]

To extend the above result to three or more players, we require that all players agree on a shared set of mutually independent beliefs,

Lemma (Aumann and Brandenburger, 1995). In finite N-person games where final choices are mutually unknown, mutual knowledge of independent and collectively congruent beliefs, payoffs, and rationality imply that said beliefs form a (possibly mixed strategy) Nash equilibrium.[34]

Indeed, if all players share the same beliefs regarding player i's play and player i's payoffs and rationality are mutually known, then said shared beliefs may only assign a positive probability to an action (of player i) if said action lies in player i's best response set (relative to the mutually known/shared beliefs). But then, if the profile of shared beliefs only assigns positive probabilities to actions in the corresponding player's best-response set, then said profile must, if its individual elements are stochastically independent, form a Nash equilibrium.[35]

Although appealing intuitively, Aumann and Brandenburger (1995) argue that the above result rests on "dubious" epistemic foundations in that it is unclear, without further justification, how players might coordinate their beliefs regarding each other's play and why such shared beliefs should satisfy independence. In turn, their main contribution lies in the proof that both of these conditions are implied by common knowledge of beliefs in conjunction with a common prior.[36]

Result 7 (Aumann and Brandenburger, 1995). In finite N-person games where final choices are mutually unknown, but players share a common prior over observed play $\sigma : \Omega \mapsto \Sigma$ (via a common prior over the set of extrinsic states of the world Ω) and, at some state, payoffs and

[33] Much like Result 1, this trivially extends to games with infinite strategy spaces. Thus, if Nash equilibrium does not exist in such a game, then the listed epistemic conditions cannot all simultaneously be satisfied.

[34] Once again, this result trivially extends to games with infinite strategy spaces. Thus, if Nash equilibrium does not exist in such a game, then the listed epistemic conditions cannot all simultaneously be satisfied.

[35] Independence is required because mixed strategy Nash equilibrium is an inherently independent concept. Indeed, if we drop the independence requirement, any such profile of shared beliefs (that satisfies mutual knowledge of payoffs/rationality) must form a correlated equilibrium.

[36] The key element driving this result lies in the impossibility of players to 'agree to disagree' when they share a common prior and individual posteriors are commonly known (Aumann, 1976). See Brandenburger (1992b) for a discussion of the implications of this result in the context of Nash equilibrium.

rationality are mutually known and beliefs are commonly known, then all players share the same beliefs (pertaining to any particular player's choice) and said beliefs form a (possibly mixed strategy) Nash equilibrium.[37]

To appreciate the practical implications of Result 7 (or the lack thereof), first note that it only speaks to players beliefs', not their actual play.[38] Thus, since actually observed play very well may not be Nash when players choose mixed strategies, even the extraordinarily demanding conditions outlined above are insufficient to imply that actually observed play is, in fact, Nash. Of course, actual play does have to be Nash if each player's best response (relative to the mutually shared beliefs) is a singleton. Indeed, in this case, the mutually shared beliefs would form a Nash equilibrium in pure strategies.[39]

3 Applications

While Section 2 serves to document that, in general, epistemic conditions *can* be used to narrow down a game's set of conceivable outcomes, Section 3 illustrates the relative merits of this approach by re-examining various well-known economic theories through an epistemic lens. For example, in our first application — the Diamond-Dybvig model of bank runs (Section 3.1) — we provide an epistemic justification for how and, thus, why the economy might arrive at equilibrium, thus strengthening the persuasiveness of the original result. In contrast, in our second application — Calvo's model of sovereign default (Section 3.2) — we show how an indiscriminate choice of Nash equilibrium (i.e. Procedure A) can yield misleading results and, ultimately, obfuscate policy-relevant aspects of the proposed theory. Finally, in Sections 3.3 and 3.4, we present an epistemic reading of the market for used cars akin to Akerlof (1970), competitive equilibrium as studied by Arrow and Debreu (1954), and the two canonical duopoly models based on Cournot (1838) and Bertrand (1883).[40]

[37] Unlike in the two prior cases, it is unclear if this result extends to games in which the strategy space is infinite.

[38] Nevertheless, so long as all players choose to act in accordance with others' shared beliefs (regarding their own play), no player has an incentive to deviate.

[39] In turn, if the same were true across all states of the world, then the common prior would only assign positive probabilities to strategy profiles that are pure Nash and so global play would assume the form of a correlated equilibrium randomizing over the relevant game's set of pure strategy Nash equilibria.

[40] Since all of our chosen theories feature strategy spaces that are uncountably infinite, the mapping from epistemic conditions to solution concepts depicted in Figure 2 may not apply. However, this does not present an issue because, as illustrated across all five examples, the power of the epistemic approach precisely lies in its independent operation from solution concepts (see Figure 1).

3.1 Bank runs (Diamond and Dybvig, 1983)

While game-theoretic reasoning is now standard across many fields of economics, this was not always the case (Samuelson, 2016). An early and highly influential contribution to bridge said gap was the Diamond-Dybvig model of bank runs.[41] Diamond and Dybvig (1983) not only explicitly appeal to Nash equilibrium, but they were also the first to embrace indeterminacy as a means to explain two highly disparate empirical phenomena — bank run vs. no bank run — within a unified economic framework. We thus choose their model as the basis for our first application.

At the heart of the Diamond-Dybvig model lies the following choice by fundamentally patient depositors: Either, they withdraw their deposits at the cost of a lower return, or they leave their deposits in the bank at the risk of having the latter default on their claim. In effect, the described game features two (stable) Nash equilibria: Only impatient depositors withdraw (no run) or everyone withdraws (run). To coordinate between these two equilibria, the literature conventionally proceeds by appealing to sunspots (see Peck and Shell, 2003) which effectively motivates coordination as follows:

(A1) *Assuming* depositors can distinguish between two extrinsic states of the world (ξ_0, ξ_1), where everyone anticipates each patient depositor not to run in state ξ_0, and to run in state ξ_1, then no investor has an incentive to deviate from said play and the resulting global play forms a correlated equilibrium. Moreover, since each depositor correctly anticipates everyone else's actions in both states, actually observed play must be Nash (in each state).

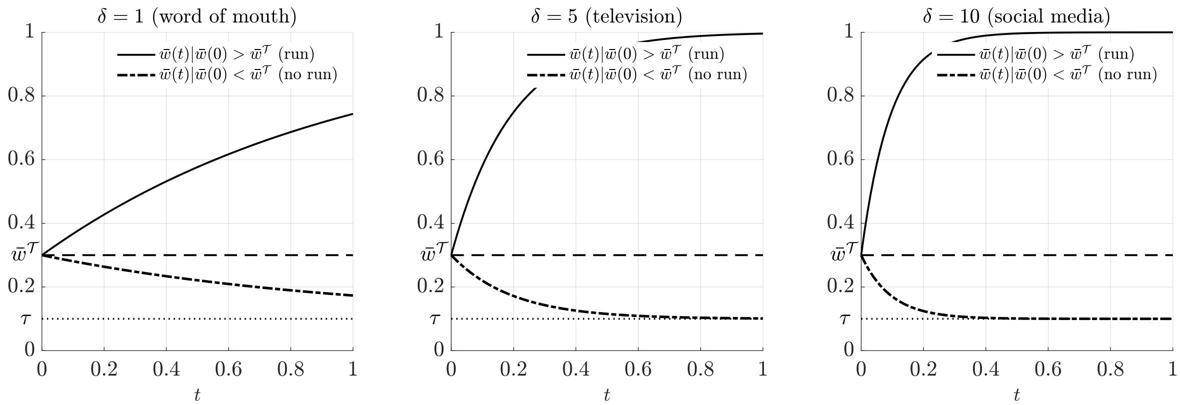
From a decision-theoretic perspective, the above narrative begs the following question: Is there reason to believe that, when deciding whether to run, real-world depositors can, in fact, correctly anticipate each others' actions? In other words, is equilibrium, correlated or Nash, an epistemically appropriate solution concept? In our view, the answer to this question is negative. But then, if equilibrium is unsuitable, why does it yield such intuitive predictions? To answer this question, Mäder (2024) notes that, during a bank run, depositors can typically *observe* each other's (tentative) actions. In turn, to the extent that actually observed play trumps any sort of pre-existing beliefs, evolutionary concepts such as best-response dynamics represent a more suitable epistemic lens:

(A1') *Assuming* depositors can observe each other's tentative actions, and that, on a continuous basis, a certain fraction revises their prevailing choice by myopically choosing a best response, tentative play assumes the form a *best-response dynamic path* (see Gilboa and Matsui, 1991). In turn, if the speed of revision is sufficiently fast, the ultimately observed play is Nash.

[41] See "Equilibrium without an auctioneer" by Peter Diamond (1987).

Thus, akin to Prisoner's Dilemma (see footnote 8), the reason why Nash equilibrium yields such intuitive predictions in Diamond and Dybvig (1983) is not that it represents an epistemically appropriate solution concept itself, but rather because the epistemically appropriate solution concept — best-response dynamics — naturally induces *outcomes* that happen to be Nash. Indeed, note that the Diamond-Dybvig model famously features a third type of equilibrium, but this 'tipping point' is dynamically unstable and so it is *not* induced by best-response dynamics (see Mäder 2024).

Figure 3. Best-response dynamics in the Diamond-Dybvig model of bank runs



Notes: Figure 3 reproduces, albeit in a slightly augmented fashion, Fig. 3 in Mäder (2024). It depicts various best-response dynamic paths of the Diamond-Dybvig model across various revision speeds δ and for two initial conditions near the unstable tipping point $w^{\mathcal{T}}$. $\bar{w}(t)$ denotes aggregate withdrawals at time t and τ is the fraction of depositors who withdraw either way (because they are fundamentally impatient).

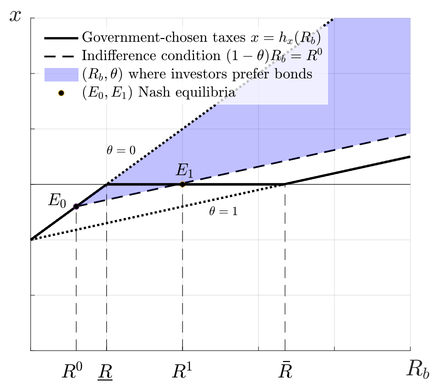
If the domain of Figure 3 is interpreted as a day, Panel A intuitively captures an 'old-fashioned' bank run whereby news primarily transmit bilaterally via word of mouth. In turn, the faster revision speeds captured in Panels B and C represent more modern economies, whereby the news of an ongoing run is broadcasted by way of television and/or social media (see Cookson et al., 2023). By the end of day, the ultimately observed outcome in economies in Panels B and C is Nash, whereas in Panel A it is not. Of course, once read through the lens of best-response dynamics, the fact that some depositors may fail to withdraw their money during an ongoing run does not (necessarily) speak to their rationality or lack thereof. Instead, such depositors might have simply heard the news too late or been unavailable otherwise.

3.2 A sovereign debt auction (Calvo, 1988)

Akin to the literature on bank runs, the contemporary literature on sovereign default conventionally follows Calvo (1988) in rationalizing the occurrence of such default by way of multiple theoretical equilibria.[42] For our second application, we thus choose Calvo's canonical model of sovereign default. This choice is mainly motivated by the fact that, in this case, our epistemically preferred solution concept — weak dominance — not only offers a novel narrative of default, but it also offers pertinent insights from the point of view of policy.

To rationalize default by way of multiple Nash equilibria, the idea at the heart of Calvo (1988) is that a sovereign bond's credit risk is increasing, at least weakly, in its own interest rate.[43] To see this, consider Figure 4, which reproduces, in a slightly altered fashion, Calvo's Figure 2.

Figure 4. Optimal repudiation θ as a function of the auction-implied yield R_b in Calvo (1988)



Notes: Figure 4 reproduces Fig. 1 from Campbell and Mäder (2023), a slightly augmented version of Fig. 2 in Calvo (1988). It depicts the government's optimal choice of taxation x and repudiation share θ as a function of the primary market yield R_b. As such, it contains three main insights: (i) credit risk is weakly increasing in R_b, (ii) investors strictly prefer bonds when $R_b \in (R^0, R^1)$, and capital when $R_b > R^1$, and (iii) equilibrium is indeterminate.

To coordinate between 'good' equilibria à la E_0 and 'bad' equilibria à la E_1, the literature conventionally proceeds by appealing to sunspots (see, for example, Cole and Kehoe, 2000):

(B1) *Assuming* investors can distinguish between two extrinsic states of the world (ξ_0, ξ_1) where it is mutually assumed that each investor anticipates everyone else to bid R^0 in state ξ_0, and R^1 in state ξ_1, then no investor has an incentive to deviate from their prescribed play and the resulting global play forms a correlated equilibrium. Moreover, since each depositor correctly anticipates everyone else's actions across both states, actually observed play is Nash.

[42] Indeed, after proving that equilibrium in Eaton and Gersovitz (1981) is unique, Auclert and Rognlie (2016) elaborate as follows: "Our objective is not to deny that sovereign debt markets can be prone to self-fulfilling crises [...]. Instead, we hope that our results may help sharpen the literature's understanding of the assumptions that are needed for such multiple equilibria to exist."

[43] In Calvo (1988), since default is certain whenever $\theta > 0$, θ does not technically reflect credit *risk*.

From a decision-theoretic perspective, the above narrative begs the following question: Is there reason to believe that, when deciding how to bid in a sovereign debt auction, real-world investors can, in fact, correctly anticipate each others' bids? In other words, is equilibrium, correlated or Nash, an epistemically appropriate solution concept? In our view, under ordinary circumstances, the answer is negative. But then, if equilibrium is not suitable, which solution concept *is* suitable? To answer this question, Campbell and Mäder (2023) note that, even though investors may not know what others are bidding, they very well may possess knowledge and/or be willing to make assumptions that allow them to narrow down the types of bids that will, or will not, be submitted. For example, so long as all investors are both rational and cautious, we can rule out a single round of weakly dominated strategies,

(B1') *Assuming each investor is both rational and cautious (and knows how the government operates)*, no bids below the risk free rate R^0 or in excess of $\underline{R}$ are submitted.[44]

Intuitively, the reason why higher yields are only appealing up to a certain point is that the government responds to higher yields via increased taxation x and, ultimately, repudiation θ.[45] In turn, although Assumption (B1') significantly narrows down the types of bids that are submitted, it is insufficient to imply that any bids are submitted in the first place.[46] Thus, to rule out rollover crises and/or obtain a unique prediction, we require additional epistemic assumptions,

(B2') *Assuming payoffs are mutually known and each investor assumes (B1')*, they correctly infer that $R^0 \leq R_b \leq \underline{R} < R^1$ and, thus, optimally choose to participate in the auction.

(B3') *Assuming each investor assumes (B2')*, they correctly infer that every investor participates in the auction. In turn, so long as neither successful, nor marginal bids can influence the ultimately transacted rate when every investor participates in the auction, bidding R^0 is the uniquely remaining, cautiously rational strategy for each investor.[47]

In summary, rationality, caution, and two layers of mutual assumption (alongside second-order mutual knowledge of the game and the government's operations) yields the 'good' Nash equilibrium (i.e. E_0 in Figure 4) via three rounds of elimination of weakly dominated strategies (3EWDS).[48]

At this point, the reader might, and hopefully does, wonder whether 3EWDS does, in fact, represent a suitable epistemic concept to describe an ordinary sovereign debt auction. In the

[44]Thus, note that the 'bad' equil. E_1 has each investor playing a weakly dom. strategy, a rather odd proposition.

[45]In fact, depending on the deadweight cost of taxation, investors may systematically avoid bids that are even lower than $\underline{R}$ (see Campbell and Mäder, 2023).

[46]In particular, this is because forgoing the auction is strictly preferred to any bid in $[R^0, \underline{R}]$ whenever $R_b > R^1$ (where investors strictly prefer capital over bonds).

[47]See Campbell and Mäder (2023) for a discussion of two kinds of auction pricing in the Calvo model..

[48]The 'bad' equilibrium is ruled out by rationality and caution as it is weakly dominated by any bid in $[R^0, R^1)$.

authors' view, given the strong nature of second-order mutual assumption of rationality and caution, 1EWDS and 2EWDS are likely more appropriate. Specifically, to err on the side of caution, we lean toward 1EWDS as our preferred concept (see Campbell and Mäder, 2023).

From an economic standpoint, the choice between 1EWDS and 2EWDS is highly consequential. Indeed, while rollover crises are a possibility under rationality and caution (1EWDS), they are ruled out as soon as rationality and caution are mutually assumed (2EWDS). In turn, from the point of view of policy, an interesting question is then whether it might be possible to rule out rollover crises *even if* rationality and caution are not mutually assumed (i.e. under 1EWDS). In Campbell and Mäder (2023), we answer this question affirmatively and, thus, further underscore the importance of considering carefully a strategic environment's prevailing epistemic state.

In summary, akin to the Diamond-Dybvig model of bank runs, we have argued that Nash equilibrium represents an epistemically unsuitable lens to study ordinary debt auctions as described by Calvo (1988). However, unlike in the case of a bank run, investors participating in such auctions realistically cannot observe each others' tentative actions. In effect, since investors must form beliefs based on their knowledge and/or assumptions about each other, we considered and weighed various levels of weak dominance.

3.3 A market for a lemon (Akerlof, 1970)

Since the seminal contribution by Akerlof (1970), it has been well-known that information asymmetries regarding a traded good's quality can lead to market collapse, but less is known about information asymmetries of the epistemic kind. For example, does the market collapse if the seller of a used car assumes that the prospective buyer is rational and cautious, but the buyer makes no such assumptions about the seller? Conversely, what if the buyer assumes that the seller is rational and cautious, but this assumption is incorrect? It is only by answering questions of this type that we can elucidate the root cause of the market's collapse.

Following the above logic, this subsection studies an Akerlof-style market for lemons through an epistemic lens. By proceeding as such, we are able to reconcile the fact that Akerlof's article, prior to acceptance, was rejected on the grounds of being both "trivial" and "too general to be true" (see Gans and Shepherd, 1994). In particular, we show that information asymmetries regarding a used car's quality are, in fact, not in and of themselves (i.e. irrespective of the chosen solution concept) sufficient to imply that the market will collapse. This serves as a powerful illustration of the broader point that economic theory does not in and of itself generate *any* predictions. Instead, it is a modeler's reading of their theory — through the lens of a solution concept, an epistemic state, or both — that generates the predictions (see Figure 1).

For purposes of illustration, let us consider a seller and a (prospective) buyer of a used car whose quality θ is distributed uniformly in $\Theta = [0, 1]$. Letting z denote an indicator of the car's sale at the price p, we assume that individual utility is given by $u_S = z[kp - \theta]$ and $u_B = z\mathbb{E}[\theta - p|p]$, where $k \geq 1$ (to reflect the seller's relatively strong preference for liquidity) and θ is known to the seller, but not to the buyer. In turn, the seller must choose a continuous, non-decreasing price mapping $p^o : \Theta \mapsto [0, 1]$ at which they wish to offer the car for sale, whereas the buyer must specify a non-increasing function $z : [0, 1] \mapsto \{0, 1\}$ such that trade occurs if and only if $z(p^o(\theta)) = 1$.

To narrow down the set of conceivable outcomes of the described game, first note that, as in the two prior applications, Nash equilibrium is contextually inappropriate because, under ordinary circumstances, there is no reason to believe that our two players can correctly anticipate each other's choices.[49] However, much like in the case of a sovereign debt auction, they very well may possess knowledge and/or be willing to make assumptions about each other:

(C1) *Assuming* the seller is rational and cautious, they will not choose to offer the car at any price $p^o(\theta) < \theta/k$.[50]

(C2) *Assuming* the buyer is rational, cautious, and aware of the true uniform prior of θ, they will reject any offer in excess of one half, i.e. $z(p^o) = 0$ for any $p^o > \frac{1}{2}$.[51]

(C3) Further *assuming* that the buyer correctly assumes (C1), they recognize that, for any actually observed offer p^o, the car's quality must satisfy $\theta \leq kp^o$. In turn, they infer $\mathbb{E}[\theta|p^o] = \mathbb{E}[\theta|\theta < kp^o] = \min\{\frac{kp^o}{2}, \frac{1}{2}\}$ and, thus, optimally choose $z(p^o) = \mathbb{1}[\min\{kp^o, 1\}/2 - p^o \geq 0]$.[52]

(C4) Further *assuming* that the seller correctly assumes (C2), they recognize that the buyer will reject any offer $p^o > \frac{1}{2}$ and, thus, only make offers satisfying $p^o(\theta) < \min\{\theta/k, \frac{1}{2}\}$.

[49] Incidentally, the set of Nash equilibria features various strategy profiles that are dubious at best (e.g. $p^o(\theta) = 1$ for any $\theta \in [0, 1]$ and $z(p^o) = 0$ for any $p^o \in [0, 1]$).

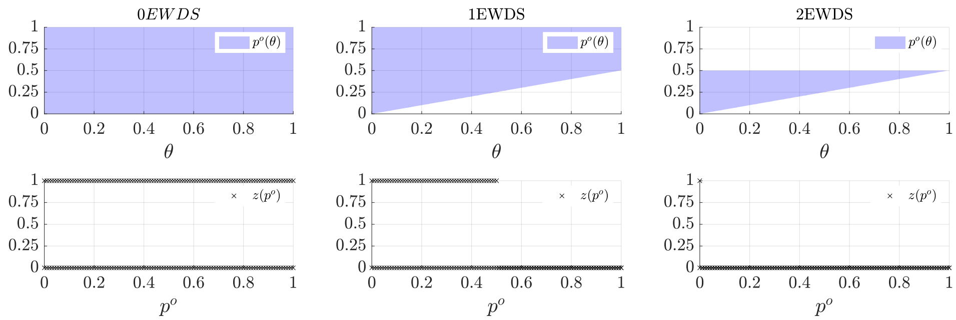
[50] Rationality is in and of itself insufficient to imply $p^o(\theta) < \theta/k$, namely because the seller might believe that the buyer will not accept any offer. In this context, caution represents a natural assumption in that it rules out such weakly dominated behavior (see Appendix B).

[51] Once again, rationality is in and of itself insufficient to imply $z(p^o) = 0$ for $p^o > \frac{1}{2}$, namely because the buyer might believe that the seller will never make such an offer (see any column of Figure 5). In this context, caution represents a natural assumption in that it rules out such weakly dominated behavior (by forcing the buyer to consider the possibility that the seller might in fact make such an offer, see Appendix B).

[52] Technically, the buyer is indifferent (between accepting and rejecting) if $u_B(p^o|z = 1) = \min\{\frac{kp^o}{2}, \frac{1}{2}\} - p^o = 0$ (see Figure 6) in which case, unless any further assumptions are made, $z(p^o) \in \{0, 1\}$ are both best responses. To avoid such indeterminacy, we assume that the buyer lexicographically prefers to buy if they are first-order indifferent.

Since (C1)-(C2) and (C1)-(C4) amount to one and two rounds of elimination of weakly dominated strategies respectively, the described strategic implications are summarized in Figure 5.

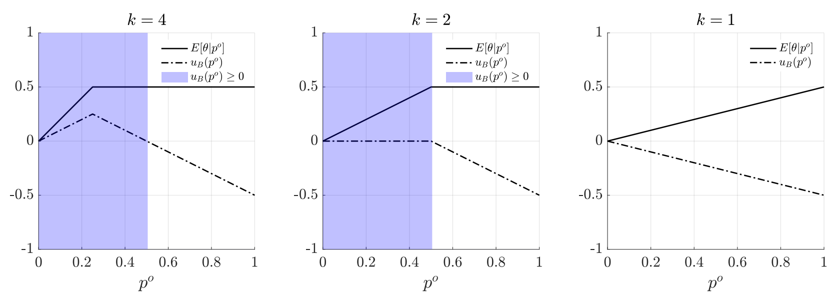
Figure 5. Iterated elimination of weakly dominated strategies (for $k = 1.99$)



Notes: Figure 5 illustrates the described game's set of conceivable outcomes under various forms of weak dominance. For example, when nothing is known about either player (i.e. 0EWDS), the set of conceivable outcomes is given by the entire strategy space. In turn, if both players are rational and cautious (i.e. 1EWDS), the seller (top row) only makes offers satisfying $p^o(\theta) \geq \theta/k$, and the buyer (bottom row) only accepts offers satisfying $p^o \leq \frac{1}{2}$. If, in addition, rationality and caution are mutually assumed (i.e 2EWDS), the buyer rejects any positive offer.

Intuitively, the reason why the market collapses under 2EWDS is that the prospective buyer leverages the seller's observed price offer to infer information about the car's quality. Specifically, as famously outlined in Akerlof (1970), a lower price offer is viewed as indicating a lower expected quality of the car. To illustrate this link, Figure 6 depicts the car's expected quality (alongside the buyer's payoff of accepting) across a range of price offers p^o.

Figure 6. Expected car quality and buyer payoff under 2EWDS



Notes: Figure 6 depicts the expected car quality as well as the prospective buyer's resulting payoff of accepting across a range of price offers p^o. In turn, since rejecting the seller's offer yields utility of zero, the blue shaded area reflects the range of offers which the buyer is willing to accept. That is, the buyer is willing to accept any offer up to and including $p^o = \frac{1}{2}$ in Panels A and B, but no (positive) price offer in Panel C.

In Figure 6, note that the prospective buyer is principally willing to buy in Panels A and B, but not in Panel C.[53] Indeed, it can be shown that the buyer are willing to accept any offer up to $p^o = \frac{1}{2}$ so long as $k \geq 2$, but no positive offer when $k < 2$ (as shown in Figure 5).

[53] The buyer is never willing to accept an offer in excess of $\frac{1}{2}$ because the car's unconditional expected value is $\frac{1}{2}$.

While the above description of market collapse is hardly novel, our precise epistemic characterization thereof is. In turn, said characterization yields two main insights.

First, while the market does collapse under 2EWDS, the same is not necessarily true under 1EWDS (see Figure 5). This is because, under 2EWDS, the buyer's reluctance to accept an offer not only reflects their rationality and caution, but also their perception of the seller. Indeed, the reason why the car's expected value is perceived to be increasing in the observed price offer is that the buyer assumes, quite reasonably, that the seller would not offer the car at a price below their own valuation.[54] Following this logic, it is tempting to conclude that rationality, caution, and mutual assumption thereof (as reflected by 2EWDS) represent the 'tightest' set of epistemic conditions that guarantee a market collapse. This is not so.

Second, by explicating and disentangling the four assumptions that collectively form 2EWDS, we can examine what would happen if a particular buyer-seller pair were most suitably described by an epistemic state that is asymmetric and/or if players made assumptions that are false. The primary benefit of adopting such a granular point of view is that it permits an even preciser, or 'tighter', description of the conditions that lead to market collapse. Perhaps unsurprisingly, whether the market collapses exclusively depends on the characteristics of the prospective buyer. Specifically, since neither disposing of (C4) nor (C1) fundamentally alters the buyer's calculus, the smallest set of assumptions to imply a market collapse is (C2) and (C3). That is, so long as the buyer is rational and cautious while assuming that the seller is as well, they will reject any offer irrespective of whether the seller is, in fact, rational, cautious, and/or what their beliefs are.

Although we have shown that rationality and caution are in and of themselves insufficient to imply a market collapse in the described model, our main point is not to call into question the prediction of such a collapse. Instead, our main point is to explicate the precise conditions that lead to market collapse. In particular, if buyer-side rationality, caution, and assumption of seller-side rationality and caution are perceived as a contextually (i.e. in the specific context of a particular buyer-seller pair) appropriate epistemic description, the prediction of a market collapse is plausible. Conversely, if the listed conditions are perceived as contextually inappropriate, the same prediction is less plausible.

[54] This assumption is correct because the seller was, in fact, assumed — in (C1) — to be both rational and cautious. If the seller were only rational, but not cautious, they very well might offer the car at a price below their own valuation, namely because they might assume that the buyer will reject their offer anyway.

3.4 Market economies (Arrow and Debreu, 1954)

For our final set of applications, in an attempt to diminish the perceived dichotomy between game theory and "classical economic theory"[55], we apply our decision-theoretic lens to various types of market economies. As part of this effort, we revisit a long-standing debate among Walrasian scholars, namely if and how one might expect such economies, competitive economies in particular, to reach 'equilibrium'.[56]

Classical competitive economies (collective price-taking)

To formally frame our discussion, we start by considering the very general, perfectly competitive framework studied by Arrow and Debreu (1954) (see Appendix C). Within this setting, the referenced authors' main contribution was to prove the existence of a competitive equilibrium as given by a price vector $p^\star$ that induces market clearing in each good's market. For this, to be able to invoke Nash equilibrium as part of their proof, Arrow and Debreu (1954) entrust the setting of prices to an auxiliary third party,

(D1) *Assuming* that prices are set by a "fictitious" third party, the Walrasian "market participant", each non-predetermined variable is chosen by a designated agent in the economy.[57] Specifically, suppose the market participant's optimization problem is given by,

$$\max_{p \in \mathbb{R}^l_{\geq 0}} \quad pz^\star \quad \text{s.t.} \quad \sum_{i=1}^{l} p_i = 1 \tag{1}$$

where $z^\star_i(p)$ denotes excess demand of good i for each market $i \in \{0, 1, ..., l\}$. The above optimization problem is constructed so as to cause prices to rise/fall when excess demand is positive/negative,

(D2) *Assuming* the Walrasian 'market participant' is unaware of the functional relationship $z^\star(p)$, but that they can observe $z^\star(p)$ following each price announcement, and that they continuously and myopically revise the prevailing price vector in the direction of each dimension's best response, we obtain the following canonical "law of supply and demand" (Arrow and Debreu, 1954),

[55] "Classical economic theory did manage to sidestep the game-theoretic aspects of economic behavior by postulating perfect competition, i.e. by assuming that every buyer and every seller is very small as compared with the size of the relevant markets, so that nobody can significantly affect [...] prices by his actions." (Harsanyi, 1995)

[56] See Walker (1987) and Walker and van Daal (2014) for a historical account of Walrasian tâtonnement and two popular interpretations thereof: a dynamic process to mimic real-world competitive markets moving *towards* equilibrium and a purely static interpretation rationalizing all observed behavior *as a manifestation of* equilibrium.

[57] Walker and van Daal (2014) describe the popular interpretation of the market participant as an auctioneer a "momentous error". Instead, this auxiliary agent is best understood, they argue, as a "crier" who simply broadcasts the prices quoted to them by the agents representing the buyers and sellers.

$$\dot{p}_h = \begin{cases} 0 & \text{if} \quad p_h = 0, z_h^\star(p) < 0 \\ H[z_h^\star(p)] & \text{otherwise} \end{cases} \tag{2}$$

where the function H with $H(0) = 0$ and $H' > 0$, as originally proposed by Samuelson (1941), can be used to parameterize the market participant's 'speed' of price adjustment. To illustrate the dynamic process induced by (2), let us examine a rudimentary one-consumer, one-producer economy. Specifically, consider a unit mass of consumers who choose labor supply so as to maximize their utility as follows,

$$n_S^\star = \underset{n_S}{\text{argmax}} \quad \frac{c^{1-\gamma} - 1}{1-\gamma} - \xi \left[\frac{n_S^{1+\frac{1}{\nu}}}{1+\frac{1}{\nu}} \right] \quad \text{s.t.} \quad c = wn_S + \pi \tag{3}$$

such that optimal labor supply $n_S^\star = \left[\frac{w^{1-\gamma}}{\xi}\right]^{\frac{\nu}{1+\gamma\nu}}$ is monotonically increasing in w so long as $\gamma < 1$. Further suppose there is a unit mass of firms who choose labor demand so as to maximize profits as follows,

$$n_D^\star = \underset{n_D}{\text{argmax}} \ \pi \quad \text{s.t.} \quad \pi = zn_D^\alpha - wn_D \tag{4}$$

such that optimal labor demand $n_D^\star = \left[\frac{\alpha z}{w}\right]^{\frac{1}{1-\alpha}}$ is monotonically decreasing in w so long as $\alpha < 1$. Thus, assuming $\alpha < 1$ and $\gamma < 1$, competitive equilibrium is unique and given by the wage that equates labor supply $n_S^\star$ and labor demand $n_S^\star$ depicted in Figure 7A. Moreover, as shown in Figure 7B, Walrasian tâtonnement asymptotically induces the competitive equilibrium $w^\star$ for initial conditions w_0 both above and below $w^\star$.[58] Indeed, this is unsurprising in that Walrasian tâtonnement represents a special instantiation of best-response dynamics, a solution concept under which any "socially stable strategy" is known to be Nash (see Gilboa and Matsui, 1991).[59]

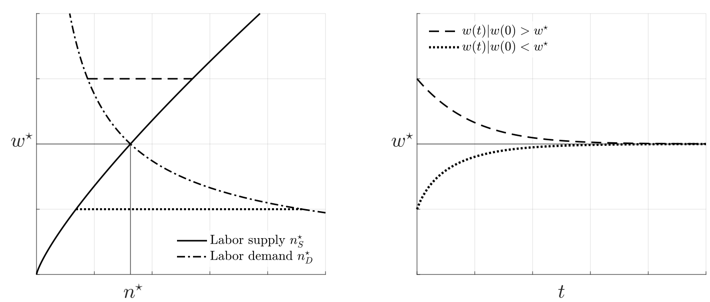
While the depicted competitive equilibrium's global stability renders it a conceptually appealing candidate for our exemplary economy's actually observed outcome, there are two factors which limit the practical relevance of Walrasian tâtonnement as a solution concept more broadly. First, many

[58] In fact, the monotonic nature of supply and demand are sufficient to imply $\lim_{t\to\infty} w_t \to w^\star$ for *any* $w_0 \geq 0$.

[59] The converse is not true such that a primitive invocation of Nash equilibrium is not only inappropriate from a decision-theoretic perspective (see Figure 2), but also because the set of Nash equilibria and the set of strategy profiles that are stable under (2) are not equivalent (i.e. some competitive equilibria are unstable). Indeed, Arrow and Debreu (1954) warn: "Neither the uniqueness nor the stability of the competitive solution is investigated in this paper".

competitive economies feature excess demand functions that are not sufficiently well-behaved to induce equilibrium that is globally stable.[60] Second, and much more importantly in our context, the epistemic assumptions that underlie Equation (2) — (D1) and (D2) — are explicitly "fictitious". Indeed, the only reason that assumptions (D1) and (D2) even qualify as epistemic, if at all, is that they were designed as conceptual placeholders in lieu of a market economy's actual epistemic setting. While this may be suitable in certain instances, namely when neither supply, nor demand get to choose the transacted price (e.g. NYSE[61]), tâtonnement is decision-theoretically inappropriate in many, if not most, other environments. But then, if not Walrasian tâtonnement, which solution concept is appropriate when prices are chosen by literal, not fictitious market participants?

Figure 7. Tâtonnement in a one-consumer, one-producer competitive economy



Notes: Panel A of Figure 7 depicts the market for labor in the described one-consumer, one-producer economy parameterized by $\gamma = \frac{1}{3}, \xi = 3, \nu = 5, \alpha = \frac{1}{3}, z = 3$. Panel B illustrates, for two separate initial conditions and $H(z) = 15z$, the process of tâtonnement induced by (2), with the Walrasian 'market participant' continuously adjusting the real wage until the labor market has cleared.

Other competitive economies (price-setting)

In many markets, Walrasian tâtonnement represents an inappropriate solution concept because prices are in fact set by actual (not fictitious) market participants. For our two final applications, we thus apply our decision-theoretic lens to the canonical market for a homogenous good and its two most well-known forms of industrial organization: Bertrand and Cournot.

Bertrand competition. In the standard Bertrand duopoly, the unique Nash equilibrium is given by both firms offering the good at a uniform price equal to marginal costs, i.e. $p_1 = p_2 = c$, where c denotes a symmetric marginal cost.[62] However, recall that Nash equilibrium is only epistemically

[60]Indeed, the same is true in pure exchange economies (see Sonnenschein, 1972; Mantel, 1974; Debreu, 1974).

[61]While the Walrasian "crier" analogy (see Walker and van Daal, 2014) exquisitely fit the New York Stock Exchange's open-outcry era, it retains its relevance under today's all-electronic trading regime.

[62]By standard Bertrand model, we mean continuous/downward sloping demand, no capacity constraints, zero fixed costs, and symmetric marginal costs c.

suitable if actual/final choices are mutually known, a fairly strong assumption in the context of Bertrand. Yet, even if Nash equilibrium is deemed to be inappropriate *as a solution concept*, there are still two epistemic arguments to be made in favor of the Bertrand-Nash *outcome*. First, if payoffs and rationality are common knowledge, the only rationalizable outcome is $p_1 = p_2 = c$ (as implied by iterated elimination of never-best responses).[63,64] In addition, and perhaps most convincingly, the latter is also the unique 'socially stable strategy' such that, if both firms can continually revise their strategy in response to the other's observed price, the resulting best-response dynamic path ultimately (i.e. asymptotically) induces $p_1 = p_2 = c$ for *any* strategic initial condition.

Cournot competition. In the standard interpretation of Cournot duopoly, firms are viewed as choosing quantities while taking prices as given, a rather artificial premise. However, Kreps and Scheinkman (1983) show that, under certain conditions, the canonical Cournot outcome can also be explained as a manifestation of Nash equilibrium *with* price-setting, namely if firms first commit to building production capacities and then, in a second stage, compete via price à la Bertrand. However, as before, Nash equilibrium only represents a suitable solution concept so long as actual/final choices are mutually known, a very strong assumption in this context. However, as before, even if Nash equilibrium is deemed to be inappropriate *as a solution concept*, there is still an epistemic argument to be made in favor of the Cournot-Nash *outcome*. Specifically, suppose it is in firms' best interest to set their price competitively so as to sell their entire inventory in the second stage *and* that this is common knowledge.[65] Then, if both firms can continually revise their production capacity in response to the other's observed capacity, the resulting best-response dynamic path ultimately (i.e. asymptotically) induces the Cournot outcome for any initial condition so long as it is globally stable.

---

[63] To obtain this result despite the game's infinite strategy space, we require a transfinite number of eliminations (see Lipman, 1994).

[64] Interestingly, the same result does not obtain under cautious rationalizability, namely because $p_i = c$ is weakly dominated by any $p_i \in (c, \bar{p})$. In effect, the set of cautiously rationalizable strategy profiles is empty, a symptom of the Bertrand model's uncountably infinite strategy space. In turn, if the price space were finite, the unique cautiously rationalizable outcome would be $p_1 = p_2 = c + \varepsilon$, the 'larger' (and more plausible) of said model's two Nash equilibria.

[65] While this is not true in general, it is true if demand is isoelastic, i.e. $p = ax^{-b}$, with $b \in (0, 1)$. In conjunction with common knowledge, this implies $p_1 = p_2 = a[x_1 + x_2]^{-b}$ in the second stage and both firms know this.

4 Discussion

On transparency (Procedure A vs. B)

For each of our applications in Section 3, we proposed an epistemic state so as to capture the specific strategic setting at hand and, in turn, directly derived a corresponding set of empirical predictions therefrom. By proceeding as such, we hope to have convinced the reader of the relative merits of Procedure C in Figure 1. However, even if the indirect path via existing solution concepts (i.e. Procedure B) remains preferred, the primary objective of Section 3 was to demonstrate the following fact: Invoking solution concepts primitively (i.e. Procedure A) effectively amounts to imposing implicit assumptions about the modeled players' rationality and/or their beliefs (see Figure 2). As such, the difference between Procedures A and B does not lie in the assumptions being made, but in the fact that all assumptions, including epistemic ones, are made explicitly under Procedure B (and C), but not under Procedure A.[66]

On versatility (Procedure B vs. C)

Aside from transparency, the main benefit of the epistemic approach lies in its versatility. Indeed, a key insight from our various applications is that, oftentimes, it is not necessary to exhaust all of a canonical solution concept's epistemic requirements to obtain the same empirical predictions. For example, in Section 3.3, we showed that 2EWDS — as induced by rationality, caution, and mutual assumption thereof via Procedure B — implies that the market for a used car must collapse, but the same is true even if the listed assumptions only apply to the buyer.[67] For example, even if the seller is not actually rational and/or cautious, the market continues to collapse so long as the buyer assumes that they are. However, since asymmetric solution concepts of the described type do not exist, it is tempting to study, or 'solve', the described game through the lens of 2EWDS, or even IEWDS. While proceeding as such is certainly possible, it effectively weakens the presented argument, namely because the epistemic conditions that underpin 2EWDS and IEWDS are stronger than the conditions that are required to imply a market collapse.[68] It is in contexts of this type that Procedure C has a distinct advantage over Procedure B.

[66] For example, by explicating the assumptions that are required to lead to sovereign default, the reader can assess more transparently whether they should, in fact, expect to observe such default (see Section 3.2).

[67] As described in Section 3.3, the minimal set of conditions that jointly imply market collapse is that the buyer is rational, cautious, and that they assume (correctly or incorrectly) that the seller is rational and cautious as well.

[68] So long as nothing is known about the buyer-seller pair, we naturally seek to know the weakest conditions that lead to a market collapse.

On Nash equilibrium

Across our five applications in Section 3, none of the proposed epistemic states mapped into Nash equilibrium as the contextually appropriate solution concept. Since Nash equilibrium requires players to know, or at least correctly anticipate, each other's actual/final choices, this is rather unsurprising. Indeed, akin to Dekel and Siniscalchi (2015), our reading of Nash equilibrium is that it is inappropriate in most real-world contexts. At the same time, it is worth noting that all predictions derived in Section 3 ultimately assumed the form of such equilibrium. To reconcile this seeming contradiction, three remarks are in order.

First, it is crucial that we distinguish between Nash equilibrium as a solution concept and strategic outcomes that are Nash. In particular, the fact that a commonly accepted outcome is Nash does not imply, even if Nash equilibrium is unique, that the relevant solution concept is in fact Nash equilibrium.[69] For example, in prisoner's dilemma (see G_2 in Appendix A), even if players' choices are mutually unknown, individual knowledge of payoffs and rationality are sufficient to deduce that the game's (unique) Nash equilibrium will obtain, namely because a single round of elimination of strictly dominated strategies implies this outcome. In effect, our cautious reading of Nash equilibrium *as a solution concept* should not be taken to mean that we advocate for a dismissal of Nash *outcomes*.

Second, since all of our derived outcomes are Nash, it is tempting to think that, even if Nash equilibrium is contextually inappropriate as a solution concept, 'equilibrium selection' à la Procedure A might still be viable. As indicated in the introduction, such inference is misguided for at least two reasons. First, as illustrated in Figure 1, the process of 'equilibrium selection' inevitably requires an assessment which equilibria are deemed to be plausible. This remains problematic because what constitutes a 'plausible' outcome is precisely what a theorist wishes to derive from their theory and, as such, should not be invoked primitively. Second, in all applications in Section 3, our goal was precisely to explicate the epistemic assumptions that are required to induce the canonical Nash outcome. In turn, if the required assumptions are perceived as plausible, said plausibility naturally extends to the resulting Nash prediction. However, if some of the required assumptions are perceived as implausible, then the reader can discern more transparently the predictions' prac-

[69] Indeed, any given outcome might feature in a large number of sets of solutions (each pertaining to a solution concept), but this does not imply that each such concept is contextually appropriate. However, if a commonly accepted outcome does not feature in a solution concept's set of solutions, the latter is likely contextually inappropriate.

tical limitations. For example, in Section 3.3, we showed that buyer-side rationality, caution, and assumption of seller-side rationality and caution are sufficient to imply market collapse. Thus, if there is reason to believe, in the context of a specific buyer-seller pair, that the buyer does *not* assume that the seller is rational and cautious, the occurrence of a sale *cannot* be ruled out.

Third, the reason why all predictions in Section 3 were Nash is because, in each application, the proposed epistemic state either implied an evolutionary convergence towards such equilibrium or it implied a 'self-evident way of play'.[70] Thus, across all of our applications, Nash equilibrium was not assumed. Instead, it was, much like in prisoner's dilemma, a natural consequence of the proposed epistemic state.

On necessary vs. sufficient conditions

Once solution concepts are recognized as reflecting an underlying epistemic state, a frequently resurfacing point of interest lies in the distinction between necessary and sufficient conditions. For example, since Aumann and Brandenburger (1995) provide sufficient conditions for Nash equilibrium, it is tempting to ask which conditions are necessary. However, in the context of solution concepts, identifying necessary conditions adds little value because they are designed to reveal circumstances under which it is unequivocally inappropriate to invoke a solution concept, not ones under which it is, in fact, appropriate. Indeed, even the most basic building block of any solution concept — rationality — is not actually a necessary condition.[71] For example, as pointed out by Aumann and Brandenburger (1995), even in the absence of rationality, players can always "stumble" into a Nash outcome by accident, a fact that is surely insufficient to warrant restricting our attention to such outcomes (by invoking Nash equilibrium) for purposes of prediction .

On infinite strategy spaces

Although most solution concepts themselves naturally extend to infinite strategy spaces, there are two main obstacles associated with their application in such contexts. First, epistemic states may not map as neatly into existing solution concepts under infinite strategy spaces.[72] Second, various finite-game results linking solution concepts to their corresponding set of solutions may

[70] That is, the proposed epistemic state was sufficient for each player to *infer* which singular strategy others would ultimately choose. In this case, rationality immediately implies that the outcome must be Nash (recall Result 1).

[71] Of course, the absence of rationality is not the same as players systematically avoiding best responses.

[72] For example, in games with infinite strategy spaces, the limit of nth order of mutual knowledge of rationality need not converge to common knowledge of rationality as required by rationalizability (see Lipman, 1994).

not extend to infinite strategy spaces.[73] That is, since both the mapping from epistemic states to solution concepts and the mapping from solution concepts to solutions may break down under infinite strategy spaces, Procedure B shown in Figure 1 may fail.

It is tempting to interpret the fact that Procedure B may fail under infinite strategy spaces as a limitation of the epistemic approach. For two reasons, the opposite is true. First, when studying an economic model through the lens of a solution concept, an elevated understanding of the latter's epistemic requirements, even if negative, can only strengthen our analysis. For example, the fact that, in games with infinite strategy spaces, the limit of nth order of mutual knowledge of rationality need not converge to common knowledge of rationality (see Lipman, 1994) as required by both rationalizability and cautious rationalizability may, and probably should, motivate a theorist to consider alternate solution concepts when studying the canonical Bertrand model. Second, a key strength of the epistemic approach is precisely that it can operate independently of existing solution concepts (i.e. Procedure C). Thus, even if a game's infinite strategy space causes the the two referenced mappings to break down, our ultimate object of interest — the predictions — can still always be derived via the direct mapping from the specified epistemic state.

Final thought

A natural reservation regarding the practical implementation of the epistemic approach is that, for many economic models, it may be challenging to specify an epistemic state that is both plausible *and* conducive to generating sharp predictions. This very well may be true, but the encountering of such obstacles represents an intended functionality, not an unintended flaw. Indeed, recall that the difference between Procedures A and B does not lie in the assumptions being made. Instead, the difference lies in the fact that all assumptions, including epistemic ones, must be disclosed under Procedure B, but not under Procedure A.

5 Conclusion

If used for purposes of prediction, a defining element of 'good' economic theory is that any predictions derived therefrom are not chiefly driven by assumptions that are implausible, implicit, or both. Following this logic, this paper proposes that applied theory, rather than primitively

[73] For example, in games with infinite strategy spaces, Nash equilibrium may not exist.

invoking solution concepts (i.e. Procedure A), derive its predictions from a primitive epistemic state instead (i.e. Procedures B and C). Given its high level of transparency and versatility, we hope to have convinced the reader of the relative merits of the epistemic approach and, ultimately, that they will consider its adoption as part of their future research.

A Two exemplary games

Figure 8. G_1: A coordination problem

	β_1	β_2
α_1	$(1, 1)$	$(0, 0)$
α_2	$(0, 0)$	$(0, 0)$

Notes: Due to Myerson (1978), the game depicted in Figure 8 features two rationalizable strategy profiles that are also Nash — (α_1, β_1) and (α_2, β_2) — but only one plausible outcome as given by (α_1, β_1). Indeed, as long as players cannot rule out that the other will choose β_1/α_1, caution dictates that they will play α_1 and β_1 (see Pearce, 1984).

Figure 9. G_2: Prisoner's dilemma

	β_1	β_2
α_1	$(-1, -1)$	$(-3, 0)$
α_2	$(0, -3)$	$(-2, -2)$

Notes: Figure 9 depicts the canonical game of prisoner's dilemma. Since both players have a strictly dominant strategy, they need not concern themselves with the other's play as a single round of elimination of strictly dominated strategies immediately yields (α_2, β_2). In effect, individual rationality is in and of itself sufficient to guarantee the outcome (α_2, β_2).

B Assumption and lexicographic probability systems

In Section 2, in our construction of Figure 2, we made frequent and extensive use of the notion of 'knowledge' so as to rule out various rounds of opponents' never-best responses. In contrast, to motivate the elimination of weakly dominated strategies in our practical applications in Section 3, we instead appealed to the more nuanced notion of (BFK-)assumption.

Assumption. An event E is said to be *(BFK-)assumed* if each element in E is perceived as infinitely more likely than each element in not-E (see Brandenburger, Friedenberg, and Keisler, 2008).

Unlike knowledge, assumption allows for events to be considered with positive probability even though they are first-order excluded and, as such, allows an elimination of *multiple* rounds of weakly dominated strategies.[74] To formalize idea, the literature typically appeals to the proposition that players' beliefs assume the form of a lexicographic probability systems (LPS).

Lexicographic probability system. A lexicographic probability system (LPS) is a finite, lexicographically ordered sequence of probability measures $\sigma \equiv (\phi^0, \phi^1, ..., \phi^k)$ for some finite $k \in \mathbb{N}$ (see Blume et al., 1991a).

Indeed, in a seminal contribution, Brandenburger, Friedenberg, and Keisler (2008) (BFK) combine LPS with lexicographic utility and the LPS-based notion of 'assumption' in hopes of providing an epistemic characterization of iterated admissibility (IA), i.e. iterated elimination of weakly dominated strategies. Technically falling just short of their objective, BFK inspired an active research program which continues to seek epistemic characterizations of IA (see Yang, 2015; Lee, 2016; Dekel et al., 2016; Catonini and DeVito, 2021; Keisler and Lee, 2023). In this context, the literature generally distinguishes between the implicit approach, whereby belief hierarchies are modeled indirectly via 'type structures', and the explicit approach, whereby belief hierarchies are modeled directly. For purposes of illustration, we adopt the explicit approach and, thus, proceed by borrowing the notation from Lee (2016). In particular, we use X denote the set of states of the world and define inductively as follows,

$$X_i^0 \equiv S_i \qquad X_i^1 \equiv X_i^0 \times \mathcal{N}(X_{-i}^0) \tag{5}$$

where we assume the set of players I to be finite and the strategy space S_i to be Polish for each $i \in I$.[75] In turn, for each $n \geq 2$, we let,

$$X_i^n \equiv \{(x_i^{n-1}, h_i^n) \in X_i^{n-1} \times \mathcal{N}(X_{-i}^{n-1}) \mid \text{marg}_{X_{-i}^{n-2}} h_i^n = h_i^{n-1} \ \} \tag{6}$$

such that, by construction, any resulting sequence of LPS $h_i \equiv \{h_i^j\}_{j \geq 1}$ lies in the set of coherent

[74] A single round of elimination of weakly dominated strategies is often motivated by way of the assumption that beliefs have full-support. For example, Pearce (1984) proposed that *after* all never-best responses have been deleted, an elimination of all weakly dominated strategies in the remaining 'rationalizable set' is warranted so long as players are cautious in that they weigh all "likely" events with positive probability. Conversely, Brandenburger (1992a) and Börgers (1994) motivate a single elimination of weakly dominated strategies *before* iterated elimination of strictly dominated strategies, a solution concept originally due to Dekel and Fudenberg (1990), by considering the case in which rationality is "common first-order knowledge" and "approximate common knowledge" respectively.

[75] Aside from defining coherency in a non-standard fashion, Lee (2016) also assumes that the strategy space is finite. Thus, while their notation naturally extends to our more general setting, their results do not.

belief hierarchies H_i^1, i.e. $h_i \in H_i^1 \subset H_i^0 \equiv \prod_{k\geq 0} \mathcal{N}(X_{-i}^k)$.[76] In effect, player i's state space is then given by,

$$X_i \equiv \{(x_i^j)_{j\geq 0} \in \prod_{k\geq 0} X_i^k \mid x_i^j = \mathrm{proj}_{X_i^j} x_i^{j+1} \ \forall j \geq 0\} \tag{7}$$

such that each element in X_i features a strategy and a corresponding coherent hierarchy of beliefs, i.e. $x_i = (s_i, h_i)$, where $s_i \in S_i$ and $h_i \in H_i^1$ for each $x_i \in X_i$. Across agents, we then say that a state $(s,h) \in X \equiv \times_{i\in I} X_i$ satisfies *BFK-rationality* if $\mathrm{proj}_{X^1}(s,h) \in \times_{i\in I} BFK_i^1$, where

$$BFK_i^1 \equiv \{(s_i, h_i^1) \in X_i^1 \mid h_i^1 \in \mathcal{N}^+(S_{-i}) \wedge s_i \in \mathrm{BR}_i(h_i^1)\ \} \tag{8}$$

for each i. In turn, for each $m \geq 1$, we say that a state $(s,h) \in X$ satisfies *BFK-rationality and mth order assumption of BFK-rationality* if $\mathrm{proj}_{X^{m+1}}(s,h) \in \times_{i\in I} BFK^{m+1}$, where for each i,

$$BFK_i^{m+1} \equiv \{(s_i, \{h_i^j\}_{j=1}^{m+1}) \in X_i^{m+1} \mid (s_i, \{h_i^j\}_{j=1}^{m+1}) \in BFK_i^m \times \mathrm{A}(BFK_{-i}^m, X_{-i}^m)\ \} \tag{9}$$

where $\mathrm{A}(E,F)$ denotes the set of LPS on F under which the event $E \subseteq F$ is (BFK-)assumed (see Lee, 2016). We are then ready to apply the above equations to our model economies from Sections 3.1. and 3.2.

Calvo's sovereign debt auction

In our Calvo application in Section 3.2, we had $R_{bi} \in \mathbb{R}_{\geq 0}$ and, thus, let X_i be determined by (5)-(7) given $X_i^0 = S_i \equiv \mathbb{R}_{\geq 0}$ for each i. In turn, we can use (8) and (9) to narrow down the relevant states of the world as follows,

1. BFK-rationality: *Assuming* x satisfies R0AR, (8) dictates $h_i^1 \in \mathcal{N}^+(S_{-i})$ for each i. That is, player i's first-order beliefs h_i^1 must assign a positive probability to each $s_{-i} \in \mathbb{R}_{\geq 0}^{n-1}$ at some lexicographic level. In turn, (8) also dictates that $s_i \in S_i^1 \equiv \{0\} \cup [R_b^0, R_b^1)$ for each i, namely because there exists no permissible LPS for which any act in $S_i \setminus S_i^1$ maximizes lexicographic expected utility (LEU).[77]
2. BFK-rationality and first-order assumption thereof: *Assuming* x satisfies R1AR, each player BFK-assumes others to be BFK-rational. That is, any state satisfying $\mathrm{proj}_{X_{-i}^1} x \in BFK_{-i}^1$ is perceived as infinitely more likely than any state outside this set. In effect, (9) implies that player i's first-order beliefs h_i^1 must feature $s_{-i} \in S_{-i}^1$ at a lower lexicographic level than $s_{-i} \in S_{-i}$. In turn, (8) dictates that $s_i \in S_i^2 \equiv [R_b^0, R_b^1)$ for each i, namely because there exists no permissible LPS for which any act in $S_i \setminus S_i^2$ maximizes LEU.
3. BFK-rationality and second-order assumption thereof: *Assuming* x satisfies R2AR, each player

[76]Given a LPS $\sigma = (\phi^0, \phi^1, ..., \phi^k)$, we define $\mathrm{marg}_X \sigma \equiv (\mathrm{marg}_X \phi^0, \mathrm{marg}_X \phi^1, ..., \mathrm{marg}_X \phi^k)$.
[77]Recall that $R_{bi} = 0$ was used to represent the option to not submit a bid.

BFK-assumes others to BFK-assume that others are BFK-rational. That is, any state satisfying $\mathrm{proj}_{X^2_{-i}} x \in BFK^2_{-i}$ is perceived as infinitely more likely than any state outside this set. In effect, (9) implies that player i's first-order beliefs h^1_i must feature $s_{-i} \in S^2_{-i}$ at a lower lexicographic level than $s_{-i} \in S^1_{-i}$. Finally, (8) dictates that $s_i \in S^3_i \equiv \{R^0_b\}$ for each i, namely because there exists no permissible LPS for which any act in $S_i \setminus S^3_i$ maximizes LEU.

Akerlof's market for a used car

In our Akerlof application in Section 3.3, the seller was tasked to choose any continuous, non-decreasing function $p^o : [0,1] \mapsto [0,1]$, whereas the buyer was tasked to choose any non-increasing function $z : [0,1] \mapsto \{0,1\}$. In the following, let S_s be the set of all continuous, non-decreasing functions from $[0,1]$ to $[0,1]$ and S_b the set of all non-increasing functions from $[0,1]$ to $\{0,1\}$ respectively. We thus let X_i be determined by (5)-(7) given $X^0_i = S_i$ for both $i \in \{s,b\}$ and, in turn, use (8) and (9) to narrow down the relevant states of the world as follows,

1. BFK-rationality: *Assuming* x satisfies R0AR, (8) dictates $h^1_i \in \mathcal{N}^+(S_{-i})$ for both $i \in \{s,b\}$. That is, both the seller and buyer's first-order beliefs h^1_i must assign a positive probability to each $s_{-i} \in S_{-i}$ at some lexicographic level.
   (a) *Seller*: The seller assigns a positive probability to each $z \in S_b$ and, thus, optimally chooses $p^o(\theta) < \theta/k$, namely to avoid having to sell at such at too 'low' of a price in case the buyer were to accept. In effect, (8) dictates that $p^o \in S^1_s$, where S^1_s is used to denote the set of all continuous, non-decreasing functions from $[0,1]$ to $[0,1]$ that satisfy $f(\theta) \geq \theta/k$ for each $\theta \in [0,1]$, where k is taken as given.
   (b) *Buyer*: The buyer assigns a positive probability to each $p^o \in S_s$ and, thus, optimally chooses to reject any offer in excess of one half, namely to avoid having to buy the car at such a price in case the seller were to offer. In effect, (8) dictates that $z \in S^1_b$, where S^1_b is used to denote the set of all non-increasing functions from $[0,1]$ to $\{0,1\}$ with $f(p^o) = 0$ for any $p^o > \frac{1}{2}$.
2. BFK-rationality and first-order assumption thereof: *Assuming* x satisfies R1AR, both players BFK-assume the other to be BFK-rational. That is, any state satisfying $\mathrm{proj}_{X^1_{-i}} x \in BFK^1_{-i}$ is perceived as infinitely more likely than any state outside this set. In effect, (9) implies that both player's first-order beliefs h^1_i must feature $s_{-i} \in S^1_{-i}$ at a lower lexicographic level than $s_{-i} \in S_{-i}$.
   (a) *Seller*: The seller perceives the event that the buyer rejects any offer in excess of one half, i.e. $z \in S^1_b$, as infinitely more likely than not and, thus, optimally chooses not to make any such offer. That is, $p^o \in S^2_s$, where S^2_s is used to denote the set of all continuous, non-decreasing functions from $[0,1]$ to $[0,1]$ that satisfy $f(\theta) \in [\theta/k, \frac{1}{2}]$ for each $\theta \in [0,1]$, where k is taken as given.
   (b) *Buyer*: The buyer perceives the event that the seller chooses $p^o \in S^1_s$ as infinitely more likely than not and, thus, as long as $k \in [1,2)$, optimally chooses not to accept any positive price offers, namely because $\mathbb{E}[\theta - p^o | \theta \leq kp^o] < 0$ for any $p^o > 0$. That is, $z \in S^2_b$, where S^2_b is used to denote the singleton consisting of the function $f(p^o) = 0$ for any $p^o \in (0,1]$.

C Arrow and Debreu

Consider a set of n firms, each of which has a closed and convex set of production plans $Y_j \subset \mathbb{R}^l$, where l is the number of goods in the economy and $y_{jh} < 0$ is used to designate the hth good as an input. In turn, $Y = \sum_{j=1}^{n} Y_j$ denotes the set of all possible production plans in the corporate sector. Assuming that firms maximize profits, we have,

$$y_j^\star \equiv \arg\max_{y_j \in Y_j} \quad py_j \tag{10}$$

where p is taken as given by each j. In addition, consider m households, each of which has a continuous, concave, and locally non-satiable utility function $u_i : X_i \mapsto \mathbb{R}$ defined on a closed and convex set of consumption bundles $X_i \subset \mathbb{R}^l$. Further suppose that each household is endowed with a vector of goods $\zeta_i \in \mathbb{R}^l$ (such that there exists $x_i \in X_i$ with $x_i < \zeta_i$) and a vector of firm shares α_i (such that $\sum_{i=1}^{m} \alpha_{ij} = 1$ for each j). Finally, assuming that households maximize utility, we have,

$$x_i^\star \equiv \arg\max_{x_i \in X_i} \quad u(x_i) \quad \text{s.t.} \quad px_i \leq p\zeta_i + \sum_{j=1}^{n} \alpha_{ij} py_j \tag{11}$$

where p is, once again, taken as given by each i.

References

Akerlof, George A. 1970. “The Market for “Lemons”: Quality Uncertainty and the Market Mechanism.” *Quarterly Journal of Economics* 84 (3):488–500.

Altinoglu, Levent and Joseph Stiglitz. 2023. “Collective Moral Hazard and the Interbank Market.” *American Economic Journal: Macroeconomics* 15 (2):35–64.

Arrow, Kenneth J. and Gerard Debreu. 1954. “Existence of an Equilibrium for a Competitive Economy.” *Econometrica* 22(3): 265-290.

Auclert, Adrien and Matthew Rognlie. 2016. “Unique Equilibrium in the Eaton-Gersovitz Model of Sovereign Debt.” *Journal of Monetary Economics* 84:134–146.

Aumann, Robert J. 1974. “Subjectivity and Correlation in Randomized Strategies.” *Journal of Mathematical Economics* 1 (1):67–96.

———. 1976. “Agreeing to Disagree.” *The Annals of Statistics* 4 (6):1236–1239.

———. 1985. *Frontiers of Economics*. Oxford: Basil Blackwell.

———. 1987. “Correlated Equilibrium as an Expression of Bayesian Rationality.” *Econometrica* 55 (1):1–18.

Aumann, Robert J. and Adam Brandenburger. 1995. “Epistemic Conditions for Nash Equilibrium.” *Econometrica* 63 (5):1161–1180.

Bernheim, Douglas B. 1984. "Rationalizable Strategic Behavior." *Econometrica* 52 (4):1007–1028.

Bertrand, Joseph. 1883. "Review of 'Théorie mathématique de la richesse sociale' and 'Recherches sur les principes mathématiques de la théorie des richesses'." *Journal de Savants* 67:499–508.

Blume, Lawrence, Adam Brandenburger, and Eddie Dekel. 1991a. "Lexicographic Probabilities and Choice under Uncertainty." *Econometrica* 59 (1):61–79.

———. 1991b. "Lexicographic Probabilities and Equilibrium Refinements." *Econometrica* 59 (1):81–98.

Börgers, Tilman. 1994. "Weak Dominance and Approximate Common Knowledge." *Journal of Economic Theory* 64 (2):265–276.

Brandenburger, Adam. 1992a. *Economic Analysis of Markets and Games*, chap. Lexicographic Probabilities and Iterated Admissibility. MIT Press, 282–290.

———. 1992b. "Knowledge and Equilibrium in Games." *Journal of Economic Perspectives* 6 (4):83–101.

Brandenburger, Adam, Amanda Friedenberg, and Jerome H. Keisler. 2020. "Two Approaches to Iterated Reasoning in Games." *Working Paper* .

Brandenburger, Adam, Amanda Friedenberg, and Terri Kneeland. 2008. "Admissibility in Games." *Econometrica* 76 (2):307–352.

Budish, Eric and Aditya Bhave. 2023. "Primary-Market Auctions for Event Tickets: Eliminating the Rents of "Bob the Broker?" *American Economic Journal: Microeconomics* 15 (1):142–170.

Calvo, Guillermo A. 1988. "Servicing the Public Debt: The Role of Expectations." *American Economic Review* 78 (4):647–661.

Campbell, Jason and Nicolas Mäder. 2023. "Rollover Crises in the Calvo Model of Sovereign Default (And How to Prevent Them)." *Working Paper* .

Carlsson, Hans and Eric Van Damme. 1993. "Global Games and Equilibrium Selection." *Econometrica* 61 (5):989–1018.

Catonini, Emiliano and Nicodemo De Vito. 2021. "Cautious Belief and Iterated Admissibility." *Working Paper* .

Cole, Harold L. and Timothy J. Kehoe. 2000. "Self-Fulfilling Debt Crises." *The Review of Economics Studies* 67 (1):91–116.

Cookson, Anthony J., Corbin Fox, Javier Gil-Bazo, Juan Felipe Imbet, and Christoph Schiller. 2023. "Social Media as a Bank Run Catalyst." *SSRN Working Paper* .

Cournot, Augustin. 1838. *Recherches sur les principes mathématiques de la théorie des richesses*. L. Hachette.

Debreu, Gerard. 1974. "Excess Demand Functions." *Journal of Mathematical Economics* 1:15–21.

Dekel, Eddie, Amanda Friedenberg, and Marciano Siniscalchi. 2016. "Lexicographic beliefs and assumption." *Journal of Economic Theory* 163:955–985.

Dekel, Eddie and Drew Fudenberg. 1990. “Rational Behavior and Payoff Uncertainty.” *Journal of Economic Theory* 52:243–267.

Dekel, Eddie and Marciano Siniscalchi. 2015. *Epistemic game theory*. Elsevier.

Diamond, Douglas W. and Philip H. Dybvig. 1983. “Bank Runs, Deposit Insurance, and Liquidity.” *Journal of Political Economy* 91 (3):401–419.

Diamond, Peter. 1987. *Advances in Economic Theory: Fifth World Congress*, chap. Equilibrium without an auctioneer. Cambridge University Press.

Eaton, Jonathan and Mark Gersovitz. 1981. “Debt with Potential Repudiation: Theoretical and Empirical Analysis.” *Review of Economic Studies* 48 (2):289–309.

Gans, Joshua S. and George B. Shepherd. 1994. “How Are the Mighty Fallen: Rejected Classic Articles by Leading Economists.” *Journal of Economic Perspectives* 8 (1):165–179.

Gilboa, Itzhak and Akihiko Matsui. 1991. “Social Stability and Equilibrium.” *Econometrica* 59 (3):859–867.

Harsanyi, John C. 1973. “Games with Randomly Disturbed Payoffs: A New Rationale for Mixed-Strategy Equilibrium Points.” *International Journal of Game Theory* 2:1–23.

———. 1995. “Games with Incomplete Information.” *American Economic Review* 85 (3):291–303.

Harsanyi, Selten Reinhard, John C. 1988. *A General Theory of Equilibrium Selection in Games*. MIT Press.

Hu, Peicong and Joel Sobel. 2022. “Getting Permission.” *AER: Insights* 4 (4):459–472.

Kamada, Yuichiro and Michihiro Kandori. 2020. “Revision games.” *Econometrica* 88 (4):1599–1630.

Keisler, Jerome and Byung Soo Lee. 2023. “Common Assumption of Rationality.” *SSRN Working Paper* .

Kreps, David M. and Jose A. Scheinkman. 1983. “Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes.” *The Bell Journal of Economics* 14 (2):326–337.

Kreps, David M. and Robert Wilson. 1982. “Sequential Equilibria.” *Econometrica* 50 (4):863–894.

Lee, Byung Soo. 2016. “Admissibility and assumption.” *Journal of Economic Theory* 163:42–72.

Lipman, Barton L. 1994. “A Note on the Implications of Common Knowledge of Rationality.” *Games and Economic Behavior* 6:114–129.

Liu, Qi and Bo Sun. 2023. “Relative Wealth Concerns, Executive Compensation, and Managerial Risk-Taking.” *American Economic Journal: Microeconomics* 15 (2):568–598.

Mäder, Nicolas. 2024. “Financial Crises as a Phenomenon of Multiple Equilibria and How to Select Among Them.” *Journal of Money, Credit, and Banking* 56 (2-3):517–536.

Mantel, Rolf. 1974. “On the characterization of aggregate excess-demand.” *Journal of Economic Theory* 12 (2):197–201.

Matsui, Akihiko. 1992. "Best response dynamics and socially stable strategies." *Journal of Economic Theory* 57:343–362.

Myerson, Roger B. 1978. "Refinement of the Nash Equilibrium Concept." *International Journal of Game Theory* 7:73–80.

Nash, John. 1950. "Equilibrium points in n-person games." *Proceedings of the National Academy of Sciences* 36 (1):48–49.

Oyama, Daisuke, William H. Sandholm, and Olivier Tercieux. 2015. "Sampling best response dynamics and deterministic equilibrium selection." *Theoretical Economics* 10 (1):243–281.

Pearce, David G. 1984. "Rationalizable Strategic Behavio and the Problem of Perfection." *Econometrica* 52 (4):1029–1050.

Peck, James and Karl Shell. 2003. "Equilibrium bank runs." *Journal of Political Economy* 111 (1):103–123.

Samuelson, Larry. 2016. "Game Theory in Economics and Beyond." *Journal of Economic Perspectives* 30 (4):107–130.

Samuelson, Paul A. 1941. "The Stability of Equilibrium: Comparative Statics and Dynamics." *Econometrica* 9 (2):97–120.

Selten, Reinhard. 1975. "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games." *International Journal of Game Theory'* 4 (1):25–55.

Sonnenschein, Hugo. 1972. "Market Excess Demand Functions." *Econometrica* 40 (3):549–563.

Thaler, Richard H. 2015. *Misbehaving: The making of behavioral economics.* W. W. Norton & Co.

van Damme, Eric. 1984. "A Relation between Perfect Equilibria in Extensive Form Games and Proper Equilibria in Normal Form Games." *International Journal of Game Theory* 13 (1):1–13.

———. 1991. *Stability and Perfection of Nash equilibria.* Springer, 2nd. ed.

Walker, Donald A. 1987. "Walras' Theory of Tatonnement." *Journal of Political Economy* 95 (4):758–774.

Walker, Donald A. and Jan. van Daal. 2014. *Léon Walras, Elements of Theoretical Economics.* Cambridge University Press.

Yang, Chih-Chun. 2015. "Weak assumption and iterative admissibility." *Journal of Economic Theory* 158:87–101.